

# **Immersive Gestaltung für Hörspiel- produktionen in mehrdimensionalen Audioformaten**

## **Bachelor-Thesis**

zur Erlangung des wissenschaftlichen Grades  
Bachelor of Science (B.Sc.)

**Adrian Immanuel Baron**



Erstprüfer: Prof. Thomas Görne

Zweitprüfer: Stefan Troschka



Hochschule für Angewandte Wissenschaften Hamburg

Fakultät Design, Medien und Information

Department Medientechnik

Hamburg, den 27.07.2022



## Abstract

Task and object of investigation of this thesis was the sound design of a radio drama in a full-spherical loudspeaker system with the aim of achieving the effect of immersion. Since this is a system that is still hardly used and I am the first person at the Hamburg University of Applied Sciences to have made a production in such a setup, it was also necessary to discuss fundamental questions and make statements that are relevant to working in such a setup. Specifically, I designed and mixed two scenes from the radio play “Die drei ??? und das Dorf der Teufel” (The Three ??? and the Devil's Village) in a fully spherical Ambisonics system, which I defined as True Spatial. In order to be able to make specific statements about the production with this system and the perception of the sounds and tones produced, I also designed and mixed both scenes in 7.1.4. My work suggests that an extension to include the lower hemisphere is not necessary for immersion to be experienced as an effect, but that this extension nevertheless allows for an easier transition into the virtual world and promotes the feeling of being enveloped by this world. It also provides greater creative and artistic freedom and allows for new artificial room acoustics that can be used for true-to-life design and narration. The thesis suggests that for effective and purposeful use of full-spherical systems, further research is needed on lower acoustic perception as well as on the creative use of the lower hemisphere.

## Zusammenfassung

Aufgabe und Untersuchungsgegenstand dieser Arbeit war die klangliche Gestaltung eines Hörspiels in einem vollsphärischen Lautsprechersystem mit dem Ziel, den Effekt der Immersion zu erzielen. Da es sich um ein noch kaum verbreitetes System handelt und ich der Erste an der Hochschule für Angewandte Wissenschaften Hamburg bin, der eine Produktion in einem solchen Setup angefertigt hat, galt es außerdem grundsätzliche Fragen zu erörtern und Aussagen zu treffen, die für die Arbeit in einem solchen Setup relevant sind. Konkret habe ich zwei Szenen aus dem Hörspiel *Die drei ??? und das Dorf der Teufel* in einem vollsphärischen und von mir als True Spatial definierten Ambisonics-System vertont und gemischt. Um spezifische Aussagen zur Produktion mit diesem System und zur Wahrnehmung der damit erzeugten Geräusche und Töne treffen zu können, habe ich beide Szenen zusätzlich in 7.1.4 vertont und gemischt. Meine Arbeit lässt darauf schließen, dass eine Erweiterung um die untere Hemisphäre nicht zwingend notwendig ist, um Immersion als

Effekt erlebbar zu machen, dass diese Erweiterung aber dennoch einen leichteren Übergang in die virtuelle Welt ermöglicht und das Gefühl fördert, von dieser Welt umhüllt zu sein. Außerdem bietet sie einen größeren kreativen und künstlerischen Freiraum und ermöglicht eine neue künstliche Raumakustik, die sich für wirklichkeitsgetreue Gestaltung und die Narration nutzen lässt. Die Arbeit legt nahe, dass für eine effektive und zielgerichtete Nutzung vollsphärischer Systeme weitere Forschungen zur unteren akustischen Wahrnehmung ebenso nötig sind wie zum kreativen Einsatz der unteren Hemisphäre.

---

# Inhaltsverzeichnis

<b>Abstract</b> .....	3
<b>Zusammenfassung</b> .....	3
<b>1 Hörspiel: Definition</b> .....	15
<b>2 Zur Geschichte des (deutschen) Hörspiels</b> .....	16
<b>2.1 Beginn des deutschen Rundfunks</b> .....	16
<b>2.2 A Comedy of Danger und Maremoto</b> .....	17
<b>2.3 Das erste deutsche Hörspiel</b> .....	17
<b>2.4 Hörspielproduktionen der 1920er Jahre</b> .....	18
<b>2.5 Aufstieg des Hörspiels</b> .....	19
<b>2.6 Das Dritte Reich – nationalsozialistische Hörspiele</b> .....	20
<b>2.7 Die zweite Blütezeit</b> .....	21
<b>2.8 Einzug des Fernsehens</b> .....	22
<b>2.9 Das neue Hörspiel</b> .....	22
<b>2.10 Stereo als technische Innovation</b> .....	23
<b>2.11 Kunstkopf-Hörspiele</b> .....	24
<b>2.12 Digitalisierung und das Internet als Plattform</b> .....	25
<b>2.13 Bedeutung des heutigen Hörspiels</b> .....	26
<b>2.14 Die drei ???</b> .....	26
<b>3 Wiedergabeformate/Geschichte der mehrkanaligen Audiosysteme</b> .....	28
<b>3.1 Théâtrophone</b> .....	28
<b>3.2 Monofonie</b> .....	29
<b>3.3 Binaural/Kopfbezogene Stereophonie</b> .....	29

---

<b>3.4</b>	<b>Stereofonie</b> .....	<b>31</b>
<b>3.5</b>	<b>Über Stereo hinaus</b> .....	<b>32</b>
<b>3.6</b>	<b>Quadrofonie</b> .....	<b>33</b>
<b>3.7</b>	<b>Surround</b> .....	<b>36</b>
<b>3.8</b>	<b>5.1 Surround</b> .....	<b>37</b>
<b>3.9</b>	<b>Surround mit Höhen</b> .....	<b>38</b>
<b>3.10</b>	<b>NHK 22.2</b> .....	<b>39</b>
<b>3.11</b>	<b>10.2</b> .....	<b>40</b>
<b>3.12</b>	<b>Ambisonics</b> .....	<b>41</b>
<b>3.13</b>	<b>True Spatial</b> .....	<b>41</b>
<b>4</b>	<b>Räumliche Wahrnehmung</b> .....	<b>42</b>
<b>4.1</b>	<b>Richtungswahrnehmung</b> .....	<b>42</b>
<b>4.2</b>	<b>Horizontale Ebene</b> .....	<b>43</b>
<b>4.3</b>	<b>Interaurale Laufzeitdifferenz (ITD)</b> .....	<b>44</b>
<b>4.4</b>	<b>Interaurale Pegeldifferenz</b> .....	<b>45</b>
<b>4.5</b>	<b>Zusammenspiel aus interauralen Laufzeit- und Pegeldifferenzen</b> .....	<b>46</b>
<b>4.6</b>	<b>Vertikale Ebene</b> .....	<b>46</b>
<b>4.7</b>	<b>Wahrnehmung der Entfernung</b> .....	<b>49</b>
<b>4.8</b>	<b>Präzedenzeffekt</b> .....	<b>50</b>
<b>4.9</b>	<b>Räumlichkeit und Umhüllung</b> .....	<b>51</b>
<b>4.10</b>	<b>Einflüsse</b> .....	<b>52</b>
<b>5</b>	<b>Immersion</b> .....	<b>53</b>
<b>5.1</b>	<b>Immersion als psychologischer Zustand</b> .....	<b>53</b>
<b>5.1.1</b>	<b>Multisensorische Stimulation</b> .....	<b>54</b>
<b>5.1.2</b>	<b>Narrative Immersion</b> .....	<b>55</b>
<b>5.2</b>	<b>Immersion als objektive Eigenschaft eines Systems/Technologie</b> .....	<b>55</b>

---

<b>6</b>	<b>Sounddesign</b> .....	57
<b>6.1</b>	<b>Auditive Szene und Klangobjekte</b> .....	57
6.1.1	Die Darstellung der realen Welt .....	57
6.1.2	Das Hyperreale .....	58
6.1.3	Imaginäre und surreale Räume .....	58
<b>6.2</b>	<b>Diegese</b> .....	59
<b>7</b>	<b>Studio Setup und Software</b> .....	60
<b>8</b>	<b>Gestalterischer Prozess</b> .....	63
<b>8.1</b>	<b>Perspektive</b> .....	63
<b>8.2</b>	<b>Erzähler</b> .....	63
<b>8.3</b>	<b>Musik</b> .....	64
<b>8.4</b>	<b>Einordnung der Szenen</b> .....	65
<b>9</b>	<b>True Spatial</b> .....	66
<b>9.1</b>	<b>Szene 11</b> .....	66
9.1.1	Autofahrt .....	66
9.1.1.1	Das Auto .....	67
9.1.1.2	Die Charaktere .....	69
9.1.1.3	Das Wetter .....	72
9.1.2	Ausstieg/Draußen .....	73
9.1.2.1	Charaktere .....	74
9.1.2.2	Die Atmosphäre .....	76
9.1.2.3	Die Schlucht .....	77
<b>9.2</b>	<b>Szene 12</b> .....	77
9.2.1	Die Schlucht .....	77
9.2.2	Die Blende .....	77
9.2.3	Die Charaktere .....	78

---

9.2.4	Atmosphäre .....	81
9.2.5	Das Autowrack .....	81
<b>10</b>	<b>7.1.4</b> .....	<b>83</b>
<b>10.1</b>	<b>Szene 11</b> .....	<b>84</b>
10.1.1	Autofahrt .....	84
	10.1.1.1 Das Auto .....	84
	10.1.1.2 Charaktere .....	85
	10.1.1.3 Das Wetter .....	86
10.1.2	Ausstieg/Draußen .....	86
	10.1.2.1 Charaktere .....	87
	10.1.2.2 Die Atmosphäre .....	87
	10.1.2.3 Die Schlucht .....	88
<b>10.2</b>	<b>Szene 12</b> .....	<b>88</b>
10.2.1	Die Blende .....	88
10.2.2	Die Charaktere .....	89
10.2.3	Die Atmosphäre .....	89
10.2.4	Das Autowrack .....	89
<b>11</b>	<b>Diskussion</b> .....	<b>91</b>
<b>11.1</b>	<b>Interviews</b> .....	<b>91</b>
<b>11.2</b>	<b>Auffälligkeiten in Produktion und Wahrnehmung</b> .....	<b>93</b>
11.2.1	Umgang mit der akustischen Perspektive .....	93
11.2.2	Umgang mit nichtdiegetischen Elementen .....	94
11.2.3	Neutrale Mischung/Sounddesign .....	96
11.2.4	Die untere Hemisphäre in True Spatial .....	98
11.2.5	Arbeit in True Spatial vs. 7.1.4 .....	101
11.2.6	Lokalisation .....	102
11.2.7	Narrative Fehler/Diegetische Inkonsistenz .....	103

---

11.2.8	Zeitaufwand/Tools .....	104
11.2.9	Size matters .....	106
<b>12</b>	<b>Fazit .....</b>	<b>107</b>
<b>13</b>	<b>Danksagung .....</b>	<b>109</b>
<b>14</b>	<b>Anhang .....</b>	<b>110</b>
<b>14.1</b>	<b>Literaturverzeichnis .....</b>	<b>110</b>
	<b>Eigenständigkeitserklärung .....</b>	<b>118</b>

## Verzeichnis der Abbildungen

<b>Abbildung 2.1</b>	Rundfunkpionier Alfred Braun bei einer Hörspielproduktion mit Geräuschemachern .....	18
<b>Abbildung 2.2</b>	Betriebsanleitung für den Volksempfänger VE 301 Dyn 1938 .....	20
<b>Abbildung 2.3</b>	Junge Menschen produzieren ein Hörspiel, 1970er Jahre .....	22
<b>Abbildung 2.4</b>	IFA Berlin 1973: Vorstellung neuer Kopfhörer in Helmoptik für Kunstkopf-Stereofonie .....	24
<b>Abbildung 3.1</b>	Zuhörer:innen mit Théâtrophone in einem Hotel 1892 .....	28
<b>Abbildung 3.2</b>	Oscar, die von Fletcher entwickelte Kunstkopffigur .....	30
<b>Abbildung 3.3</b>	Stereo-Konfiguration .....	31
<b>Abbildung 3.4</b>	Darstellung von Cinerama .....	33
<b>Abbildung 3.5</b>	Quadrofonie-Setup .....	34
<b>Abbildung 3.6</b>	5.1-Konfiguration .....	38
<b>Abbildung 3.7</b>	Pro Logic IIz in 5.1.2 (links) und 7.1.2 (rechts) .....	39
<b>Abbildung 3.8</b>	NHK 22.2-Aufstellung .....	39
<b>Abbildung 3.9</b>	Lautsprecherverteilung in 10.2 nach ITU .....	40
<b>Abbildung 4.1</b>	Kopfbezogenes Koordinatensystem .....	42
<b>Abbildung 4.2</b>	Lokalisationsunschärfe in Abhängigkeit von der Frequenz .....	43
<b>Abbildung 4.3</b>	Summenlokalisationskurve für Laufzeitdifferenz für Breitbandsignale, Zuhörerkopf fixiert .....	44
<b>Abbildung 4.4</b>	Summenlokalisationskurve für Pegeldifferenz für Breitbandsignale, Zuhörerkopf fixiert .....	45
<b>Abbildung 4.5</b>	Minimale Lokalisationsunschärfe auf der Medianebene für Sprache eines bekannten Sprechers; Kopf fixiert .....	47
<b>Abbildung 4.6</b>	Die richtungsbestimmenden Bänder und der Anhebungsbereich auf der Frequenzachse .....	48
<b>Abbildung 7.1</b>	Lautsprecher-Konfiguration und Aufbau in Regie 3 des Tonlabors der HAW DMI .....	60

---

<b>Abbildung 7.2</b>	Lautsprecheranordnung aus Regie 3 im AllRADecoder .....	61
<b>Abbildung 7.3</b>	Anordnung der virtuellen 7.1.4-Lautsprecheranordnung im MultiEncoder nach ITU-Norm .....	62
<b>Abbildung 8.1</b>	Der Halo-Trick; Halo-Upmix (links) und MultiEncoder (rechts) .....	65
<b>Abbildung 9.1</b>	Spurenverteilung in der True-Spatial-Mischung mit jeweiligem En-/Decoder .....	66
<b>Abbildung 9.2</b>	IEM RoomEncoder (links) und der IEM MultiEncoder (rechts) .....	70
<b>Abbildung 9.3</b>	Darstellung der virtuellen Umgebung in der ersten Hälfte von Szene 11 .....	71
<b>Abbildung 9.4</b>	Darstellung der virtuellen Umgebung in der zweiten Hälfte von Szene 11 .....	75
<b>Abbildung 9.5</b>	Darstellung der virtuellen Umgebung in Szene 12 .....	79
<b>Abbildung 9.6</b>	IEM FdnReverb (links) und IEM DualDelay (rechts) mit den Einstellungen für Justus und Bob aus Szene 12 .....	80
<b>Abbildung 9.7</b>	IEM MultiEQ auf dem Kanal von Justus und Bob aus Szene 12 .....	80
<b>Abbildung 10.1</b>	Spurenverteilung in der 7.1.4-Mischung mit jeweiligem En-/Decoder (links) und der ReaSurroundPan der Firma Cockos (rechts) .....	83
<b>Abbildung 11.1</b>	Beispiel-Projekt der Interviews. Obere Gruppe (A): True Spatial; untere Gruppe (B): 7.1.4; linksseitig die zehn Samples, rechtsseitig die beiden Mischungen .....	92



„Das Hörspiel als Schauspiel für Blinde“

(Kälin, 1991, S. 53)



# 1 Hörspiel: Definition

Was ist ein Hörspiel? Literatur mit Antworten auf diese Frage gibt es, wie das Hörspiel selbst, seit nahezu einem Jahrhundert. Seit jeher hat sich das Hörspiel stetig weiterentwickelt, die Abgrenzung seiner diversen Formen und Arten ist nicht leicht (Köhler, 2005). Die akustische Gattung des Hörspiels entstammt dem Broadcast-Medium Rundfunk, nicht dem Printwesen (Bung & Schrödl, 2017), sie sollte dem Rundfunk als funktionierendem massenkommunikativem Medium eine weitere, eigene Kunstform bieten (Kälin, 1991).

Im Duden heißt es, das Hörspiel sei eine „an die technischen Möglichkeiten des Rundfunks gebundene, auf das Akustische ausgerichtete dramaturgische Gattung“ (Dudenredaktion, o. D.). Anfang der 1920er Jahre bekam diese neue Gattung schließlich ihren Namen: Hörspiel. Der Begriff wurde geprägt durch die Zeitschrift „Der Deutsche Rundfunk“. Das Hörspiel sollte vom anfangs üblichen Sendespiel abgegrenzt werden. Im Sendespiel wurden lediglich Bühnen- und Theaterstücke mittels Rundfunk übertragen. Eigens geschriebene Stücke für den Rundfunk gab es noch nicht (Heister, 1925).

Unter den Begriff Hörspiel werden heute sämtliche Typen und Genres der Gattung im Rundfunk subsumiert (Schanze, 2002). Vorlagen für ein Hörspiel können literarischen Ursprungs sein, sie können aber auch auf Theateraufführungen, Fernsehserien, Kinofilmen oder eigens angefertigten Hörspieldrehbüchern basieren.

Die verschiedenen Typen der Gattung haben unterschiedliche Produktionsformen. Bei deren Einteilung spielen Fragen eine Rolle wie: Ist die Musik dominant, und in welchem Maße? Überwiegt der gesprochene Anteil oder die klangliche Gestaltung der Szenerie? Ein klassisches Hörspiel grenzt sich beispielsweise dadurch vom Sende-, Klang- oder Schallspiel ab, dass beim Hörspiel der Fokus auf dem gesprochenen Wort liegt (Köhler, 2005).

In der Frage, was ein Hörspiel von einem Hörbuch unterscheidet, gehen die Meinungen auseinander. Nach Köhler (2005) sind Begriffe wie Hörbuch oder Audiobuch lediglich Oberbegriffe, die das Hörspiel, neben Vorträgen und Lesungen, miteinbeziehen. Ähnlich sehen das einige Verlage und der Hörspielpreisträger Andreas Ammer, der einmal schlicht feststellte: „Ein Hörspiel ist dann ein Hörspiel, wenn es eine Hörspielabteilung bezahlt“ (Krug, 2008, S. 11). Gisselbrecht (o. D.): „Warum ein Hörspiel die Untergattung eines Hörbuchs sein soll, bleibt mir verborgen“. Gisselbrecht (o. D.) und Blum (2019) machen den Unterschied daran fest, dass das Hörbuch meist von einem/einer Sprecher:in vorgetragen wird und Geräusche bzw. Soundeffekte eine untergeordnete oder sehr geringe Rolle spielen. Anders das Hörspiel, das von der Musik und Klangeffekten lebt.

## 2 Zur Geschichte des (deutschen) Hörspiels

Bevor die klangliche Gestaltung eines modernen Hörspiels in einem mehrdimensionalen Audioformat thematisiert wird, soll zunächst erörtert werden, wie sich das Hörspiel im Zuge der verschiedenen technischen Innovationen der vergangenen 100 Jahre entwickelt hat.

### 2.1 Beginn des deutschen Rundfunks

Hörspiele und Rundfunk gehören untrennbar zusammen. Der Beginn der Geschichte des Hörspiels in Deutschland ist mit der Geschichte des Rundfunks verwoben (Köhler, 2005). Das erste Radioprogramm der deutschen Geschichte wurde am 29. Oktober 1923 live aus Berlin gesendet (Krug, 2008).

„Drei Minuten vor acht Uhr! Alles versammelt sich im Senderaum. Erwartungsvoll beobachtet man das Vorrücken des Zeigers der Uhr ... Acht Uhr! Alles schweigt. In das Mikrofon ertönen nun die Worte: Achtung! Hier Sendestelle Berlin Voxhaus Welle 3.000. Wir bringen die kurze Mitteilung, dass die Berliner Sendestelle Voxhaus mit dem Unterhaltungsfunk beginnt.“ (Leonhard, 1997, S. 23)

In Deutschland nach dem Ersten Weltkrieg herrschten bittere Armut und eine immens hohe Arbeitslosigkeit. Daraus leitete sich auch ein wichtiger Auftrag des Radios ab, wie ihn Hans Bredow auffasste, der damalige Staatssekretär für das Telegrafien-, Fernsprech- und Funkwesen, der auch Vater des Rundfunks genannt wird. Es gilt „die Einsamkeit aus dem Leben des Menschen zu verbannen“, so Bredow in einem Interview mit dem Südwestfunk im Jahr 1954 (Dietze, 1954).

Neben der Unterhaltung habe der Rundfunk aber auch einen Kultur- und Bildungsauftrag, so der damalige Geschäftsführer der Reichs-Rundfunk-Gesellschaft und Radiopionier Kurt Magnus. Deutschland habe viele große Schriftsteller und Denker hervorgebracht, es sei die Verpflichtung des Rundfunks, deren Werke unter das Volk zu bringen (Magnus, 1928). Im Sinne dieses Bildungs- und Unterhaltungsauftrags wurden anfangs vor allem Opern- und Theateraufführungen, Konzerte, Vorträge und literarische Lesungen übertragen. Mit Letzteren war das sogenannte Sendespiel geboren. Bei diesen Sendespielen wurden Texte, die nicht speziell für den Rundfunk geschrieben wurden, von mehreren Personen vorgelesen. Außerdem versuchte man es aber auch mit klassischen literarischen Texten. So wurde zum Beispiel Goethes „Faust“ am 24. November 1924 von Hamburg aus gesendet (Krug, 2008).

Das Hörspiel sollte bereits existierende Kunstformen nicht ändern oder weiterentwickeln. Vielmehr sollte es als neue Kunstform etabliert werden (Kälin, 1991).

## 2.2 A Comedy of Danger und Maremoto

Es ist der 15. Januar 1924. „A Comedy of Danger“ ist beim britischen Sender BBC zu hören. Das europäische Hörspiel oder – wie der Autor Richard Hughes es nannte – das „listening play“ erlebt seine Geburtsstunde. Durch die Wahl des Handlungsorts für das Hörspiel kompensiert Hughes raffiniert das Fehlen eines Bildes. Das Stück spielt in einer eingestürzten Kohle-Mine, in die Wasser dringt. Die Protagonisten sind ein junges Paar und ein älterer Bergmann, die sich in Todesgefahr befinden. Die drei Personen treten in einen Dialog über die Existenz, das Sein und das Nicht-Sein, in einem ständigen Wechsel zwischen Hoffnung und purer Angst. Wie das Drama für die drei ausgeht, bleibt ungewiss. Hughes hat damals auf ein offenes Ende gesetzt (Hörburger, 2006).

„A Comedy of Danger“ soll gehört, nicht gesehen werden. Wäre es eine Theateraufführung, müsste es im Raum komplett dunkel sein. Insofern sei „A Comedy of Danger“ ideal für den Rundfunk, so Nigel Playfair, Produzent der BBC. Jeder, der dieses Stück gehört hat, konnte mit den drei Verschütteten mitfühlen und sich in deren Situation hineindenken. Das Bild der Szenerie entstand in den Köpfen der Zuhörer:innen (Radio Drama at 90, o. D.).

Zur selben Zeit gab es auch in Frankreich ähnliche Versuche im Rundfunk, die ebenso als Vorreiter des heutigen Hörspiels gelten. Am 21. Oktober 1924 wurde das Stück „Maremoto“ gesendet. Dieses eigentlich für einen Radio-Literatur-Wettbewerb geschriebene Stück von Maurice Vinot, der unter dem Namen Gabriel Germinet schrieb, handelt von einem Schiffsunglück. Das auditive Drama sorgte für Aufsehen, da die verwendeten Effekte und die Darbietung des Unterganges so realistisch und beklemmend wirkten, dass einige Hörer:innen sich an die Behörden wandten, um die Ausstrahlung zu stoppen. In dem Hörspiel wurde ‚live‘ von dem Schiffsunglück berichtet, wie man das aus den amerikanischen Nachrichten kannte, nicht aber wie dort durch einen Erzähler, der über allem steht. Als Reaktion auf die Beschwerden der Zuhörer:innen verbot das französische Marineministerium die Ausstrahlung von „Maremoto“ (Cusy & Germinet, 1994).

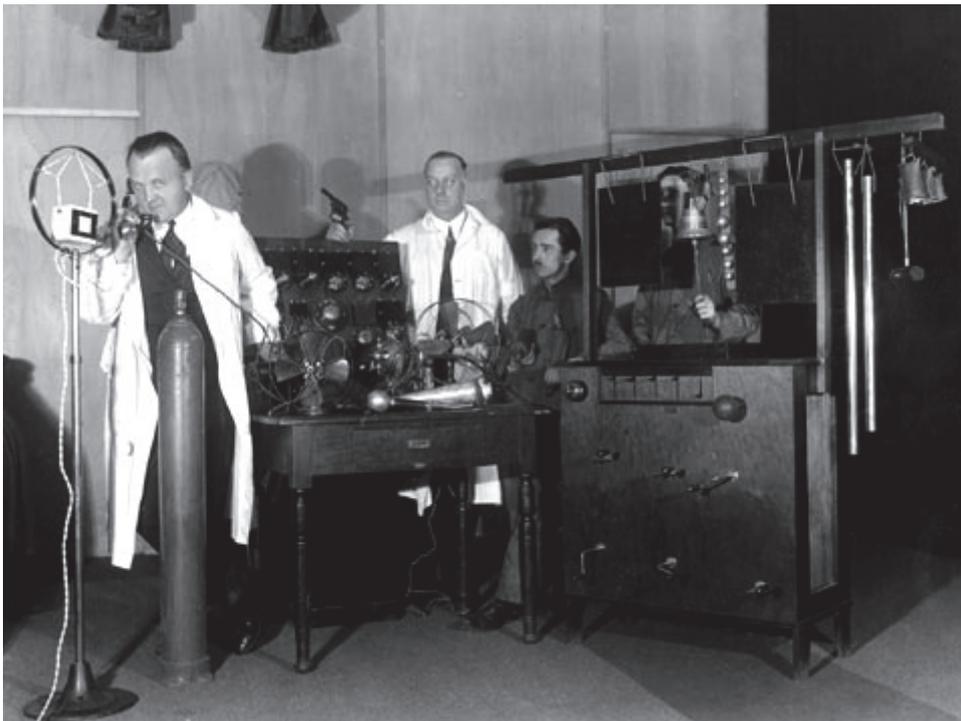
## 2.3 Das erste deutsche Hörspiel

Das erste deutsche Hörspiel feierte am 24. Oktober 1924 Premiere. Geschrieben hat es Dr. Hans Flesch, der damalige künstlerische Leiter des in Frankfurt ansässigen Südwestdeutschen Rundfunkdienstes. Seine „Zauberei auf dem Sender“ mit dem Untertitel „Versuch

einer Rundfunkgroteske“ wurde in diesem Sender uraufgeführt. Das Hörspiel handelt von einer Märchentante, die den Sendeablauf eines Rundfunkhauses stört, indem sie Kindern eine Gute-Nacht-Geschichte erzählen will. Mit dieser „Rundfunkgroteske“ wurden die ersten Versuche einer akustischen Gestaltung auditiver Räume unternommen, weshalb „Zauberei auf dem Sender“ nach Aussagen Fleschs als Bühnenstück nicht funktioniert hätte (Hörburger, 2006).

#### 2.4 Hörspielproduktionen der 1920er Jahre

Frühe Hörspielproduktionen waren enorm aufwändig und konnten nur in den Funkhäusern hergestellt werden. Dabei wurden die Stücke live aufgeführt und aufgenommen. Man orientierte sich stark am Theater und dessen Dramaturgie. So waren die Sprecher:innen vor den Mikrofonen mit Kostümen bekleidet, da die Vorstellung herrschte, dass die Aufführungen auf diese Weise an Authentizität gewinnen würden. Die Schauspieler:innen verharrten während ihrer Auftritte auch keineswegs an einem festen Ort, wo sie ihre Texte vorlasen. Es gab Auftritte und Abgänge wie im Theater (Krug, 2008).



**Abbildung 2.1:** Rundfunkpionier Alfred Braun bei einer Hörspielproduktion mit Geräuschemachern, <https://bit.ly/3zd9q3v>

Erich Kästner hielt 1929 seine Eindrücke fest, als er einer Produktion beiwohnen durfte:

„Kein Wort darf gesprochen oder auch nur geflüstert werden. Zwanzig Menschen, über zwei Räume und einen Flur, der die Säule verbindet, verteilt, und jeder hält ein Textbuch in der Hand, in dem der Regisseur mit Blau- und Rotstift inszeniert hat, und jeder wartet auf bestimmte Winke, winkt weiter, winkt wieder, führt Winkbefehle aus! ... Er selber, der Regisseur, sitzt inzwischen in seiner Isolierzelle, hört per Radio, was außerhalb seiner Zelle geschieht, gibt durch ein Fenster Wink-Kommandos, jagt seine Sendboten zu den Inspizienten, sie möchten den Regen das nächste Mal besser machen, und zu der Schauspielerin X, sie möge lauter sprechen oder eindringlicher weinen.“ (Blaes & Kraft, 2002, S. 28)

Den gelernten Bühnen-Darsteller:innen fiel es schwer, sich an das neue Medium anzupassen. Dem Schauspieler Paul Bildt ging es wie vielen seiner Kolleg:innen: Er war es nicht gewohnt, in absoluter Stille zu spielen. Kein Publikum, das Geräusche macht, und sei es nur das Quietschen von Stühlen oder ein Räuspern. Hinzu kam die Herausforderung, sämtliche Emotionen rein auditiv zu vermitteln. Als zusätzliche Erschwernis wirkte die schlechte Übertragungsqualität. Die genaue Aussprache wurde deshalb noch wichtiger (Krug, 2008).

Doch nicht nur die Schauspieler:innen hatten Probleme. So erklärte beispielsweise der Regisseur und Autor Fritz Peter Buch, dass es nicht darauf ankomme, möglichst viele Hörereignisse in dem Hörspiel unterzubringen, sondern man müsse „... ihn [den Zuhörer] eigentlich vergessen machen, dass er zu Hause in seinen vier Wänden sitzt, den Kopfhörer über den Ohren oder den Lautsprecher vor sich auf dem Tisch“ (Buch 1930, S. 9).

In den frühen Hörspieljahren entstanden hauptsächlich Stücke von Mitarbeitern der Rundfunkanstalten (Leonhard, 1997). Die etablierten Schriftsteller:innen hielten sich mit eigens für das neue Medium geschriebenen Werken zurück.

## 2.5 Aufstieg des Hörspiels

Hörspiele gewannen in der Folge zusehends an Popularität. 1928 soll es insgesamt mehr als drei Millionen Hörer:innen gegeben haben. Heinz Schwitzke (1963) schreibt: „1929 beginnt mit einem Schlag die Zeit des ... Hörspiels in Deutschland.“ Die Produktionszahl stieg seither sprunghaft an: 1930 wurden 854 Hörspiele gesendet, 1932 waren es schon 1.400. Auch die sich anfangs eher zurückhaltenden Schauspieler:innen begannen das Hörspiel zu schätzen, nicht zuletzt auch wegen steigender Gagen. Das durchschnittliche Honorar eines Sprechers lag bei 500 Reichsmark (Krug, 2008), was heute etwa 2.150€ entspräche (Bundestag, 2016).

Seit den Anfängen des Hörspiels wurde ständig mit dem neuen Medium experimentiert. Und es wurden neue Möglichkeiten ausprobiert, bis zum Ende der Weimarer Republik.

„Ob als akustisches Schallspiel, wenn auch mit begrenzten technischen Voraussetzungen, als literarisches Hörspiel oder als Sendespiel, in Anlehnung an das Theater, viele der innovativen Ansätze, aber auch die bereits realisierten und verschiedensten Ausprägungen des Hörspiels fanden ihr vorläufiges Ende mit Hitlers Machtergreifung.“ (Köhler, 2005, S. 24)

## 2.6 Das Dritte Reich – nationalsozialistische Hörspiele

Als die Nazis an die Macht kamen, hatte das zunächst einschneidende Auswirkungen auf das Personal der Rundfunkanstalten, wo Zensur an der Tagesordnung war (ebd.). Rundfunkleiter und Mitarbeiter wurden entlassen, etliche deportiert. Auch der „Vater des Rundfunks“ Hans Bredow und der oben erwähnte Dr. Hans Flesch kamen in Haft (Döhl, 1992). „Bereits Ende 1939 wurden beim Berliner Deutschlandsender (DS) 98 Prozent der eingereichten Hörspielmanuskripte sofort abgelehnt, 1938 wurden gar 99,5 Prozent der Hörspielsendungen über Aufträge gedeckt.“ (Leonhard, 1997, S. 1195)



**Abbildung 2.2:** Betriebsanleitung für den Volksempfänger VE 301 Dyn 1938, <https://bit.ly/3GOYUL3>

Das nationalsozialistische Regime nutzte eigene und überarbeitete Hörspiele wie „Der Ruf“ von Hermann Kasack (1932) zu Propagandazwecken. Kasacks Stück um das Schicksal des Arbeitslosen Martin Keller wurde wie andere Stücke herangezogen, um das Weltbild des Führers zu verbreiten (Kälin, 1991), etwa die Gemeinschaftsideologie, „die dem einzelnen lediglich eine Ersatzindividualität in Form eines imaginären ‚Wir‘ zuspricht ...“ (Kälin, 1991, S. 97). Propaganda-Hörspiele wie „Deutsche Passion 1933“ von Richard Euringer, „Der Weg ins Reich“ von Eberhard Wolfgang Möller, „Mysterium, die Geburt des Reiches“ eines SA-Mannes und viele weitere wurden ausgestrahlt.

Die Zahl der Zuhörer:innen hatte auch unter der Herrschaft der Nationalsozialisten ein hohes Niveau. So verzeichnete man in den frühen

1930er Jahren zwischen sechs bis sieben Millionen Hörer:innen. Damit dies so blieb und möglichst viele Menschen der Propaganda zuhören konnten, führte das Regime 1933 den „Volksempfänger“ ein (Köhler, 2005). Mit solchen Geräten ließen sich auch diejenigen erreichen, die auf dem Lande wohnten und keinen Zugang zu anderen für Propagandazwecke genutzten Stätten wie dem Theater oder dem Kino hatten (Krug, 2008).

Doch die ins Exil vertriebenen Autor:innen und Künstler:innen blieben nicht stumm. Hörspiele wurden zur Gegenpropaganda eingesetzt und zur „Aufrechterhaltung eines nicht faschistischen kulturellen Deutschlands“ (Köhler, 2005, S. 30). Als wichtigste Autor:innen solcher Exilhörspiele gelten Bertolt Brecht („Das Verhör des Lukullus“), Anna Seghers („Der Prozess der Jeanne d’Arc zu Rouen 1431“) und Ernst Ottwalt („Kalifornische Ballade“) (Krug, 2008).

## 2.7 Die zweite Blütezeit

1945, Nachkriegsdeutschland. In einer Zeit, in der Theater, Kinos, Zeitschriften und Bücher rar waren, erlebte das deutsche Hörspiel eine zweite Blüteperiode (Krug, 2008). Wo zuvor noch die Nazis maßgeblichen Einfluss hatten, waren es nun die alliierten Besatzungsmächte, die die Inhalte der Hörspiele wesentlich beeinflussten. Nun sollten nicht mehr nazi-propagandistische Hörspiele in den Radios ausgestrahlt werden, sondern politisch bildende, vor allem zu Demokratie erziehende Stücke. Das Funkmedium sollte unbedingt souverän und unabhängig sein, damit sich die verhängnisvolle Geschichte nicht wiederholte (Köhler, 2005).

Vor allem der Nordwestdeutsche Rundfunk (NWDR) mit Sitz in Hamburg hatte sich die politische Aufklärung zur Aufgabe gemacht. Der NWDR verhalf 1947 Wolfgang Borchert zu seinem großen Durchbruch. Der später als Schriftsteller bekannte Borchert schrieb in nur acht Tagen das eigentlich fürs Theater konzipierte Stück „Draußen vor der Tür“ (Krug, 2008). In diesem Hörspiel verarbeitet er die verheerenden Auswirkungen des Krieges, die nicht nur die Physis der Deutschen betrafen, sondern auch ihren mentalen Zustand und ihr existenzielles Gefühl von Heimatlosigkeit. Das Stück fand sensationelle Resonanz und rührte an den Nerv der Zeit (Wellershoff, 1985). Die Ära des Nachkriegshörspiels war angebrochen (Krug, 2008).

Ebenso aufsehenerregend wirkte das Hörspiel „Träume“ von Günter Eich. „Kein Hörspiel konnte später solch geballte Reaktionen hervorrufen wie jene Erstsending.“ (Krug, 2008, S. 55) Noch während „Träume“ am 19. April 1951 erstmals ausgestrahlt wurde, beschwerten sich beim NWDR zahllose Anrufer:innen, dass man so etwas nicht senden könne. Eichs Stück sorgte für Entsetzen, da es unter anderem in Kinderschlachthäusern eines Vampirs spielt (ebd.). Mit dieser alptraumhafte Szenerie verstand es Eich, die Angst der Hörer:innen künstlerisch sehr wirksam zu schüren (Köhler, 2005). In der Hörspielforschung zählt „Träume“ bis heute zu den wichtigsten Hörspielen (Krug, 2008).

## 2.8 Einzug des Fernsehens

Bis 1952 war das Radio und somit auch das Hörspiel als Unterhaltungsmedium konkurrenzlos (Bannasch, 2020). Mit der Einführung des Fernsehens, das das Kino in die eigenen vier Wände brachte, ließ das Interesse am Hörspiel mit der Zeit nach. Der nun größte Konkurrent raubte dem Radio zusehends die Hörerschaft (Krug, 2008). Dies hatte zur Folge, dass auch das Personal zum neuen visuellen Massenmedium wechselte (Köhler, 2005). Wo das Hörspiel zuvor eine wichtige Rolle in den Sendeprogrammen einnahm, liefen ihm nun „rock- und poporientierte Jugendprogramme wie ‚Teens-Twens-Top-Time‘“ (Krug, 2008, S. 81) den Rang ab. Der Anteil an gesendeten Hörspielen lag etwa beim Hessischen Rundfunk (HR) 1967 bei nur noch 0,2 Prozent, der Südwestfunk (SWF1) nahm Hörspiele komplett aus dem Programm (ebd.).

## 2.9 Das neue Hörspiel



**Abbildung 2.3:** Junge Menschen produzieren ein Hörspiel, 1970er Jahre, <https://bit.ly/3xhMY7R>

Ende der 1960er Jahre hatte sich der Stereohörfunk so weit etabliert, dass jede öffentlich-rechtliche Sendeanstalt mindestens ein stereofones Programm hatte (Krug, 2008). Mit diesem technischen Durchbruch begann die Zeit des sogenannten „Neuen Hörspiels“ (Bannasch, 2020), das viele Hörspielautor:innen anzog, die der neuen Gattung eine eigene, offene Dramaturgie

gaben (Schöning, 1982). Zusätzlich ermöglichte es die Etablierung des Tonbandgerätes, Klänge, Ortsgeräusche und Sprache oder Töne von Menschen aufzunehmen, die sogenannten Originaltöne (Krug, 2008). Unter Originalton (O-Ton) versteht man eine „unveränderte, nicht übersetzte, nicht manipulierte Tonaufzeichnung“ (Dudenredaktion, o. D.).

Diese Aufzeichnungen ließen sich später schneiden und völlig neu arrangieren. Paul Wühr erhielt 1971 den Hörspielpreis der Kriegsblinden für sein Originalton-Hörspiel „Preislied“. Dabei hat er die Tonbänder der zwölf Stunden langen Sprachaufnahmen zerschnitten und neu arrangiert. Bei Originalton-Hörspielen fiel kein Honorar für Schauspieler:innen an, da sie schlicht nicht benötigt wurden. Die Arbeit eines Hörspielautors ähnelte der eines Reporters: Aufgezeichnet und gesprochen wurde das, was einem vor das Mikrofon lief (Krug, 2008).

Wo zuvor die deutsche Literatur Vorbild war, suchte man sich seine Vorbilder nun in der experimentellen Kunst der 1920er Jahre, etwa in Kurt Schwitters' Lautmalereien oder Pierre Schaeffers „Musique concrète“. Dieser Typ Hörspiel fand bei der breiten Masse allerdings kaum Anklang und wanderte in die Kulturprogramme ab (Krug, 2008).

## 2.10 Stereo als technische Innovation

Die technischen Innovationen der Stereophonie, deren Etablierung in den Rundfunkanstalten Mitte der 1960er Jahre stattfand, und der Klangbearbeitung machten es möglich, Stücke besonders realitätsnah oder eben auch realitätsfern zu gestalten (Köhler, 2005). Die „Klarheit“ nimmt zu, „störende Nebengeräusche“ nehmen ab (Keckeis, 1973, S. 35). Stimmen ließen sich nun im Raum verteilen oder stark verfremden (Köhler, 2005). Der Hörspielautor und ehemalige Redaktionsleiter des WDR, Klaus Schöning, erläutert:

„Die Stereophonie machte den Hör-Raum als Spielfläche zwischen den beiden Lautsprechern erfahrbar, auf der nun die einzelnen akustischen Elemente choreografiert werden konnten. Aus dem einseitigen Instrument der Mikrofonie wird in der Stereophonie ein Instrument über zwei Seiten. Der zwischen den beiden Seiten existierende Raum wird zum Spiel-Raum, die Gleichzeitigkeit ablaufender Ereignisse durch die stereofonen Positionen gleichsam sichtbar verdinglicht. Eine Erfahrung, die den konkreten Hör-Raum des Radio-Zuhörers, der in der monotonen Rezeption durch entsprechende Stücke als innere Bühne vermännlicht worden war, ins Spiel brachte.“ (Stuhlmann, 2001, S. 255 f.)

Doch der Vormarsch der Stereophonie in der Hörspiel-Welt hatte nicht nur Anhänger. Einige Skeptiker wollten an der „Dramaturgie und Ästhetik des traditionellen Hörspiels festhalten“ (Köhler, 2005, S. 44). Eben diese Dramaturgie wurde allerdings schon in den Anfängen des Hörspiels kritisiert. Da sich das Hörspiel früh Stücke aus Literatur und Theater zum Vorbild nahm oder sie adaptierte, entwickelte sich zunächst keine eigene Dramaturgie des Rundfunk-Mediums (ebd.). Mit dem Rückzug der Literatur aus den Hörspielen setzte ein weiterer Niedergang ein. Die einstige Radiokunst konnte sich zwischen Nachrichten und Musik nicht mehr behaupten (Krug, 2008). Daran änderte weder die Einführung der Kunstkopf-Stereophonie noch die der Quadrofonie etwas (ebd.) (siehe Kapitel 3: Wiedergabeformate).

„Ein vierseitiges Lautsprecher-Instrument hätte den Spielraum erweitern können, doch konnte sich die Quadrofonie für die Radio-Rezeption nicht durchsetzen. Auch Versuche mit der so genannten Kunstkopf-Stereophonie, über Kopfhörer

den ganzen Raum plastisch zu erfahren, wurde kaum von den Hörern angenommen.“ (Stuhlmann, 2001, S. 255 f.)

## 2.11 Kunstkopf-Hörspiele

Nachdem der Kunstkopf 1973 auf der Internationalen Funkausstellung (IFA) in Berlin vorgestellt worden war, kam das erste Kunstkopf-Hörspiel „Demolition“ im selben Jahr auf den Markt. Der Science-Fiction-Krimi, der vom ehemaligen leitenden Regisseur des Berliner Radiosenders RIAS (Rundfunk im amerikanischen Sektor), Ulrich Gerhardt, geschrieben wurde, ist 1973 in Zusammenarbeit mit dem Heinrich-Hertz-Institut (heutiges Fraunhofer-Institut) entstanden. Als Vorlage diente Gerhardt das von Alfred Bester geschriebene Stück „The Demolished Man“ von 1953 (Arnold, 2013). Nach Gerhardts Einschätzung scheiterte die Gattung der Kunstkopf-Hörspiele vor allem an zwei Faktoren: am Fehlen von Autor:innen und an den hohen Produktionskosten:

„Es gab keine Autoren dafür, also was Hörspiel anging, gab es keine Stücke. Die interessantesten Sachen waren eigentlich Bearbeitungen aus der Literatur, da konnte man ziemlich viel an Effekten herausholen. Die Autoren selber waren daran irgendwie nicht interessiert, die waren literarischer interessiert.“ (Arnold, 2013, 18:32–18:52)



**Abbildung 2.4:** IFA Berlin 1973: Vorstellung neuer Kopfhörer in Helmoptik für Kunstkopf-Stereofonie, <https://bit.ly/3aHddMj>

„Zu der Zeit sind für eine Produktion fünf bis zehn Tage benutzt worden, also bei Mono waren's fünf Tage, mit Einführung von Stereo waren es dann schon das Doppelte, die zehn Tage, und für Kunstkopf je nach Größe des Stückes und Aufwand war's dann noch mal doppelt so viel oder dreimal so viel.“ (Arnold, 2013, 18:59–19:24)

Allmählich stagnierte in der Folge die Produktion der Kunstkopf-Hörspiele (Arnold, 2013). Trotz der neuen technischen Möglichkeiten und unter dem Eindruck des Streits, den die Verfechter des alten mit denen des neuen, modernen Hörspiels austrugen, verlor das Hörspiel in den 1960er und 1970er Jahren an Popularität. Erst in den frühen 1980er Jahren konnte der Dissens mit dem Ergebnis geklärt werden, dass jeder Hörspiel-Gattung ihre Daseinsberechtigung zugestanden wurde (Köhler, 2005). Die öffentlich-rechtlichen Rundfunkanstalten der ARD, RIAS Berlin und der Deutschlandfunk strahlten in den späten 1970ern und Anfang der 1980er jährlich etwa 1.500 Hörspiele aus. Darunter befanden sich Kinderhörspiele, etwa zu Büchern von Astrid Lindgren, die von der ARD schon in den 1950er Jahren gesendet wurden, Kriminalhörspiele, Sozialhörspiele, Science-Fiction-Stücke, aber auch künstlerische Tonbandwerke. Die Hörspielsparte im Rundfunkprogramm war abwechslungsreich geworden wie nie zuvor (Krug, 2008).

## 2.12 Digitalisierung und das Internet als Plattform

Spätestens seit der Etablierung des Internets durchlebt das Hörspiel einen grundlegenden Wandel. „Die Radiokunst [das Hörspiel] versucht das Internet als neue Bühne zu nutzen. Die Entwicklung steht erst in den Anfängen, aber nicht mehr nur gelegentlich wird heutzutage von einer ‚Fusion von Radio und Internet‘ – und der ‚Neuerfindung des Radios‘ – gesprochen.“ (Krug, 2008, S. 10). Im Zuge dieses Wandels hat sich auch die Rolle des Hörspiels verändert.

„Gerade in einer Sozialisationsphase, wo junge Menschen mit einer immer umfangreicheren und schnelleren Entwicklung der elektronischen Medien, wie Computertechnologie und Fernseher, konfrontiert waren, hat sich dieses Medium dennoch seinen Wert erhalten. Es bietet vor allem eine attraktive, pädagogisch wertvolle und phantasievolle Alternative zur Bilderflut des Fernsehens oder zu den hektischen Computerspielen ... Das Hörbuch oder Hörspiel tritt oft an Stelle der Eltern, die früher aus einem Buch vorgelesen hatten und durch neue Rollenverhältnisse in der Familie (z. B. zunehmende Berufstätigkeit beider Elternteile) gar keine Zeit

mehr haben, ihren Kindern Geschichten mündlich vorzutragen und ihnen damit Literatur näher zu bringen.“ (Köhler, 2005, S. 83 f.)

Die Produktion von Hörspielen ist durch die Digitalisierung in den letzten Jahrzehnten viel einfacher geworden. Aufnahmen sind jederzeit möglich und leicht verfügbar. Die umständliche Arbeit mit Tonbändern ist obsolet geworden. Auch die Anzahl der Spuren und Sprecher hängt nicht mehr von der Breite des Tonbands ab (Krug, 2008).

### **2.13 Bedeutung des heutigen Hörspiels**

Alles in allem hat das Hörspiel heutzutage keine große Bedeutung mehr. Genaue Zahlen zu Verkäufen oder Downloads sind nicht bekannt (Krug, 2008). Der Anteil der Hörspiele am Hörbuchmarkt macht wohl etwa 20% aus (Natusch, 2014). Der Hörspielmacher Oliver Wenzlaff sagte in einem von der Hörspiel-Gemeinschaft e.V. 2013 veranstalteten Vortrag, dass am Hörspielmarkt Die drei ??? und TKKG einen dominierenden Anteil von 60 bis 80% haben, wovon alleine Die drei ??? 40% ausmachen (Wenzlaff, 2013). In den ohnehin seltenen Erhebungen wird allerdings nicht zwischen dem Hörbuch und dem Hörspiel unterschieden, was eine genauere Betrachtung des Hörspiels erschwert (Natusch, 2014).

### **2.14 Die drei ???**

Das Abenteuer der drei ???, denen ich mich in dieser Arbeit schwerpunktmäßig widmen werde, begann in den frühen 1960er Jahren.

Die Erfolgsgeschichte der drei Privatdetektive Justus Jonas, Peter Shaw und Bob Andrews beginnt in den USA. Dort schuf der an der Ostküste in Cape May lebende Robert Arthur das Trio 1964 unter dem ursprünglichen Serientitel „Alfred Hitchcock and the Three Investigators“. Bis zu seinem Tod fünf Jahre später schrieb er elf Folgen. Noch zu Arthurs Lebzeiten kamen „Die drei ???“ 1968 auf den deutschen Markt, vorerst aber nur in Buchform (Die drei ???, o. D.). Außer in Deutschland wurden die Kriminalromane mit den drei Jungs als Hauptfiguren in dreißig weiteren Ländern veröffentlicht (Akstinat, 2008).

Über ein Jahrzehnt später startete die Hörspielserie mit „Die drei ??? und der Super-Papagei“ als erste Folge, die am 12. Oktober 1979 ausgestrahlt wurde. Bemerkenswert ist, dass bis zum heutigen Tag die deutschen Sprecher der Protagonisten dieselben geblieben sind: Seit über vierzig Jahren leihen Oliver Rohrbeck (Justus Jonas), Jens Wawrczeck (Peter Shaw) und Andreas Fröhlich (Bob Andrews) den Ermittlern aus Kalifornien ihre Stimmen.

Der 10. Januar 2005 war ein Schock für die Fans der Krimi-Reihe. „Der schwarze Skorpion“ wurde als 120. und vorerst letzte Hörspielfolge der drei ??? ausgestrahlt (Die drei ???, o. D.). Das Hörspiellabel Europa konnte aufgrund eines Lizenzrechtsstreits mit dem Kosmos Verlag die Serie nicht mehr fortsetzen und stoppte die Produktion nach dieser Folge (Gänßle & Kuchinke, 2020). Doch sollte es nicht auf Dauer still bleiben um die drei. Am 13. Oktober 2006 wurde ein neues Abenteuer aus Rocky Beach ausgestrahlt, dem fiktiven Heimatort der drei Jugenddetektive. Aus rechtlichen Gründen wurden „Die drei ???“ umbenannt in „Die Dr3i“. Auch die Namen Justus Jonas, Peter Shaw und Bob Andrews wurden als Folge des Rechtsstreits geändert. Dabei bekamen sie die Originalnamen zurück, die Arthur ihnen ursprünglich gegeben hatte: Aus Justus Jonas wurde Jupiter Jones, aus Peter Shaw wurde Peter Crenshaw, einzig Bob Andrews durfte seinen Namen behalten.

Nach drei Jahren wurde der Lizenzstreit endlich beigelegt, Anfang April 2008 ein neuer Fall der drei ??? gesendet. 2009 ließ sich Sony BMG, dessen Label Europa unter anderem die Audionutzungsrechte innehat, etwas einfallen und veranstaltete „Die drei ??? und der seltsame Wecker“ live. Dabei waren außer den drei Sprechern auch ein Geräuschemacher und Live-Musik zu hören. Ein Jahr darauf folgte ein Weltrekord: 2010 schauten 15.211 Menschen den Detektiven live in der Berliner Waldbühne zu. Seit 2015 lösen Die drei ??? ihre Fälle auch in 3D-Audio. Im Planetarium und im Mediendom in Kiel und im Planetarium Hamburg kann man unter „3D Sound“ beim Lösen der Fälle mitfiebern (Die drei ???, o. D.). Bis heute haben die Jungs aus Rocky Beach in 214 Hörspiel-Folgen plus 20 Spezialfolgen Verbrecher gejagt und Fälle gelöst (Stand: 9. März 2022).

## 3 Wiedergabeformate / Geschichte der mehrkanaligen Audiosysteme

Für die Wiedergabe der im Rahmen dieser Arbeit gestalteten zwei Hörspielszenen nutze ich ein recht komplexes mehrdimensionales Wiedergabesystem. Zum besseren Verständnis solcher Systeme soll zunächst deren Entstehung näher beleuchtet werden. Gleichzeitig versuche ich aufzuzeigen, warum sich bestimmte Systeme durchgesetzt haben, andere aber nicht.

### 3.1 Théâtrophone

Wenige Jahre nach der Erfindung des Telefons (Boren, 2017) stellte Clément Ader bei der Exposition International d'Électricité 1881 in Paris das „Théâtrophone“ (Hawes, 1991) vor. Bei diesem System wurden die live aufgeführten Opern- und Theaterstücke (ebd.) von Mikrofon-Paaren an der Bühne aufgenommen und kabelgebunden über längere Distanzen an je zwei Telefonempfänger (Boren, 2017) übertragen, die sich die Zuhörer:innen an die Ohren hielten (Görne, 2015). Nicht nur die Schauspieler:innen und das Orchester waren dabei zu hören, auch die Reaktionen des Publikums waren Teil der Übertragung (Hawes, 1991). Die erste mehrkanalige Musikübertragung war geboren (Görne, 2015).



**Abbildung 3.1:** Zuhörer:innen mit Théâtrophone  
in einem Hotel 1892, <https://bit.ly/3PTNkZE>

### 3.2 Monofonie

Die ersten verfügbaren Lautsprechersysteme waren jedoch monofon. Bei diesen Mono-Systemen diente ein einziger Lautsprecher, an dem sämtlicher aufgenommener Schall über ein einzelnes Signal anliegt, der Klangwiedergabe. Ein Problem bei Systemen mit nur einem Lautsprecher ist die Wiedergabe von Rauminformationen (Watkinson, 1998). Nach Steinberg und Snow ist es möglich, mit einem Lautsprecher den Eindruck von Tiefe zu vermitteln, doch sei die Lokalisation immer eindimensional und beziehe sich stets auf die Richtung des Lautsprechers (Steinberg & Snow, 1934, S. 12). Um Räumlichkeit zu vermitteln, seien zwei oder drei Lautsprecher nötig oder eine sogenannte binaurale Wiedergabe (ebd.).

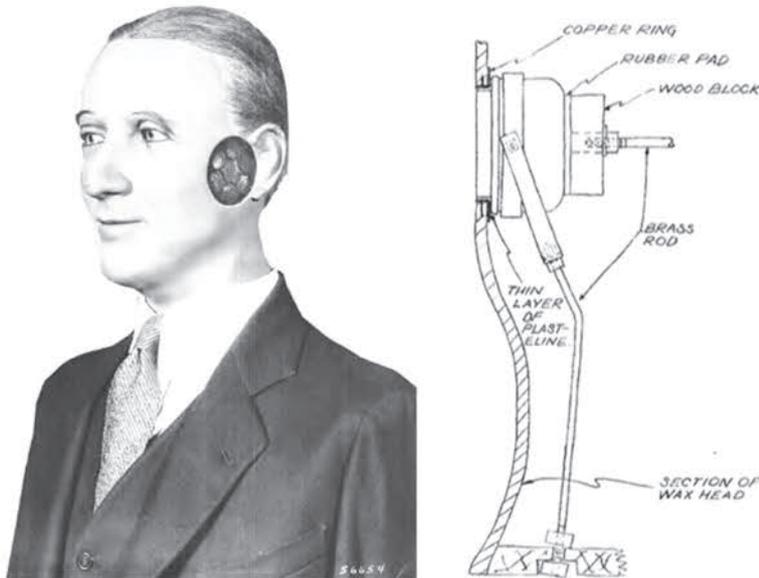
### 3.3 Binaural/Kopfbezogene Stereofonie

Der Begriff binaural wurde 1861 von Somerville Scott Alison geprägt, der damit zum Ausdruck brachte, dass am menschlichen Hören zwei Ohren beteiligt sind. Entsprechend wurde dieser Begriff bis in die 1970er Jahre hinein häufig für Techniken verwendet, die zwei Signale aufzeichnen oder wiedergeben, welche für beide Ohren bestimmt sind, aber nicht unbedingt für Signale, die vom menschlichen Körper modifiziert werden und daher den Signalen am Trommelfell des Hörers entsprechen (Paul, 2009). Theile beschreibt die kopfbezogene Stereofonie als ein Verfahren, bei dem

„das Schallfeld, das im Aufnahmeraum am Ort der beiden Ohren eines Hörers herrschen würde, ... durch einen sog. Kunstkopf aufgenommen [wird] und an den Ohren des Hörers mit Hilfe von Kopfhörern reproduziert [wird]. Der Kunstkopf ist in akustischer Hinsicht dem menschlichen Kopf weitestmöglich nachgebildet. Statt Trommelfelle trägt der Kunstkopf an entsprechender Stelle im Gehörkanal Mikrofone, deren resultierende Richtcharakteristik und Pegel- und Laufzeitdifferenzen den Verhältnissen an den menschlichen Ohren entsprechen“ (Theile, 2008, S. 296).

Die ersten binauralen Übertragungen mit einer Art Kunstkopf wurden bereits 1927 patentiert, eines von Harvey Fletcher in Zusammenarbeit mit Leon Sivian und ein weiteres System zur Aufnahme und Wiedergabe von W. Bartlett Jones (Fletcher & Sivian, 1927; Paul, 2009). Beide verwendeten einfache kugelförmige Objekte als Kunstkopf. Fletchers Forschungen in den Bell Laboratories sollten im Jahr 1931 ein ausgefeilteres binaurales Aufnahmegerät hervorbringen: eine Schneiderpuppe mit dem Spitznamen „Oscar“ (Paul, 2009). Die in Oscar untergebrachten 1,4-Zoll-Mikrofone waren zu groß für den menschlichen Gehörgang; daher wurden sie

auf den Wangenknochen der Puppe direkt vor den Ohren angebracht (Boren, 2017). 1931 und 1932 fanden Test-Übertragungen aus der Philadelphia Academy of Music statt, bei denen Oscar als Aufnahmegerät Verwendung fand.



**Abbildung 3.2:** Oscar, die von Fletcher entwickelte Kunstkopffigur (Paul, 2009, S. 770)

Auf der Weltausstellung 1933 in Chicago wurde Oscar der Öffentlichkeit präsentiert. Dabei wurde die Puppe auf eine verglaste Bühne gestellt, über Kopfhörer konnte man hören, was Oscar hört. Eine Person ging sprechend um Oscar herum, was die begeisterten Hörer:innen ebenfalls vernehmen konnten. Doch die Anwesenden konnten die sich bewegende Person nicht nur hören, sondern auch sehen, was die auditive Lokalisierung verfälscht. Bei der Lokalisation traten überdies Irrtümer auf. Quellen, die sich vorne befanden, wurden beispielsweise so wahrgenommen, als befänden sie sich hinten. Auch bei der Einschätzung von Entfernungen unterliefen Fehler (Paul, 2009). Solche Irrtümer sind auf die nicht-individualisierten binauralen Aufnahmen und die Mikrofone zurückzuführen, die außerhalb der Ohrmuschel an Oscar angebracht waren (Boren, 2017; Paul, 2009).

Die Mikrofone wurden mit der Zeit kleiner, zudem wurden Kunstköpfe entwickelt, bei denen das Mikrophon, wie Theile dies beschreibt, im Gehörgang angebracht sind (2008). Ungeachtet technischer Fortschritte blieb die binaurale Technologie für den größten Teil des 20. Jahrhunderts ein Nischensektor der Audiotechnik (Boren, 2017). Die Gründe dafür sind vielfältig. Für eine genaue Reproduktion der Signale muss das Außenohr des Kunstkopfes und dasjenige des Hörers idealerweise gleich sein, andernfalls kommt es zu Beeinträchtigungen (Theile,

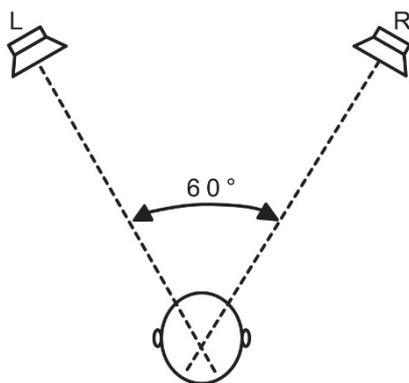
2008). Der ehemalige Hörspielleiter des Hessischen Rundfunks (HR), Christoph Buggert, gibt einen weiteren Grund für das Scheitern des Kunstkopfes an:

„In den vollen Genuss der totalen Raum-Abbildung kommen nur Hörer, die über eine Stereo-Empfangsanlage mit Stereo-Kopfhörern, am besten sogenannten offenen Kopfhörern verfügen. Ein relativ befriedigender Empfang ist noch mit zwei Stereo-Lautsprechern möglich, die einander gegenübergestellt werden, so dass sie eine Art überdimensionale Kopfhörer-Anlage ergeben, wenn der Hörer zwischen ihnen sitzt.“ (Arnold, 2013, 13:51–14:16)

Der erste offene Kopfhörer kam nur wenige Jahre vor Einführung des Kunstkopfes auf den Markt: der HD 414 der Firma Sennheiser (Sennheiser, o. D.).

### 3.4 Stereofonie

Nach ersten Versuchen und Experimenten mit der Stereofonie Ende des 19. Jahrhunderts (Görne, 2015) begann die kommerzielle Nutzung von Stereosystemen, wie wir sie heute kennen, in den 1930er Jahren mit Alan Blumlein und Harvey Fletcher (Power, 2015). In dem Patent, das der Elektroingenieur Blumlein am 14. Dezember 1931 anmeldete, macht er nicht nur genaue Angaben zu der stereofonen Aufnahme, sondern auch zur Übertragung und Wiedergabe (Blumlein, 1931). „Hier hat Blumlein die beiden entscheidenden Motivationen der Mehrkanaltechnik beschrieben: die Phantomschallquellenortung und die Abbildung des Aufnahme-raums durch die Richtungsinformation der seitlichen Reflexion.“ (Görne, 2015, S. 9) Doch die neue Technik setzte sich erst in den 1950er Jahren als Standard durch (Torick, 1998; Görne, 2015).



**Abbildung 3.3:** Stereo-Konfiguration

Die Aufstellung der Stereoabhöre für den Rundfunk ist festgehalten in den Empfehlungen der International Telecommunication Union (ITU) ITU-R BS.2051-2. Dabei empfiehlt die ITU, die Lautsprecher auf einem Azimut von  $\pm 30^\circ$  vor den Hörer:innen aufzustellen (ITU, 2018). Werden bei dieser Aufstellung zwei identische Signale abgespielt, nimmt man nicht zwei getrennte Schallquellen wahr, „sondern nur eine einzige fiktive Schallquelle, eine sog. Phantomschallquelle“ (Theile, 2008, S. 187). Ob sich dieser Phantom-

Effekt jedoch einstellt, ist stark von der Position der Hörer und der Raumakustik abhängig (Holman, 2008). Trotz des „Quantensprungs“ (Görne, 2015, S. 304) der zweikanaligen Stereophonie zur Monophonie wurde weiter nach Möglichkeiten gesucht, die Wiedergabe zu verbessern (ebd.).

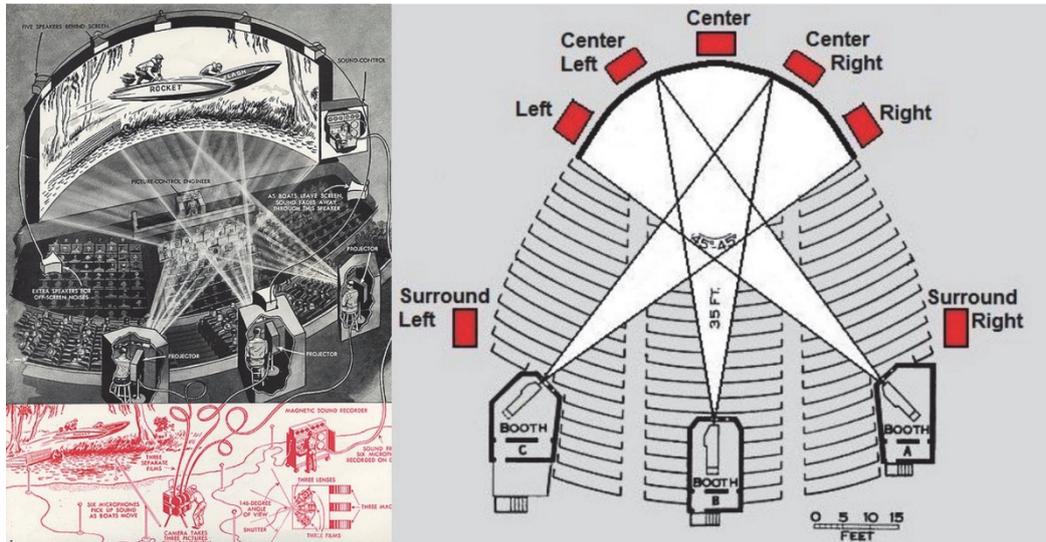
### 3.5 Über Stereo hinaus

Während Blumlein sich mit der Stereophonie beschäftigte, befasste sich Harvey Fletcher 1933 mit mehrkanaligen Aufnahmen und deren Wiedergabe. In einem seiner Versuche nahm er ein Orchester mit einer Wand aus Mikrofonen auf. Jedes dieser Mikrofone war mit einem dedizierten Lautsprecher für die Wiedergabe verbunden. Die Anordnung brach er allerdings später auf drei Kanäle herunter, wegen der unzulänglichen Fähigkeit des menschlichen Gehörs, auf der vertikalen Ebene zu lokalisieren (Streicher & Everest, 1998). Doch einige der Zuhörer:innen aus dem Publikum gaben an, die neue Technik mit drei Lautsprechern vermittele eine größere Emotionalität als ein Live-Auftritt (Torick, 1998).

Ein weiterer wichtiger Schritt in der Entwicklung der Mehrkanaltechnik war 1940 der Film „Fantasia“, der mit Fantasound realisiert worden war, einer Technik, die W. E. Garity und J. N. A. Hawkins bei Disney entwickelt hatten (Torick, 1998). Dieses System arbeitete mit zehn Lautsprechern: drei Lautsprechern vorne (vorne links, vorne Mitte, vorne rechts) sowie sieben weiteren Lautsprechern, die um das Publikum verteilt waren (vordere linke/rechte Ecke, seitlich links/rechts, hinten links/rechts und an der Decke). Die seitlichen und hinteren Lautsprecher wurden nur dezent eingesetzt. Synchron zum Filmband lief eine vierkanalige Tonspur mit drei Audiospuren und eine Kontrollspur. Mit Fantasound ließen sich Klänge über die Lautsprecher schwenken (eng. *panning*) (Davis, 2003). Da es für eine Person unmöglich war, die Vorstellungen jedes Mal gleich zu gestalten, entwickelte Garity „Togad“, ein „tone-operated gain-adjusting device“, das diese Arbeit übernahm (Torick, 1998). Das Surround-System war finanziell kein Erfolg (Holman, 2008) und fand nur wenig Interesse. Fantasound wurde daraufhin nach Russland verkauft, ging aber bei der Überfahrt verloren (Davis, 2003). Ungeachtet dessen ist Fantasound der Wegbereiter für sämtliche Surround-Systeme, die wir heute kennen (Holman, 2008).

Am 30. September 1952 kam mit dem Dokumentarfilm „This Is Cinerama“ ein Audiosystem zum Einsatz, das ähnlich aufgebaut war wie Fantasound. Fred Waller hatte dafür ein Videosystem entwickelt, bei dem drei Projektoren zu jeweils einem Drittel synchron eine Leinwand bestrahlen. Das zugehörige mehrkanalige Soundsystem, konzipiert von Hazard E. Reeves, bestand aus einem Magnetband mit sieben Kanälen: sechs Audiokanälen und einem Kontrollkanal. Ähnlich wie bei Fantasound gab es Lautsprecher vorne, in diesem Fall fünf, und

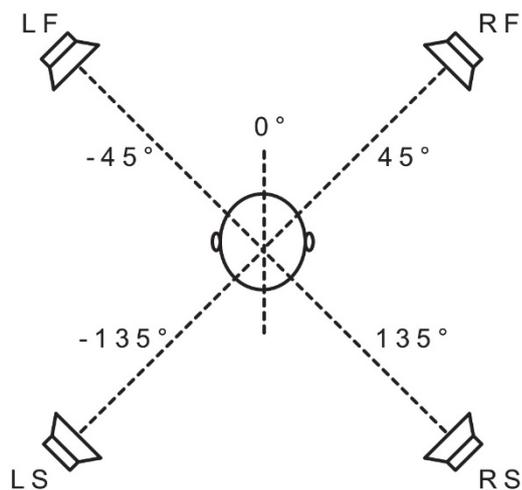
Lautsprecher, die um das Publikum verteilt waren, die Surround-Lautsprecher. Zuvor bekamen die Surround-Lautsprecher ihr Signal aus demselben Kanal, das änderte sich bei dem neuen Format, bei dem die Quellkanäle separat den Surround-Lautsprechern zugespielt werden konnten, was bestimmte Szenen noch lebhafter wirken ließ. Da aber die Kosten, das System zu betreiben, für das Kino zu hoch waren, wurde es 1963 wieder abgeschafft (Davis, 2003).



**Abbildung 3.4:** Darstellung von Cinerama (Davis, 2003, S. 560 & <https://bit.ly/3NiGnji>)

### 3.6 Quadrofonie

Während sich in den 1950er Jahren das Stereoformat zusehends etabliert hatte, vor allem durch Innovationen wie die Stereo-Platte 1958 und die Einführung des FM-Rundfunks, standen die 1970er Jahre ganz im Zeichen der Quadrofonie (Davis, 2003). Robert Berkovitz führte einige Testaufnahmen an der Acoustic Research Corporation durch, bei denen er zusätzlich zu dem vorderen Lautsprecherpaar ein weiteres Paar hinter die Hörer:innen platzierte. Die vorderen Lautsprecher standen, anders als bei Stereo, in einem Winkel von  $\pm 45^\circ$ , die hinteren Lautsprecher in einem Winkel von  $\pm 135^\circ$  (Power, 2015). Nach erfolgreichen Tests kam es rasch zu einer kommerziellen Vermarktung in großem Stil. Als neue Herausforderung stand den Ingenieur:innen nun bevor, zwei zusätzliche Kanäle auf die verbreiteten Stereo-Datenträger zu bekommen (Davis, 2003; Power, 2015). Peter Scheiber, ein professioneller Fagotist, fand schließlich eine Möglichkeit, mit einem Matrixverfahren vier Audiokanäle auf zwei Kanäle zu komprimieren und daraufhin wiederherzustellen (Torick, 1998).



**Abbildung 3.5:** Quadrofonie-Setup

Matrixsysteme entwickelten, alle aber basierten auf dem System von Peter Scheiber (Torick, 1998). Bei diesen sogenannten 4-2-4-Matrix-Verfahren entsprechen die aufeinanderfolgenden Zahlen der Anzahl der Kanäle in den Prozessen der En- und Decodierung (Breh, 1973). Als Scheiber 1968 sein System vorstellte, war die Resonanz eher verhalten. Doch hatte er das Interesse des Ingenieurs Benjamin Bauer von CBS geweckt. Bauer entwickelte auf Basis von Scheibers Matrix ein eigenes 4-2-4-System, das zu den wichtigsten Quadrofonie-Systemen zählt: das Stereo-Quadrophonic- oder kurz SQ-System (Torick, 1998). Weitere bekannte 4-2-4-Systeme sind das Sansui-Quadrophonic-Sound(QS)-System, das Electro Voice 4-Channel Stereo (EV-4) und das Dynaquad System (Hönemann, 2011; Davis, 2003). Alle diese Systeme verwendeten die oben beschriebene quadrofone Lautsprecheraufstellung in Form eines Quadrats (Davis, 2003).

Doch die 4-2-4-Matrixsysteme waren mit einigen Problemen behaftet; von Nachteil war vor allem die geringe Kanaltrennung, die in der Regel nur etwa 3 dB betrug. „Zur Übertragung  $n$  unabhängiger Signale sind  $n$  Kanäle erforderlich. Überträgt man mit weniger Kanälen, so sind die rekonstruierten Signale nicht mehr unabhängig; sie unterscheiden sich deutlich von den Ursprungssignalen“ (Görne, 2015, S. 307). Die Folge ist, dass das Signal der vorderen Kanäle auf den hinteren Lautsprechern mit geringerer Lautstärke und gedrehter Phase auftaucht (Hönemann, 2011). Um dem abzuwehren, fügte man verstärkungsgesteuerte Logiksysteme hinzu, um dominante Signale besser zu isolieren. Diese waren aber nur teilweise effektiv und erzeugten manchmal unerwünschte Pump-Artefakte. Die Matrixsysteme neigten

„Er unterteilte die 360° des akustischen Umfelds in vier Hauptsektoren und legte mittels eines mathematischen Koordinatensystems feststehende Winkelwerte fest. Jede aufgezeichnete Raumrichtung erhielt so eine feste Phasendefinition in Bezug auf alle anderen Richtungen. Somit konnten die hinteren Kanäle in das vordere Klangpanorama eingeschachtelt und mittels eines Decoders, der die Phasenbeziehungen der einzelnen Signale erkannte, wieder separiert werden.“ (Hönemann, 2011, Absatz 12)

Neben Scheiber gab es weitere Ingenieur:innen, die solche Amplituden-Phasen-

außerdem dazu, sehr empfindlich auf Phasen- und Amplitudenunterschiede zu reagieren (Hönemann, 2011; Davis, 2003).

Angesichts der Tatsache, dass die 4-2-4-Systeme derart fehleranfällig waren, taten sich die japanische Firma JVC und das amerikanische Plattenlabel RCA zusammen und entwickelten 1971 die CD-4 (Compatible Discrete 4), auch „Quadradis“ genannt (Davis, 2003; Breh, 1973; Hönemann, 2001). CD-4 ist, im Unterschied zu den 4-2-4-Formaten, ein diskretes Quadrofonie-Format: Vier Audiokanäle wurden getrennt aufgenommen, gespeichert und wiedergegeben. Dabei wurde das Summensignal der vorderen und hinteren Kanäle im hörbaren Frequenzbereich gespeichert (20–15.000 Hz) und deren Differenzsignal mit einem 30 kHz-Trägersignal in den Ultraschallbereich gemischt (Davis, 2003). Der große Vorteil dabei: CD-4 war mit Stereoanlagen kompatibel und konnte auch mit einem normalen Stereo-Abspielgerät gehört werden. Die schlechte Kanaltrennung war nun dank diskret getrennter Kanäle auch kein Problem mehr (Hönemann, 2011). Doch gab es auch in diesem Fall ein entscheidendes Manko: Für CD-4 war ein spezieller Tonabnehmer nötig, der im Ultraschallbereich arbeiten kann, in dem das Differenzsignal liegt. Hierfür wurde eigens die Shibata-Nadel entwickelt, benannt nach ihrem Erfinder Norio Shibata (ebd.), die jedoch kaum verbreitet war (Davis, 2003).

Die Produzent:innen waren sich uneins, wie mit der Quadrofonie umgegangen werden sollte. Einige verwendeten die zusätzlichen Lautsprecher, um den Raumeindruck dem des Aufnahmeorts anzunähern, zum Beispiel dem eines Konzertsaals. Bei diesem Ansatz beließen die Produzent:innen die Band oder das Orchester im vorderen Stereofeld und brachten lediglich den Hall des Raumes auf die hinteren Kanäle. Ein anderer Ansatz war, die Zuhörer akustisch inmitten der Band zu platzieren (Holman, 2008).

Um im Quadrofonie-Format produzieren zu können, waren die Tonstudios jedenfalls benötigt, aufzurüsten. Neben der erforderlichen Aufstellung zusätzlicher Lautsprecher mussten die Aufnahmeräume von deutlich höherer Qualität sein, da die Quadrofonie die Effekte des Raums auf den Klang signifikant verstärkte (Schabbing, 2005). Dass die Produktionsstätten tiefer in die Tasche greifen mussten, war ein Grund für das Aussterben der Quadrofonie. Doch es gab weitere Gründe.

Wollte sich ein Endkonsument zu Hause etwas Quadrofonies anhören, brauchte er zusätzlich zwei Lautsprecher, und falls bereits zwei vorhanden waren, benötigte er einen Vorverstärker mit vier Kanälen und passende Decoder (ebd.). Denn die Kompatibilität unter den einzelnen Formaten war eher gering. Darum waren stets quadrofonische Effekte zu hören, wenn man ein Format mit dem Decoder eines anderen Formats abspielte (Breh, 1973). Und solche Decoder waren alles andere als günstig. Der SQ-Decoder „Space & Image Composer“ kostete 1981 mehr als 1.000 US-Dollar (Hönemann, 2011). Und wer sich vier Lautspre-

cher mit passendem Verstärker und Decoder zugelegt hatte, musste ausreichend Platz haben, um ein solches quadrofonies Lautsprecher-Quadrat aufzustellen. Und selbst wenn es genug Platz gab, war die ideale Abhörposition (eng. *Sweetspot*) nochmal schmaler, als es bei Stereo der Fall war (Schabbing, 2005).

Ein Problem für die Konsumenten waren nicht nur die hohen Kosten und der erforderliche Raum. Die Quadrofonie war mit erheblichen Wahrnehmungsproblemen verbunden. Für die Lokalisation von Klängen nutzt und verarbeitet der Mensch verschiedene Informationen, auf die ich in Kapitel 4 genauer eingehen werde. Zu diesen Informationen gehören die interaurale Pegeldifferenz (eng. *interaural level difference*, ILD), die interaurale Laufzeitdifferenz (eng. *interaural time difference*, ITD) und der HRTF (eng. *head-related transfer function*). Diese Funktion fasst den Frequenzgang für jeden Ankunftszeitpunkt zusammen, der mit einem Schallfeld in Verbindung steht, das mit dem Kopf, den äußeren Ohren, der Schulter usw. interagiert (Holman, 2008). Wie Philipson et al. in Hörversuchen herausfanden, ist eine Lokalisierung bei der Quadrofonie an den Seiten (etwa zwischen 60° und 150° Azimut) kaum möglich (2002). Vor sich nahm man hauptsächlich das Stereobild wahr und hinter sich ein engeres Stereobild, so Eberhard Sengpiel in einem Interview (Schabbing, 2005). Zusätzlich fand man heraus, dass die Phantomschallquellen durch das verbreiterte Stereodreieck vorne (sonst  $\pm 30^\circ$ , hier  $\pm 45^\circ$ ) weiter oben wahrgenommen wurden (Ratliff, 1974).

### 3.7 Surround

Während die Schwierigkeiten bei der Anpassung von Stereomedien an die Quad-Technologie noch einige Mühe bereitete, gab es im Bereich der Kinotontechnik größere Fortschritte. 1975 wurde mit Dolby Stereo der Grundstein des heutigen 5.1-Systems gelegt (Power, 2015). Bei Dolby Stereo, auch bekannt als Dolby Surround (Görne, 2015), handelt es sich um ein 4-2-4-Matrix-Verfahren, ähnlich denen, die bei der Quadrofonie benutzt wurden (Power, 2015; Davis, 2005). Anstelle des bereits vorhandenen Mono-Systems wurden bei Dolby Stereo zwei Spuren auf eine Filmspur gelegt. Dazu mussten vier Informationskanäle auf zwei Spuren untergebracht werden (Power, 2015). Dafür gab es neben den Stereo-Lautsprechern Links/Rechts nun einen zusätzlichen Center-Lautsprecher. Hinten wurde ein einzelner Surround-Lautsprecher aufgestellt (Power, 2015; Görne, 2015).

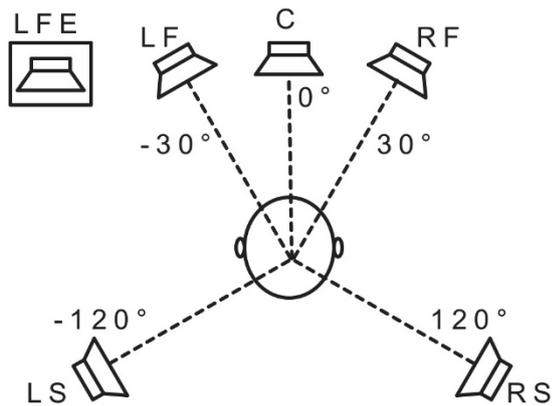
Neben der geringen Datenmenge ist die Kompatibilität mit dem Stereoformat ein weiterer Vorteil. Hierfür werden die vier Kanäle zu zwei Signalen zusammengefasst: Lt (Left total) und Rt (Right total) (Görne, 2015). Die anfängliche Reaktion der Industrie war eher zurückhaltend, bis im Mai 1977 der Film „Star Wars“ und einige Monate später „Close Encounters of the Third Kind“ das Potenzial des Systems demonstrierten. „Superman“ war 1978 der erste Film nach

„This Is Cinerama“, bei dem die Surround-Lautsprecher auf zwei Kanäle geteilt wurden (Holman, 2008). Im selben Jahr wurde die Grundmatrix von Dolby verfeinert, um die Kanaltrennung zu verbessern. Vier bzw. fünf Jahre drauf, 1982 und 1983, wurde die Verstärkungsregelung in den Decodern mit dem von Douglas Mandell entwickelten Pro Logic-Algorithmus aktualisiert. Dieses System nutzt eine dynamische Matrix mit konstanter Leistung anstelle einer variablen Verstärkung, die auf eine feste Matrix angewendet wird, um Pump-Artefakte zu reduzieren (Davis, 2003). „Die Kanaltrennung verbessert sich dadurch zwar signalabhängig auf bis zu 30 dB – trotzdem leiden Richtungsortung und Klangqualität unter der Matrizierung.“ (Görne, 2015, S. 307)

### 3.8 5.1 Surround

Es bestand mit anderen Worten die Notwendigkeit, zu einem diskreten System überzuwechseln, da die Probleme, die durch die Matrixcodierung verursacht wurden, zu groß waren. Neben den genannten Problemen der Klangqualität und Richtungsortung (Görne, 2015) war die Encodierung ursächlich dafür, dass die vom Center-Kanal wiedergegebene Sprache zu einer Modulation der Hintergrundmusik führte (Power, 2015); da es nur einen Surround-Kanal gab, ergaben sich in der Folge Probleme bei der Vermittlung des Raumklangs (Steinke, 1996). Als Lösung wurde ein diskretes Modell vorgeschlagen, dessen Anordnung wie folgt aussah: drei Frontkanäle, der Center (C) bei  $0^\circ$  und die beiden äußeren, links/rechts (L/R),  $\pm 30^\circ$ , zwei Surroundkanäle links und rechts (LS/RS) bei  $\pm 120^\circ$  und ein optionaler Subwoofer-Kanal (*Low Frequency Effects*, LFE) (Power, 2015; Görne, 2015).

Diese Anordnung wurde erstmals von Tomlinson Holman von Lucasfilm als 5.1-Kanal-System bezeichnet und im Juni 1991 in Ottawa als ITU-R Rec. BS.775-1 standardisiert (Davis, 2003). Das diskrete 5.1-System bot gegenüber anderen Mehrkanalformaten wie Stereo und Quadrofonie erhebliche Wahrnehmungsvorteile (Power, 2015). Der vordere Center-Kanal festigt das Stereoklangfeld für die Hörer:innen, die nicht in der Mitte des Raumes sitzen, und sorgt für eine bessere tonale Ausgewogenheit, als dies bei einer Phantomschallquelle zwischen dem linken und rechten Lautsprecher der Fall ist (Holman, 2008). Die hinteren Kanäle bieten darüber hinaus den Vorzug gegenüber Stereo, dass sie die Hörer:innen auf der Horizontalen einhüllen (Martin, 2005) und eine neue künstliche Raumakustik erzeugen (Theile, 1993). Nach Stuart (1996) erlaubt der Übergang von Stereo zu Mehrkanal auch eine bessere Trennung der Schallquellen mit niedrigeren Maskierungsschwellen. Doch blieben auch hier Probleme nicht aus, die in früheren Mehrkanalformaten auftraten. So führt nach Holman (2008) das Schwenken zwischen den vorderen und hinteren Lautsprechern über die



**Abbildung 3.6:** 5.1-Konfiguration

Seite des Zuhörers dazu, dass Klänge zwischen den vorderen und hinteren Kanälen hin und her springen, ähnlich wie bei der Quadrofonie.

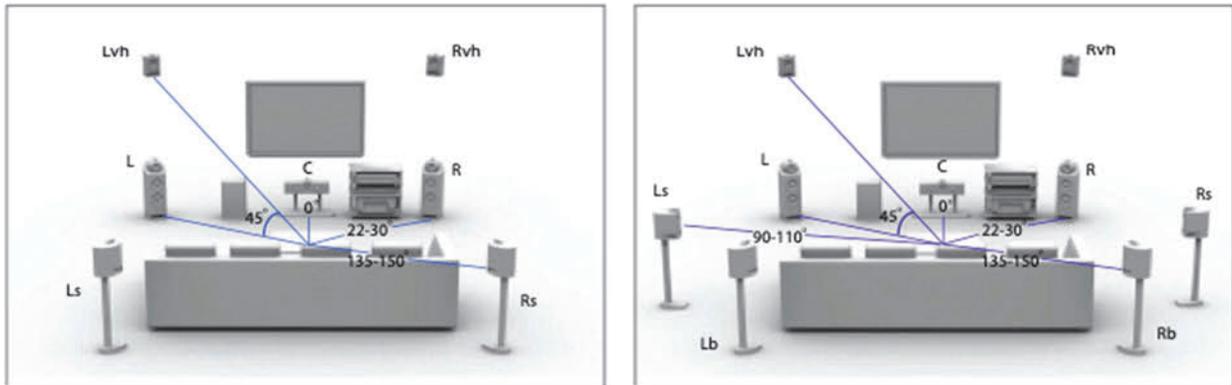
Die 5.1-Surround-Systeme wurden weiterentwickelt, weitere Lautsprecher hinzugefügt (Power, 2015). Beim 7.1-Surround-Verfahren SDDS (Sony Dynamic Digital Sound) von Sony zum Beispiel wurden zwei zusätzliche Lautsprecher für Mitte links/rechts (LC/RC) hinzugefügt (Görne, 2015). Darüber hinaus entwickelte man Upmixing-Verfahren wie das Dolby EX-System und Lexicons Logic 7-System, die es ermöglichen, 5.1-Material hochzumischen, um zusätzliche Kanäle zu erhalten. Diese umfassen in der Regel einen hinteren Kanal direkt hinter dem/der Zuhörer:in sowie zusätzliche Seitenkanäle, die darauf abzielen, das Problem der Erzeugung von Bildern seitlich des Zuhörers zu mildern (Power, 2015).

Chaffey und Shirley (2007) haben einige dieser Upmixing-Systeme getestet, um herauszufinden, ob sich die Lokalisierung gegenüber dem aktuellen 5.1-System verbessert. Ihr Ergebnis: Die Upmixing-Systeme verbessern die Lokalisierung seitlicher Quellen gegenüber der aktuellen 5.1-Konfiguration deutlich. Außerdem gab es Untersuchungen dazu, inwieweit die Erhöhung der Anzahl und die Positionierung der horizontalen Kanäle den räumlichen Eindruck optimiert. Das 5.1-System, so wurde ermittelt, ist eine Lautsprecheranordnung, die den räumlichen Eindruck eines diffusen Schallfeldes optimal zu reproduzieren vermag (Hiyama et al., 2002).

### 3.9 Surround mit Höhen

In den oben beschriebenen Surround-Systemen hat sich die Wiedergabe auf die horizontale Ebene beschränkt. Doch „wir nehmen unsere Umgebung akustisch nicht nur von vorne, hinten und seitlich auf, sondern auch über und unter uns“ (Fuchs, 2018, S. 22). Bei der Entwicklung von Surround-Systemen spielte seit den 1970er Jahren die Erwägung eine Rolle, Höhenkanäle einzubringen und 3D-Formate zu schaffen (Power, 2015). Eines der ersten kommerziellen Surround-Systeme mit Höhen war das Dolby Pro Logic IIz, das es als 5.1 und 7.1 mit Höhenlautsprechern gab. Bei diesem System gibt es zwei Höhenkanäle, die in 45°-Elevation über dem linken und rechten Stereo-Lautsprecher angebracht werden. Das Audiomaterial für

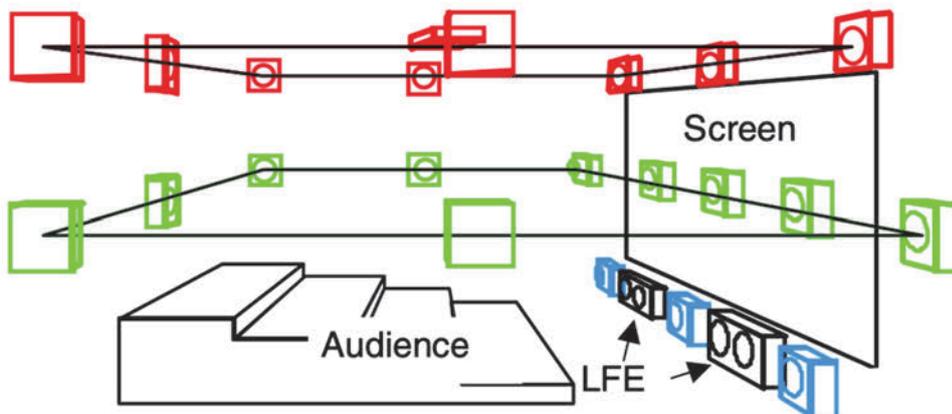
die beiden vorderen Höhenkanäle hat hauptsächlich ungerichteten diffusen Inhalt (Power, 2015). Tsingos et al. fanden bei Versuchen heraus, dass die Erweiterung der Höhenkanäle Hörer:innen ein intensiveres Hörerlebnis, mehr Tiefe und ein besseres Gefühl für den akustischen Raum bei Soundeffekten und Musik vermittelt (2011).



**Abbildung 3.7:** Pro Logic IIz in 5.1.2 (links) und 7.1.2 (rechts) (Tsingos et al., 2011, S. 2)

### 3.10 NHK 22.2

Zwei weitere nennenswerte Surround-Systeme mit Höhen sind das NHK 22.2 und das von Holman vorgeschlagene 10.2. Das NHK 22.2 ist für den Kinogebrauch entwickelt worden und verwendet zehn Lautsprecher auf Ohr-Höhe, neun über dem Publikum und drei unten vor dem Publikum (Power, 2015). Einige Studien zur Wahrnehmung legen nahe, dass die große Anzahl von Front- und Höhenkanälen die Lokalisierung und Tiefe über einen größeren Bereich verbessert, jedenfalls im Vergleich zum 5.1-System. Allerdings wurde auch festgestellt, dass dieser Effekt von der verwendeten Klangszene abhängt (Hamasaki et al., 2004).



**Abbildung 3.8:** NHK 22.2-Aufstellung, <https://bit.ly/3NUwrfD>



### 3.12 Ambisonics

Das System, das ich in meiner Arbeit verwendet habe, entstand in der Zeit der Quadrofonie. Nachdem die Scheiber-Matrix auf den Markt gekommen war, versuchten sich, wie oben beschrieben, andere mit der Entwicklung eigener Systeme. Der Mathematiker Michael Gerzon verfolgte dabei einen anderen Ansatz als seine Kolleg:innen (Davis, 2003). 1973 schlug er ein Kodierungsschema vor, das später Ambisonics genannt wurde. Es sah die Verwendung von sphärischen harmonischen Basisfunktionen vor, mit denen die Teile eines Schallfeldes kodiert werden konnten, die aus vielen verschiedenen Richtungen im Umfeld der Position des Hörers stammten (Gerzon, 1973). Gerzon hat dieses System für eine beliebige Anzahl von Basen (Ambisonics  $n$ -ter Ordnung) ausgearbeitet. Niedrigere Ordnungen von Ambisonics stellen eine stärkere Abschneidung der sphärischen harmonischen Zerlegung dar und ergeben eine geringere räumliche Präzision bei der Wiedergabe (Boren, 2017).

Die Ambisonics-Kodierung gehört zur Gruppe des szenenbasierten Audios (eng. *scene-based audio*, SBA). Anders als bei kanalbasiertem Audio (eng. *channel-based audio*, CBA), wo jedem Lautsprecher ein Kanal zugeteilt und explizit für ein Lautsprecher-Setup gemischt wird, ist szenenbasiertes Audio eine Audiodarstellung, bei der der Inhalt durch eine Reihe von Koeffizientensignalen dargestellt wird. Diese Koeffizientensignale sind die linearen Gewichte räumlicher orthogonaler Basisfunktionen. Die Szene kann durch Wiedergabe dieser Koeffizientensignale auf einer Lautsprecheranordnung oder auf Kopfhörern wiedergegeben werden. Die Produktion ist von der Wiedergabe entkoppelt, wobei Anzahl und Position der Ziellautsprecher keine Rolle spielen (ITU, 2018).

### 3.13 True Spatial

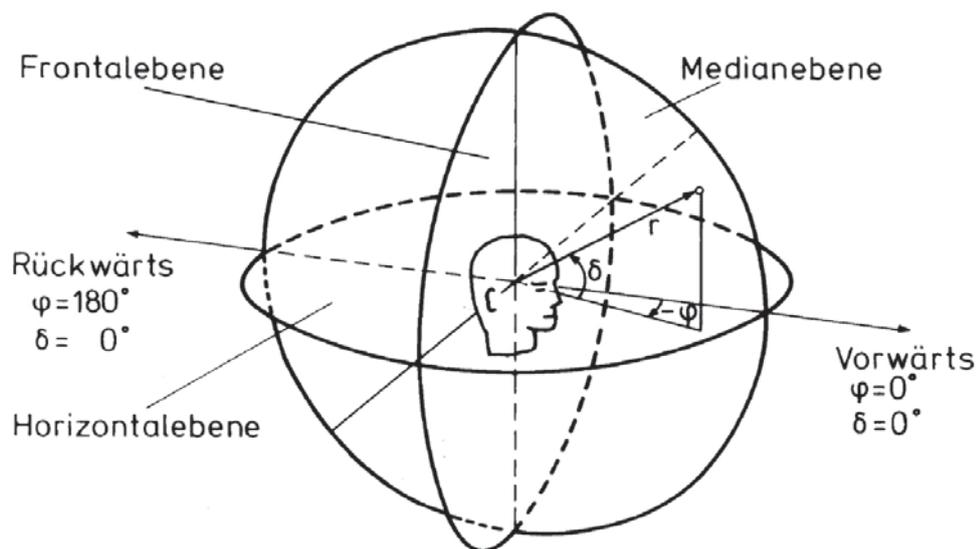
Was das in meiner Arbeit verwendete Ambisonics-Lautsprecher-setup ausmacht und was ich von nun an als „True Spatial“ bezeichne, ist die vollsphärische Wiedergabe, die durch Lautsprecher unterhalb des/der Hörenden erzielt wird. Dies ermöglicht die Platzierung von Klangobjekten sowohl auf der oberen wie auf der unteren Hemisphäre.

## 4 Räumliche Wahrnehmung

Ich befasse mich in dieser Arbeit mit Surround-Systemen, die sich nicht auf die horizontale Wiedergabe beschränken, sondern auch die vertikale Achse einbeziehen. Daher gilt es zu erörtern, wie diese raumbezogenen Informationen vom menschlichen Gehör verarbeitet werden. Die Kenntnisse über die dabei wirkenden Mechanismen beeinflussen auch meine akustische Gestaltung.

### 4.1 Richtungswahrnehmung

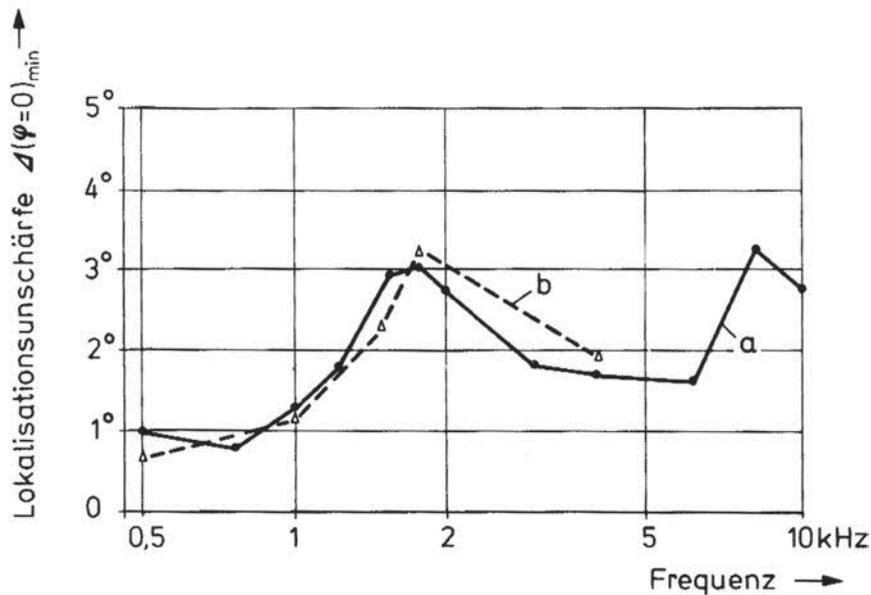
„Richtungswahrnehmung oder Lokalisation ist die Zuordnung der vom Gehör zum Bewusstsein gebrachten Hörereignisrichtung zur Schallereignisrichtung“ (Dickreiter, 2008, S. 105). Die Fähigkeit, Schallereignisse auf Richtung und Entfernung einzuschätzen und den uns umgebenden Raum akustisch zu erfassen, wird durch das binaurale Hören ermöglicht. Doch besitzen auch Menschen diese Fähigkeit, die nur mit einem Ohr hören können (monaural) (ebd.; Fuchs, 2018). Um die Lokalisation zu beschreiben, verwendet man ein kopfbezogenes Koordinatensystem, bestehend aus der horizontalen Ebene, der Vertikal- oder auch Medianebene und der Frontalebene (Dickreiter, 2008).



**Abbildung 4.1:** Kopfbezogenes Koordinatensystem;  $r$  Entfernung,  $\delta$  Erhebungswinkel,  $\varphi$  Seitenwinkel (Blauert, 1974, S. 11)

Während die Horizontalebene den Raum in Oben und Unten teilt, unterteilt die Medianebene den Raum in Links und Rechts. Die Frontalebene spaltet den Raum in Vorne und Hinten

(Sengpiel, o.D.). Nach Blauert (1996) ist Lokalisierung das Gesetz oder die Regel, nach der der Ort eines Hörereignisses mit einem bestimmten Attribut – oder auch mehreren Attributen – eines Schallereignisses in Beziehung steht. Kommt es dabei zu gewissen Veränderungen einzelner oder auch mehrerer dieser Attribute, bewirkt das eine Lokalisationsunschärfe, die den akustischen Raum weniger differenziert klingen lässt. Das menschliche Gehör braucht bestimmte Informationsquellen, um die Raumposition einer Schallquelle auszumachen. Diese Informationsquellen bieten bestimmte Positionsreize, die sich in zwei Arten unterteilen lassen: binaurale, bei der beide Ohren beteiligt sind, und monaurale, die von einem Ohr abhängig sind (Goldstein, 2015).



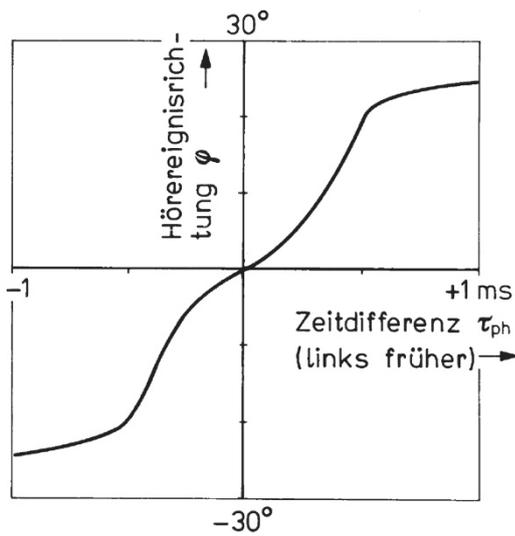
**Abbildung 4.2:** Lokalisationsunschärfe in Abhängigkeit von der Frequenz für  $\Delta(\varphi = 0)_{min}$  und  $\delta = 0$ ; Kurve a: Dauertöne, Kurve b: Gauß-Töne (Blauert, 1974, S. 32)

## 4.2 Horizontale Ebene

Das menschliche Gehör ist für die räumliche Wahrnehmung auf der Horizontalebene optimiert und kann dort Richtungsänderungen von nur  $1^\circ$  feststellen (Görne, 2015). Dieser Wert wird allerdings nur dann erreicht, wenn es sich um eine einzelne impulsartige Schallquelle mit kleinem Seiten- und Erhebungswinkel handelt (Blauert, 1974). Für die Verortung nutzen wir hauptsächlich die Differenzen jener Signale, die zwischen den beiden Ohren ankommen. Dabei ist zwischen zwei spezifischen Differenzen zu unterscheiden: den interauralen Laufzeitdifferenzen (eng. *Interaural Time Differences*, ITD) und den interauralen Pegeldifferenzen (*Interaural Level Differences*, ILD) (Görne, 2015).

### 4.3 Interaurale Laufzeitdifferenz (ITD)

Die ITD ist die Zeitdifferenz, die der Schall benötigt, um beide Ohren zu erreichen. Trifft zum Beispiel ein Schallereignis zuerst auf das linke Ohr, muss die Schallquelle linksseitig liegen. Anders als bei der ILD ist die ITD kaum distanzabhängig (Blauert, 1996). Wie John William Strutt (1842–1919), besser bekannt als Lord Rayleigh, schon in einem Experiment Ende des 19. Jahrhunderts feststellen musste, kann es bei der ITD zu Verwechslungen zwischen Vorne und Hinten kommen.



**Abbildung 4.3:** Summenlokalisationskurve für Laufzeitdifferenz für Breitbandsignale, Zuhörerkopf fixiert (Blauert & Braasch, 2008, S. 101)

Trifft Schall beispielsweise vorne links auf das Ohr, erzeugt er die gleiche ITD wie ein Schall, der von hinten links kommt. Derselbe Fall tritt ein, wenn der Schall in 0° bzw. 180° Azimut auf den/die Hörer:in trifft (Görne, 2015). Geht man von einem durchschnittlichen Ohrabstand von 17 cm und somit von einem maximalen Schallumweg von 21 cm aus (Dickreiter, 2008; Görne, 2015), so beträgt die maximal resultierende Laufzeitdifferenz

$$\tau = \frac{r}{c} = \frac{0,21 \text{ m}}{343,2 \frac{\text{m}}{\text{s}}} \approx 0,612 \text{ ms.}$$

Das entspricht wiederum einer Frequenz von

$$f = \frac{1}{\tau} = \frac{1}{0,612 \text{ ms}} \approx 1,63 \text{ kHz.}$$

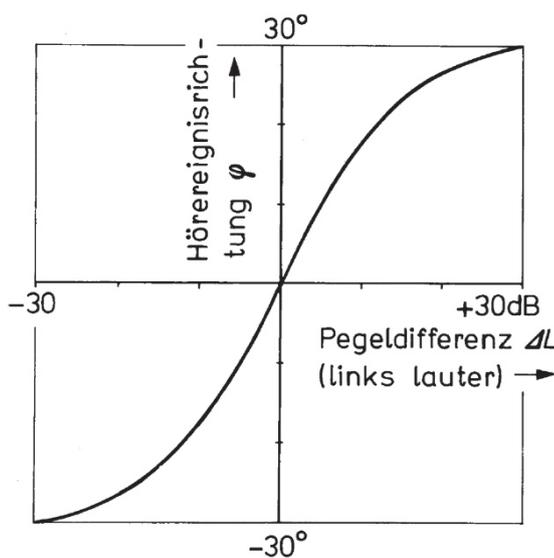
Trifft nun beispielsweise ein reiner Sinuston mit  $f \geq 1,63 \text{ kHz}$  auf unsere Ohren, kann unser Gehör keine Zeitdifferenz wahrnehmen mit der Folge, dass es zu falschen Verortungen kommen kann (ebd.). „Dieser Fall der mehrdeutigen Ortung ... tritt ein, wenn das Signal hochfrequent [ $f \geq 1,63 \text{ kHz}$ ] ... und quasistationär ist [wie z. B. bei Flöte und Violine]“ (Görne, 2015, S. 127). Das bedeutet, dass die Periodendauer des Signals größer sein muss als die Laufzeitdifferenz, damit man die Schallquelle eindeutig über die ITD orten kann. Denn sonst lässt sich kein eindeutiger Unterschied der Phasenlage der Signale feststellen.

Doch auch bei Schallereignissen mit Frequenzen jenseits der 1,63 kHz vermag unser Gehör, sofern es sich um ein zeitveränderliches Signal handelt (Görne, 2015), Zeitdifferenzen wahrzunehmen (Dickreiter, 2008). Bei solchen Signalen, deren Schwingungsperiode sich nicht identisch wiederholt, entnimmt unser Gehör die Zeitdifferenz aus dem Vergleich

der Hüllkurven der Signale (ebd.). Bei akustischen Ereignissen in unserem Alltag, wie zum Beispiel Sprache und Musik, handelt es sich um eben solche zeitveränderlichen Signale, die breitbandig sind und eine „stark strukturierte“ (Dickreiter, 2008, S. 107) Hüllkurve besitzen. Dem Menschen fällt die Lokalisation breitbandiger und impulsartiger Schallsignale deutlich leichter als die eines Sinussignals (ebd.).

#### 4.4 Interaurale Pegeldifferenz

Als zweites Maß der Lokalisation in der Horizontalen dient uns die interaurale Pegeldifferenz, die Blauert (1974) als die wichtigere der beiden Mechanismen wertet. Bei dieser Art der Richtungsortung werden die Pegelunterschiede der auf die Ohren treffenden Signale verglichen (Dickreiter, 2008). Die Pegelunterschiede werden durch Abschattungen des Kopfes und der Ohrmuschel selbst verursacht. So entsteht an dem schallzugewandten Ohr ein Druckstau mit mehr Pegel als am abgewandten Ohr, an dem ein Schallschatten entsteht (Görne, 2015). Dank der sich so ergebenden Pegeldifferenz zwischen den beiden Ohrsignalen ist unser Gehör in der Lage, die Richtungsinformationen über den gesamten hörbaren Frequenzbereich zu entnehmen (Dickeiter, 2008).



**Abbildung 4.4:** Summenlokalisationskurve für Pegeldifferenz für Breitband-signale, Zuhörerkopf fixiert (Blauert & Braasch, 2008, S. 101)

Doch diese Differenzierung ist frequenzabhängig. Dabei entsteht durch die Beugungserscheinung des Schalls um den Kopf, die zu starkem Übersprechen führt, bei Frequenzen unterhalb von 300 Hz „praktisch keine Pegeldifferenz mehr“ (Dickreiter, 2008, S. 107). Anders ist das bei 2 kHz, wo die Empfindlichkeit am höchsten ist (Blauert, 1974). Ein allgemeingültiger Zusammenhang zwischen der Pegeldifferenz und der Schallereignisrichtung lässt sich jedoch nicht feststellen, da dieser von der Art des Signals abhängt (Dickreiter, 2008). Blauert (1974) zeigte in Versuchen, dass die akustische Lokalisationsschärfe neben der Frequenz auch stark pegelabhängig ist. So kommt es ab 90 dB zu einer Ermüdung des Gehörs, in deren Folge Höreignisse näher der Medianebene verortet werden.

Herrscht wiederum ein Signalpegel von unter 30 dB, nimmt die Unschärfe ebenso zu.

#### 4.5 Zusammenspiel aus interauralen Laufzeit- und Pegeldifferenzen

Wie eben beschrieben, arbeiten die beiden Mechanismen der Lokalisation auf der horizontalen Ebene weitgehend unabhängig voneinander. Die interaurale Laufzeitdifferenz ist maßgeblich, wenn es um die Lokalisation von Signalen mit Frequenzen unterhalb von 1,63 kHz geht. Treten Hörereignisse mit Frequenzkomponenten über 1,63 kHz auf, wie es bei Sprache, Musik und Geräuschen häufig der Fall ist, wird die Laufzeitdifferenz über die Hüllkurve ermittelt und die Richtung des Hörereignisses hauptsächlich über die interaurale Pegeldifferenz bestimmt (Dickreiter, 2008). Sowohl für die ITD als auch für die ILD lässt sich aber festhalten, dass die Lokalisationsschärfe zu den Seiten hin ungenauer wird (Blauert, 1974).

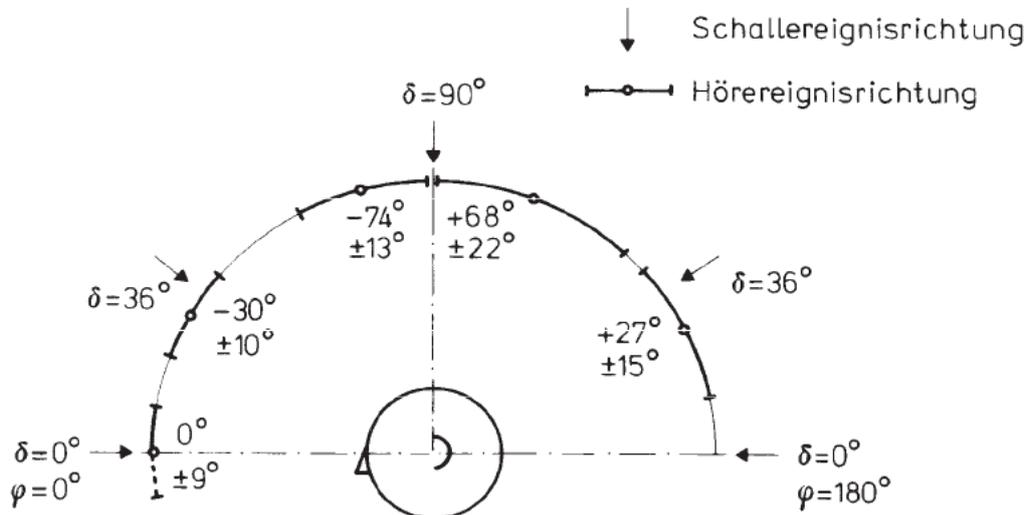
„Ist zu einem Punkt im dreidimensionalen Raum die Entfernung (oder auch Laufzeit bzw. Pegel) von 2 Punkten bestimmt (die Ohren), deren Position bekannt ist, kann die Position des gesuchten Punktes nur auf einer Kreisbahn, die parallel zur Medianebene liegt, bestimmt werden. Wo der Punkt tatsächlich auf der Kreisbahn liegt, kann nicht festgestellt werden. Beim Hören kommt es zusätzlich zu dem Problem, dass nur die Differenz von Laufzeit und Pegel zwischen den Ohren bestimmt werden kann. Dadurch ergibt sich zusätzlich eine Hyperbelfunktion für Entfernungen, bei denen die Differenz gleich ist. So wird aus der Kreisbahn ein Hyperbelkegel, auf dessen Fläche das Hörereignis liegt.“ (Falke, 2016, S. 26)

Dieser Hyperbelkegel ist weithin als „cone of confusion“ bekannt. Tritt eine Lokalisation auf diesem Kegel ein, kommt Blauerts (1974) Drehtheorie ins Spiel. Danach vollzieht der Mensch kleine Kopfbewegungen, immer in Richtung des Hörereignisses, um auf diese Weise genauer zu lokalisieren. In späteren Forschungen stellten Blauert & Braasch (2008) außerdem fest, dass eine Auslenkung auf der horizontalen Ebene über Lautsprecher nicht größer als  $\pm 30^\circ$  sein kann, da sonst der Zusammenhang zwischen der ITD/ILD und dem Hörereignis verlorengeht. In der Folge beginnt das Signal zwischen den Lautsprechern zu springen.

#### 4.6 Vertikale Ebene

Schallquellen lassen sich auch auf der vertikalen, der Medianebene lokalisieren, allerdings funktioniert die Verortung hier auf grundlegend andere Weise. Bleibt eine Kopfbewegung aus, ergibt sich durch die Symmetrie der Medianebene keine Laufzeit- und Pegeldifferenz der am Ohr anliegenden Signale (Dickreiter, 2008). Blauert (1974) fand bei Experimenten heraus,

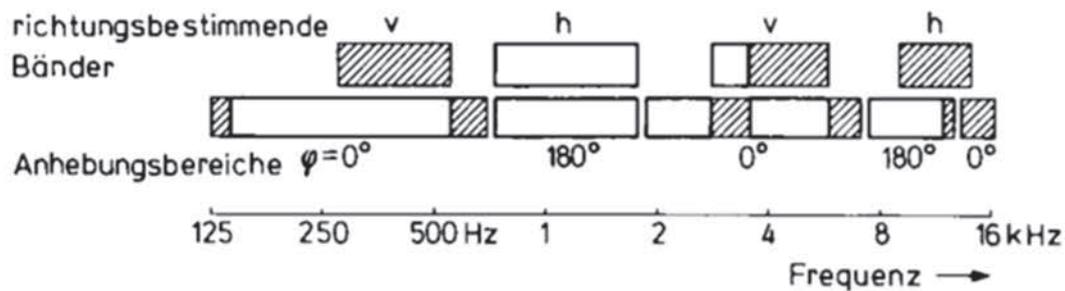
dass die maximal zu erreichende Lokalisationsschärfe bei  $0^\circ$  auf der Medianebene  $\pm 4^\circ$  beträgt<sup>1</sup>, wenn der Kopf weder seitlich gekippt oder gedreht wird. Ist eine seitliche Bewegung des Kopfes möglich, lässt sich die Höhenlokalisierung grob über die ITD und ILD ermitteln, zumindest in dem Maße, wie der Mensch erkennt, dass ein Hörereignis in der oberen Hemisphäre stattfindet (ebd.).



**Abbildung 4.5:** Minimale Lokalisationsunschärfe auf der Medianebene für Sprache eines bekannten Sprechers; Kopf fixiert (Blauert, 1974, S. 36)

Für eine genauere Verortung von Hörereignissen in der Vertikalen spielt die Frequenzverzerrung bzw. die Klangfarbe (eng. *Timbre*) eine entscheidende Rolle (Blauert, 1974). Über die HRTF (eng. *Head Related Transfer Function*) lässt sich beschreiben, welche Frequenz bei welchem Raumwinkel wie verzerrt wird. Diese Funktion, die bei jedem Menschen individuell unterschiedlich ausgeprägt ist, vermittelt Informationen über Überlagerungen von Direktschall und die Reflexionen vor allem der Ohrmuschel und den daraus entstehenden richtungsabhängigen Kammfilter (Görne, 2015). Für die räumliche Wahrnehmung spielt die Verarbeitung dieser spektralen Veränderungen im Zusammenhang mit dem Prozess des Lernens, mit Erfahrungen, besonders in der Kindheit (Goldstein, 2015), und mit der Vertrautheit eine sehr wichtige Rolle (Dickreiter, 2008; Blauert, 1996). Die Präzision der Lokalisation in der Medianebene ist deutlich schlechter als auf der Horizontalen, ihre Spanne beträgt laut Blauert (1974) zwischen  $4^\circ$  und  $22^\circ$ .

<sup>1</sup> Die  $\pm 4^\circ$  sind bei dem Versuch mit weißem Rauschen erzielt worden. Neben weißem Rauschen ist die Lokalisationsschärfe auch mit Sprache ( $\pm 9^\circ$ ) und anderen akustischen Reizen ( $\pm 17^\circ$ ) getestet worden (Blauert, 1974).



**Abbildung 4.6:** Die richtungsbestimmenden Bänder und der Anhebungsbereich auf der Frequenzachse (Blauert, 1974, S. 92)

Dabei hängt die Genauigkeit sowohl vom Erhebungswinkel als auch vom Schallsignal selbst ab. Blauert (1974) stellte bei Versuchen fest, dass je nach Einfallswinkel bestimmte richtungsbestimmende Frequenzbänder angehoben werden. So lokalisierten alle Versuchspersonen eine Schallquelle oben, wenn der Pegel bei 8 kHz angehoben wurde. Bei der Manipulation weiterer Frequenzbänder zeigte sich eine Veränderung der Wahrnehmung bezüglich vorne/hinten. Blauert (1996) betont jedoch, dass dieser Befund keine Allgemeingültigkeit beanspruchen könne, da der Test mit nur zehn Proband:innen durchgeführt wurde. Allerdings schreibt auch Dickreiter, dass „eine spezifische Kombination von spektralen Gipfeln und Einbrüchen die Hörereignisrichtung bestimmt“ (Dickreiter, 2008, S. 109). Bei den Versuchen Blauerts kam jedenfalls (1974) heraus, dass der Mensch nur breitbandige Signale, die mindestens 1–2 Terz umfassen, in der Medianebene zu lokalisieren vermag, da sich sonst keine Veränderung des relativen Pegels über die Frequenzen ermitteln lässt.

Dieses Wissen führt zu der Erkenntnis, dass ein Panning auf der Vertikalen mittels Laufzeit- oder Pegeldifferenzen nicht möglich ist, da das Hörereignis sonst zwischen den oberen und unteren Lautsprechern zu springen beginnt.

„Das Erkennen einer Höhenlage zwischen den Lautsprechern ist insbesondere nicht möglich, da die monaurale Frequenzverzerrung für die jeweilige Winkelrichtung nicht korrekt ist. ... Zur Positionierung braucht es daher mehrere Lautsprecher in der Vertikalen, die über Schallfeld-Syntheseverfahren, wie Higher Order Ambisonic (HOA) oder Wellenfeldsynthese WFS, angesteuert werden. Dabei muss an dem Ohr des Zuhörers ein Schallfeld erzeugt werden, das dem einer realen Schallquelle ähnlich ist“ (Falke, 2016, S. 23 f.).

Bei einer experimentellen Mischung Falkes et al. (2014) mit HOA, die entsprechend arbeitete, konnte festgestellt werden, dass eine Genauigkeit der Lokalisation auf der Vertikalen mittels

Höhenpanning von  $\pm 20^\circ$  möglich ist. „Mit einer größeren Anzahl Lautsprechern in der Vertikalen ließe sich die Lokalisationsschärfe sicherlich verbessern und dadurch ein Springen vermeiden“ (Falke, 2016, S. 24).

Doch ob mehr Lautsprecher vorteilhaft für die Lokalisation sind, ist unklar und umstritten, jedenfalls in der horizontalen Ebene. So deuten Versuche von Martin (2002) darauf hin, dass der Sweet Spot durch mehr Lautsprecher schmaler wird und außerdem eine größere Zahl von Lautsprechern mehr Übersprechungen bewirken, was Klangfärbungen verursacht. Doch vertreten andere die Ansicht, dass mehr Lautsprecher auf der Horizontalen für eine stabilere Lokalisation der Seiten sorgen würde. Wie erwähnt fand Blauert (1996) heraus, dass die Lokalisation an seitlichen Positionen sehr ungenau ist – einen Befund, den Theile & Plenge (1977) mit ihren Untersuchungen untermauerten. Lund (2000) empfahl nach eigenen Versuchen, die Anzahl der Lautsprecher in einem Surround-System rückseitig zu erhöhen, denn dies würde die Lokalisierung von Phantombildern an den Seiten stabiler machen.

#### 4.7 Wahrnehmung der Entfernung

Entscheidend für die Abschätzung der Entfernung sind hauptsächlich die Erfahrung und die Bekanntheit des Signals. Von Belang sind dabei die Lautstärke und spektrale Veränderungen des Schallsignals, was wiederum bedeutet, dass das Entfernungshören sehr fehleranfällig ist (Blauert, 1974; Dickreiter, 2008). Bei einer Entfernungsverdopplung nimmt ein allseitig abstrahlendes Schallsignal um 6 dB ab, bei gerichtetem Schall ist dieser Wert geringer. Die Lautstärke der Quelle des Signals erschließt sich der Mensch aus seiner Erfahrung und im Zuge eines Lernprozesses.

Doch lässt sich eine Entfernung allein auf Basis der Lautstärke eines Signals nur sehr ungenau abschätzen, eine solche Einschätzung ist zuverlässiger bei Entfernungen zwischen 3 und 15 m (Blauert, 1996). Hinzu kommt die spektrale Veränderung des Signals, die ausschlaggebend ist bei Entfernungen, die größer sind als 15 m und kleiner als 3 m (ebd.). Während tiefe Frequenzen ungerichtet abstrahlen, werden höhere Spektralanteile mit zunehmender Frequenz gerichtet. Dadurch entsteht eine Klangfärbung, die deshalb deutlich verstärkt wird, weil der Mensch für die Wahrnehmung tiefer Frequenzen auf einen höheren Lautstärkepegel angewiesen ist (Dickreiter, 2008). Durch die Luft werden die hochfrequenten Anteile stärker gedämpft als die tieferen. Mit der Distanz ändert sich folglich das Spektrum des Signals, das am Ohr anliegt.

Eine genaue Entfernungseinschätzung ist bei Distanzen jenseits von 15 m nicht mehr möglich (Blauert, 1974). Im Unterschied zu Entfernungen über 15 m entstehen bei geringeren Distanzen, unter 25 cm, die spektralen Veränderungen durch Beugungs- und Abschattungs-

effekte des Kopfes. Ist die Entfernung eines Hörereignisses jedoch kleiner als der Kopfradius, kann es zur sogenannten „Im-Kopf-Lokalisation“ (IKL) kommen, bei der das Hörereignis im Kopf verortet wird. Dieser Effekt tritt allerdings eher bei der Übertragung mittels Kopfhörer auf (ebd.).

In geschlossenen Räumen spielt schließlich auch noch das Verhältnis des direkten Schalls zum diffusen eine Rolle. Wie erwähnt verändert sich der Lautstärkepegel der Schallquelle mit der Distanz. Anders ist das bei diffusem Schall, der unabhängig von der Schallquellenentfernung gleich bleibt. Da dieses Verhältnis durch die Nachhallzeit und die Größe des Raums verändert wird, spielen auch hier Erfahrung und Lernen bei der Einschätzung der Distanz eine entscheidende Rolle (Dickreiter, 2008).

#### 4.8 Präzedenzeffekt

In der bisherigen Untersuchung haben wir den Schall unabhängig vom Raum betrachtet, in dem das Hörereignis stattfindet. Doch der Mensch hört immer räumlich (Blauert, 1974), gewöhnlich vernehmen wir Schall stets in einer gewissen Umgebung. Das kann ein kleines Zimmer sein, eine Kirche oder die freie Natur. Der Schall wird von Objekten in dieser Umgebung reflektiert (indirekter Schall), bevor er auf das Ohr trifft (Goldstein, 2015).

In einer Umgebung, in welcher der indirekte Schall im Mittel mehr Pegel aufweist als der Direktschall und dazu noch aus vielen unterschiedlichen Richtungen kommt, ist der Mensch jedoch ebenfalls in der Lage, die Schallquelle zu verorten (Görne, 2015). Die Lokalisation funktioniert hier auf Basis des Präzedenzeffekts. Die Richtung, aus der die erste auf das Gehör treffende Wellenfront kommt, ist die Richtung der Schallquelle (deshalb auch: Gesetz der ersten Wellenfront). Wie groß der Anteil und die Dauer des indirekten Schalls sind, wird von der Nachhallzeit bestimmt.

Liegt die Verzögerung des direkten Schallsignals in einem Bereich von  $\tau \leq 0,6 \text{ ms}$ , wird die Schallquelle durch die ITD zwischen den Richtungen des Direktschalls und des indirekten Schalls lokalisiert (Phantomschallquellenortung). Befindet sich das verzögerte Signal in einer Bandbreite zwischen  $0,6 \text{ ms} \leq \tau \leq 30 \text{ ms}$ , werden die reflektierten Schallsignale nicht separat wahrgenommen, sie lassen den Klang vielmehr lauter und voller wirken und veranschaulichen uns den Raum in Form und Größe. Bei  $\tau \geq 30 \text{ ms}$  nehmen wir den reflektierten Schall als Echo wahr. Der zeitliche Grenzwert des Echos hängt wiederum vom Pegel und dem Signal an sich ab (Görne, 2015). Bei einem einzelnen Klicken kann die Schwelle, unter bestimmten Umständen, schon bei 2 ms liegen. Bei Sprache liegt sie etwa bei 20 ms, bei Musik gar bei über 50 ms (Blauert, 1996). Ist die Anzahl der Reflexionen durch den Nachhall zu groß, entstehen Überlagerungen, die eine Lokalisation unmöglich machen (Blauert, 1974).

## 4.9 Räumlichkeit und Umhüllung

Die bisherige Diskussion konzentrierte sich auf die Lokalisierung einer Schallquelle an einem Punkt im Raum. Quellen in der realen Welt können jedoch größer und verstreuter sein als Punktquellen. In diesem Fall haben wir es mit diffusen Schallquellen zu tun, die bei Nachhall und in diffusen Umgebungen auftreten (Holman, 2008). Wir sind ständig von Geräuschen und Klängen umgeben, doch diese klangliche Umhüllung nehmen wir erst dann bewusst wahr, „wenn es außergewöhnlich klingt, wenn es in einer Orientierungsreaktion unsere Aufmerksamkeit auf sich zieht (*Cocktailparty-Effekt*) – oder wenn wir bewusst in unsere Umgebung hineinhören“ (Görne, 2017, S. 162). Dabei kann das Sounddesign die Aufmerksamkeit und den Fokus gezielt auf bestimmte Ereignisse lenken (ebd.).

Für die Beschreibung von Raumklang werden zwei Begriffe verwendet: Räumlichkeit (eng. *Spaciousness*) und Umhüllung (eng. *Envelopment*). Diese Begriffe werden oft gleichbedeutend verwendet, sind aber voneinander zu unterscheiden: Räumlichkeit bezieht sich auf die Ausdehnung des darzustellenden Raums und kann über ein 2- oder 5-Kanal-System gehört werden. Sie wird prinzipiell durch das Verhältnis von Direktschall zu Reflexionen und Nachhall bestimmt. Bei einem 2-Kanal-System ist das Schallfeld in der Regel auf den Bereich zwischen den Lautsprechern beschränkt; dabei stellt sich der Eindruck von Räumlichkeit ein, nämlich dass zwischen den Lautsprechern ein physischer Raum entsteht. Die Tiefendimension wird zwar mit einbezogen, doch erstreckt sich die Tiefe wiederum nur auf den Bereich zwischen den Lautsprechern. Umhüllung hingegen bezieht sich auf den Eindruck, von Klang umgeben und somit in den Raum der Aufnahme eingebunden zu sein. Zur Wiedergabe ist ein Mehrkanal-Soundsystem erforderlich. Zweikanal-Stereo kann das Gefühl erzeugen, in einen Raum jenseits der Lautsprecher zu schauen; Mehrkanal-Stereo kann das Gefühl erzeugen, dort zu sein (Holman, 2008; Görne, 2017).

Gerade Ambisonics eignet sich aufgrund der Erfassung und Kodierung von 360-Grad-Klangfeldern gut für die Erzeugung einer einhüllenden Umgebung. Für die Produktion von Räumlichkeit und Umhüllung ist aber vor allem auch auf die Auswahl oder Schaffung von geeignetem Klangmaterial zu achten. Dies zeigt sich nach Field (2002) daran, dass es oft ausreicht oder sogar vorzuziehen ist, Quellen für die Surround-Kanäle zu verwenden, die zwar mit den vorderen Kanälen verbunden sind, in Wirklichkeit aber separate Aufnahmen sind. Das Raumgefühl, das bei diesem Ansatz erzeugt würde, sei nur zum Teil auf die fehlende Korrelation zwischen den Quellen zurückzuführen: Der Hauptwahrnehmungseffekt würde durch die Auswahl der Klänge selbst erzeugt. Würde man statt nur eines einzigen Klangfelds, das von einer einzigen Aufnahme abgeleitet ist, diskrete Quellen verwenden, um

ein die Hörer:innen umgebendes Feld zu schaffen, biete dies die Möglichkeit, die Realität kreativ zu bearbeiten.

#### **4.10 Einflüsse**

Bei der Betrachtung der räumlichen Wahrnehmung spielt der gesamte Wahrnehmungsapparat eine Rolle, wobei sich unterschiedliche Sinne wechselseitig beeinflussen. Insofern ist auch die visuelle Wahrnehmung und ihre Wechselwirkung mit der auditiven Wahrnehmung zu beachten (Scherzer, 2010; Flückiger, 2006).

„Wird der Raum zu einer narrativen Größe, muss auch z. B. im Hörspiel sowohl der Wiedergaberaum als auch die Situation der Präsentation in das inszenatorische Kalkül mit einbezogen werden. Für diese Hörsituation muss der Präsentationsraum also entweder mit inszeniert oder aus dem Bewusstsein ausgeblendet werden. Die Situation des Hörers selbst und seine räumliche Position in Relation zum fiktiven narrativen Raum muss in jedem Fall berücksichtigt werden, wenn der akustische Raum zu einer narrativen Größe wird.“ (Scherzer, 2010, S. 38)

Wie bereits erwähnt, ist auch die Erfahrung entscheidend und eng mit der Fähigkeit zur Lokalisierung verknüpft. So blicken wir intuitiv in die Höhe, wenn ein Vogel zu hören ist. Einfluss hat auch die Plausibilität: Ein Flugzeug, das von unten kommt, erscheint uns unwahrscheinlich. Die Einschätzung der Wahrscheinlichkeiten bestimmter Szenarien ist durch soziale, kulturelle, historische und anderer Einflüsse in unserem Leben geprägt (Scherzer, 2010).

## 5 Immersion

In dieser Arbeit geht es um immersive Gestaltung. Was aber ist Immersion überhaupt? Der Begriff wird in einer Vielzahl von Texten verwendet, die unterschiedliche Wahrnehmungserfahrungen beschreiben (Agrawal et al., 2019). Meist wird er gleichgesetzt mit Realismus, Natürlichkeit, Präsenz und dem Gefühl, von Geräuschen, Tönen, Bildern oder anderen erfahrbaren Realitäten umgeben zu sein, mit anderen Worten: Die Begrifflichkeit ist eher vage (McMahan, 2003). Zudem werden Immersion und Umhüllung häufig synonym verwendet, zu Unrecht (Scherzer, 2010). Der Schwede Jan Berg (2009) zeigte die verworrene Natur des Konzepts der Immersion für Audioanwendungen auf und wies nach, dass neben dem Klang auch andere Modalitäten zur Immersion beitragen. Immersion ist demnach ein komplexes Phänomen, das mehr umfasst als die Tatsache, dass Hörer:innen von irgendeiner Art von Klang umgeben sind.

Abgesehen von Audioanwendungen wurde Immersion in vielen weiteren Bereichen erforscht, so zum Beispiel in Videospiele, Virtual Reality (VR), Musik, Film und Literatur (Agrawal et al., 2019). Da der Begriff „Immersion“ zur Beschreibung einer Vielzahl unterschiedlicher Erfahrungen genutzt wird und es keinen Konsens über die Definition dieses Begriffs gibt, kann dies dazu führen, dass die zu untersuchende Theorie mit den angewandten Forschungsmethoden nicht zusammenpasst (Nilsson et al., 2016). Darüber hinaus besteht mit dem Aufkommen der virtuellen Realität und der Verwendung austauschbarer Begriffe wie „Präsenz“, „Involvierung“ und „Engagement“ die Gefahr, dass Konzepte verworren und unscharf werden (Agrawal et al., 2019). Nach Agrawal et al. (2019) lassen sich im Wesentlichen zwei Sichtweisen auf das Phänomen der Immersion unterscheiden: eine, die sich auf den psychologischen Zustand einer Person richtet, und eine zweite, die eine objektive Eigenschaft einer Technologie oder eines Systems fokussiert.

### 5.1 Immersion als psychologischer Zustand

Psychologische Immersion wird als der psychologische Zustand von Nutzer:innen verstanden, die involviert, absorbiert, beteiligt oder versunken sind (Lombard et al., 2009). Laut Alison McMahan (2003) entsteht Immersion durch die mentale Absorbierung der Nutzer:innen von der Welt. Die Rolle des Systems/der Technik sei dabei nebensächlich, es komme lediglich auf das Individuum und die Art der Erzählung etwa in einem Computerspiel an. Ähnlich sieht das Janet H. Murray (2017), der zufolge Immersion ein metaphorischer Begriff sei, der sich von der physischen Erfahrung des Eintauchens in Wasser ableite. Dabei würden wir bei

einer psychologisch-immersiven Erfahrung das gleiche Gefühl erstreben, das wir bei einem Sprung ins Meer oder in ein Schwimmbecken verspüren: das Gefühl, von einer völlig anderen Realität umgeben zu sein, die sich ähnlich unterscheidet wie das Wasser von der Luft und die unsere volle Aufmerksamkeit, unseren gesamten Wahrnehmungsapparat in Anspruch nimmt<sup>2</sup>. Wie McMahan ist auch Murray der Auffassung, dass Immersion nicht einfach das Ergebnis einer intensiven sensorischen Stimulation ist. Eine mitreißende Erzählung könne laut Murray auch mit einem begrenzten sensorischen Input immersive Erfahrungen vermitteln, wie im Fall von Büchern (2017).

Der Norweger Jan-Noël Thon (2008) ist ebenfalls der Ansicht, dass es für die Immersion wichtiger sei, *was* präsentiert wird; weniger wichtig sei, *wie* es präsentiert wird. Auch er gibt als Beleg für diese These Bücher an, bei denen der sensorische Input begrenzt ist, die Leser:innen aber in ihren Köpfen Szenarien entwerfen können und so in den Akt des Lesens eintauchen. Zusätzlich sei, so Thon, die „Aufmerksamkeitsverschiebung“ (eng. *shift of attention*) von zentraler Bedeutung für das Konzept der Immersion. Diese Verschiebung der Aufmerksamkeit führe in Verbindung mit dem Konstruieren einer mentalen Repräsentation im Gehirn zu einer immersiven Erfahrung.

Laut Agrawal et al. (2019) können anerkanntermaßen zwei Faktoren unabhängig oder zusammen mit weiteren Faktoren psychologische Immersion bewirken: das subjektive Gefühl, umgeben zu sein oder multisensorisch stimuliert zu werden, und die Absorption von einer Erzählung oder der Darstellung dieser Erzählung.

### 5.1.1 Multisensorische Stimulation

Eine der gängigsten Konzeptualisierungen von Immersion ist das Gefühl, umgeben zu sein oder eine multisensorische Stimulation zu erleben. Biocca und Delaney (1995) bezeichnen diese Form des Eintauchens in die Umgebung als Wahrnehmungsimmersion (*perceptual immersion*). McMahan (2003) zufolge lässt sich Wahrnehmungsimmersion erreichen, indem man die Außenwelt ausblendet und die Wahrnehmung der Nutzenden von Videospielen auf den dargebotenen Stimulus einschränkt. Die Wahrnehmungsimmersion ist dabei keine eigene Gattung der Immersion. Vielmehr unterstützt sie die psychologische Immersion, da

<sup>2</sup> „*Immersion* is a metaphorical term derived from the physical experience of being submerged in water. We seek the same feeling from a psychologically immersive experience that we do from a plunge in the ocean or swimming pool: the sensation of being surrounded by a completely other reality, as different as water is from air, that takes over all of our attention, our whole perceptual apparatus.“ (S. 99)

sie die Aufmerksamkeit von der Aktivität abzieht und dadurch die psychologische Immersion fördert (Agrawal et al., 2019).

### 5.1.2 Narrative Immersion

Ein zweiter für immersive Erfahrungen maßgeblicher Faktor ist die Erzählung. Adams und Rollings (2006) sehen die Immersion im Falle einer Geschichte als das Gefühl an, sich in einer Geschichte zu befinden, vollständig darin involviert zu sein und die Welt der Geschichte und deren Ereignisse als real zu akzeptieren. Arsenault (2005) ist der Ansicht, dass man seine Vorstellungskraft nicht einsetzen muss, um in die Geschichte einzutauchen; insofern sei von einer spezifischen fiktionalen Immersion zu sprechen. Ryan (2003) teilt die narrative Immersion auf in räumliche, zeitliche und emotionale Immersion. Räumliche Immersion wird erlebt, wenn eine Person ein starkes Gefühl für den Raum hat und den Prozess genießt, diesen Raum zu erkunden. Bei der zeitlichen Immersion konzentriert sich die Aufmerksamkeit auf die sich entfaltende Geschichte, getrieben von der individuellen Neugier, zu erfahren, was als Nächstes passiert. Die emotionale Immersion entsteht durch die emotionale Bindung der wahrnehmenden Person an die Charaktere oder die Geschichte. Sie ist am stärksten, wenn diese Person sich mit der dargestellten Situation identifizieren kann und emotional in die Geschichte oder die Figuren involviert ist.

Agrawal et al. schlagen folgende gut brauchbare Definition für Immersion vor: Immersion ist ein Phänomen, das von einer Person erlebt wird, die sich in einem Zustand tiefer geistiger Beteiligung befindet, in dem ihre kognitiven Prozesse (mit oder ohne sensorische Stimulation) eine Verschiebung ihrer Aufmerksamkeit mit der Folge bewirken, dass sie sich von der Wahrnehmung der physischen Welt abkoppeln kann (2019).<sup>3</sup>

## 5.2 Immersion als objektive Eigenschaft eines Systems/Technologie

Mel Slater hat einen völlig anderen Ansatz. Seiner These nach verstärkt die Erhöhung der Zahl von Kanälen und Lautsprechern die Immersion, unabhängig von Inhalt, Kontext und individuellen Vorlieben (2003). Slater lehnt daher auch die Vorstellung ab, Immersion sei eine subjektive Erfahrung. Ein offenkundiges Manko dieses Konzepts besteht aber genau

<sup>3</sup> „Immersion is a phenomenon experienced by an individual when they are in a state of deep mental involvement in which their cognitive processes (with or without sensory stimulation) cause a shift in their attentional state such that one may experience disassociation from the awareness of the physical world.“ (S. 5)

darin, dass es die Grenzen der Wahrnehmung ebenso wenig berücksichtigt wie den Inhalt oder den Kontext des Wahrgenommenen und individuelle Faktoren wie Vorlieben und augenblickliche Stimmung. Die Rolle des Individuums ist aber insofern von größter Bedeutung, als Erfahrungen von Natur aus sehr subjektiv sind.

Die Überzeugung, Immersion sei eine objektive Eigenschaft eines Systems oder einer Technologie, wird in der Literatur indes nur von einer kleinen Minderheit vertreten. Slaters Konzept der Immersion ist auf die Verfügbarkeit eines externen Systems angewiesen, da es sich stark auf Anwendungen zur Erzeugung und Vermittlung virtueller Realität konzentriert, und das schränkt die Generalisierbarkeit seines Konzepts ein. Der Begriff Systemimmersion kann folglich für die von Slater und anderen vertretene Ansicht verwendet werden, die sich von dem oben beschriebenen Konzept der Immersion als psychologischer Zustand unterscheidet (Agrawal et al., 2019).

## 6 Sounddesign

### 6.1 Auditive Szene und Klangobjekte

Wie kann das Sounddesign die Wirkung der Immersion und das Hörerlebnis mit einem dreidimensionalen Mehrkanal-Wiedergabesystem steigern, das den Klang aus allen Richtungen zu erzeugen vermag? Um diese Frage zu beantworten, gilt es zunächst zu klären, wie das Sounddesign in der Lage ist, unterschiedliche Arten von Räumen zu schaffen. Denn die Art der Realität, die für das Publikum konstruiert werden soll, bestimmt die Auswahl der enthaltenen Klänge und deren Bewegung im Raum (Field, 2002). Nach Field (2002) lassen sich vier darzustellende Realitäten bzw. Räume unterscheiden:

- die reale Welt
- das Hyperreale
- imaginierte, aber realistische Räume
- das Surreale

#### 6.1.1 Die Darstellung der realen Welt

„Die auditive Szene besteht aus Klangobjekten in räumlicher und zeitlicher Beziehung.“ (Görne, 2017, S. 84). Ein Klangobjekt kann dabei entweder dinghaft – ein konkreter Klang, den wir in unserer Wahrnehmung mit einem Objekt in Verbindung bringen – oder abstrakt und somit „losgelöst vom klangerzeugenden Ding oder Ereignis“ (Görne, 2017, S. 84) sein.

In realen Räumen enthalten Klänge wichtige kontextuelle Hinweise auf die Umgebung. Ein mehrkanaliges System ermöglicht es uns, die Größe, den Umfang und die Funktion von Klangmaterial sowohl durch die Klangfarbe als auch durch den Raum zu bestimmen. Diese beiden Elemente, die sich in der Realität nicht voneinander trennen lassen, können zu kreativen Zwecken frei zusammengesetzt und wieder demontiert werden (Field, 2002).

Nach der Wahrnehmungstheorie von James J. Gibson sind für die Wahrnehmung der Realität unserer äußeren Umgebung die Struktur der Sinnesreize und darin vorkommende Invarianzen maßgeblich. Die menschliche Wahrnehmung beruht laut Gibson nicht auf der inneren Repräsentation des Wahrgenommenen, das die Einzelnen mit gespeichertem Wissen vergleichen und interpretieren, sondern auf einer direkten Resonanz der für uns wichtigen Invarianzen der äußeren Umgebung. Die Umwelt bietet dem Menschen dabei über die Reize vermittelte spezifische Handlungsangebote (eng. *Affordances*) (Mausfeld, 2005).

Wendet man diese Idee auf die Produktion einer immersiven auditiven Umgebung in der „realen Welt“ an, verweist sie auf Möglichkeiten, anspruchsvolle und detaillierte auditive Realitäten für die Hörer:innen zu schaffen, indem man räumliche Beziehungen zwischen Klängen herstellt, die denen ähneln, die wir täglich in der Welt erleben (Field, 2002).

### 6.1.2 Das Hyperreale

Hyperreale Räume weisen die Merkmale der Realität auf, allerdings in einer gesteigerten oder größer dimensionierten Form. Die eindrucksvollsten hyperrealen Räume sind oft diejenigen, die aus einer persönlichen Perspektive erlebt werden. Da sich in dieser Perspektive die Beziehungen zwischen den Hörer:innen und dem Raum ändern, müssen sich auch die kontextuellen Hinweise in der Umgebung ändern.

Aus gestalterischer Sicht kann die Glaubwürdigkeit hyperrealer Umgebungen nicht auf dem expliziten Inhalt beruhen, sie beruht vielmehr auf dem, was in einer Klanglandschaft nicht vorhanden ist. Indem die räumlichen Beziehungen der Realität beibehalten, aber wichtige Kontextinformationen vorenthalten werden, ist es möglich, die Hörer:innen zu erhöhter Aufmerksamkeit zu stimulieren (Field, 2002; Görne, 2002). In diesem Zusammenhang hat die Verwendung eines Surround-Systems den Effekt, dass die Kontraste zwischen den räumlichen Details der Realität und dem oft leeren Raum des Hyperrealen entsprechend vergrößert werden. Um diese Kontraste glaubhaft zu machen, ist es von zentraler Bedeutung, dass jeder/jede Hörer:in den gleichen Zugang zu den vom Surround-System präsentierten räumlichen Details erhält (Field, 2002).

### 6.1.3 Imaginäre und surreale Räume

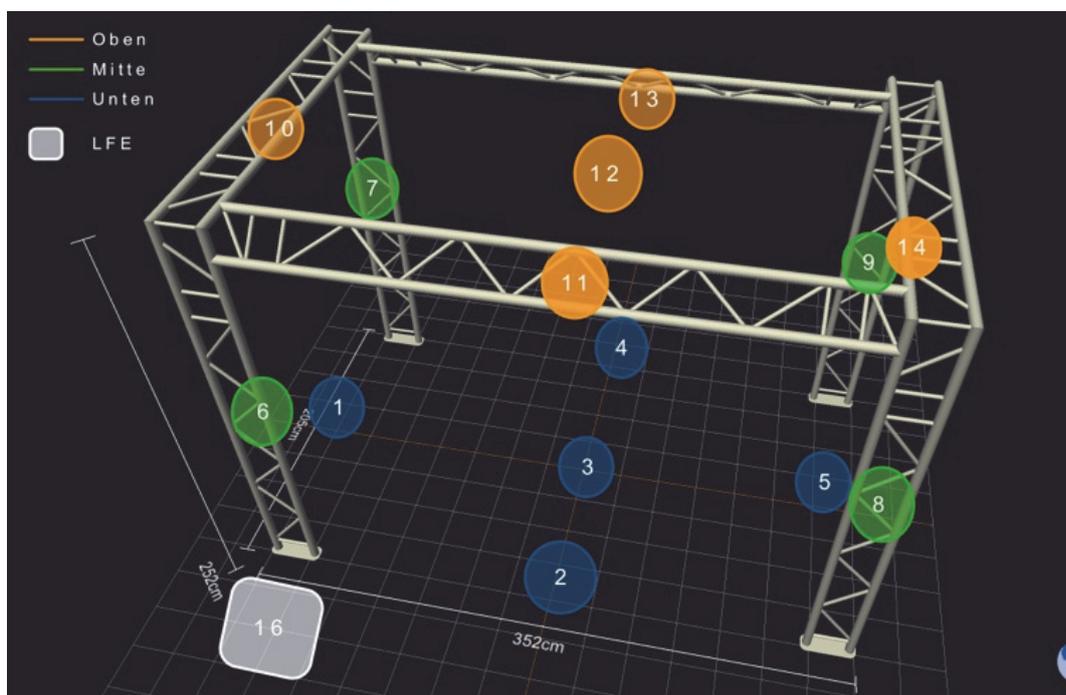
Imaginäre Räume beziehen ihre Glaubwürdigkeit direkt aus der Realität. Mit zunehmender Entfernung von der Realität gelangt man an einen Punkt, an dem der Raum auf die Hörer:innen eher surreal wirkt und nicht imaginär. Dieser Zustand ist erreicht, wenn sich die Beziehungen zwischen Objekten und ihren Räumen in einem Klangfeld aufzulösen beginnen. Field (2002) gibt dazu folgendes Beispiel: Die Klänge einer großen Umgebung werden in den Raum einer kleinen Telefonzelle geleitet. Die zugrundeliegenden Geräusche entstammen dann zwar noch denselben Quellen, aber ihre räumlichen Beziehungen entsprechen nicht mehr unseren Erwartungen an den ursprünglichen Kontext, in dem diese Klänge auftraten. Funktion, Kontext, Raum und unsere Erwartungen sind also eng miteinander verknüpfte Parameter. Durch die Bereitstellung einer umhüllenden Richtungsflexibilität bieten sich für das Sounddesign mehr Ansätze, diese Grenzen aufzubrechen und neu zu ziehen.

## 6.2 Diegese

Für die Differenzierung der Arten möglicher Räume bietet sich der aus der Erzähltheorie stammende Begriff Diegese an. Die „Diegese wird oft verkürzt als Bezeichnung der raumzeitlichen Beziehung der erzählten Welt, als Bezeichnung ihrer modellhaften Einheit, als räumlich-zeitliches Universum der Figuren bzw. Charaktere im Sinne des weicheren Begriffs der ‚erzählten Welt‘ gebraucht“ (Wulff, 2022). Bei der Kontextualisierung des erzählerischen Diskurses lassen sich drei Dimensionen unterscheiden: die extradiegetische, die alle Ereignisse bezeichnet, welche außerhalb der erzählten Welt stattfinden, die intradiegetische/diegetische, die alle Ereignisse meint, die in der erzählten Welt stattfinden, und die metadiegetische Ebene, die unter anderem Aussagen bezeichnet, die diegetische Elemente im Rahmen der Erzählung lokalisieren (ebd).

## 7 Studio Setup und Software

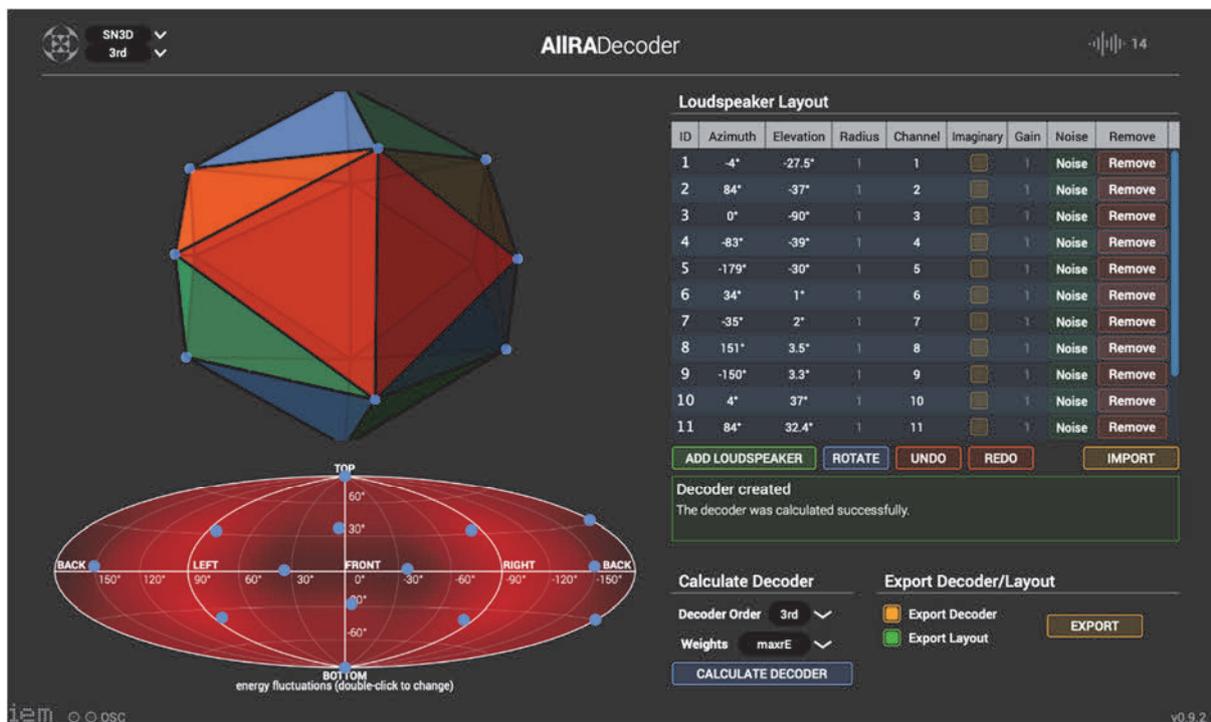
Das Studio, in dem ich das Sounddesign und die Mischung der drei Fragezeichen gemacht habe, ist die Regie 3 im Tonlabor der HAW DMI. Das dortige Lautsprechersetup besteht aus 14 Genelec 8331A, die in drei Ebenen an einem 3,52 m \* 2,05 m \* 2,52 m Käfig aus Traversen befestigt sind, plus einem Genelec 7370A Subwoofer in der linken unteren Ecke. Eine genauere Beschreibung des Aufbaus und der Verteilung der Lautsprecher ist in Troschka et al. (2021) nachzulesen. Über zwei Yamaha R18-D's sind die Lautsprecher mit Dante verbunden, das als Interface dient, und von Kanal 1 bis 16 durchgepatcht.



**Abbildung 7.1:** Lautsprecher-Konfiguration und Aufbau in Regie 3 des Tonlabors der HAW DMI

Als DAW habe ich REAPER des Unternehmens Cockos genutzt. Denn dies ist, nach meiner Kenntnis, die einzige DAW, die es ermöglicht, einen Kanal mit bis zu 64 Spuren zu erzeugen, was bei der Arbeit mit HOA unumgänglich ist. Da mir physisch 16 Kanäle Ausgangsseitig zur Verfügung standen, konnte ich mit dritter Ordnung Ambisonics arbeiten ( $\text{Anzahl der Lautsprecher} = (M + 1)^2$ ;  $M = \text{Ordnung}$  (SSA Plugins, 2017)). Allerdings wurden nur 14 dieser 16 Kanäle genutzt, so dass wir uns physisch nicht wirklich in dritter Ordnung befinden. Ungeachtet dessen habe ich sowohl en- als auch decodierseitig in dritter Ordnung gerechnet. Für den LFE habe ich einen eigenen Bus angelegt, der über Kanal 16 in Dante auf den Genelec 7370A geroutet ist. Bei der En- und Decodierung habe ich ebenso wie für die Bearbeitung der Spuren die IEM Plugin-Suite verwendet. Um spezifische Aus-

sagen zur Produktion und Wahrnehmung in einer vollsphärischen Lautsprecherumgebung treffen und weiterführende Fragestellungen erarbeiten zu können, habe ich dieselben beiden Szenen in 7.1.4, einem der komplexeren Konsumerformate, vertont und gemischt. Für das 7.1.4-Setup habe ich außerdem den ReaSurroundPan genutzt, den mehrkanaligen Panning-Tool der Firma Cockos. Um mögliche Lautstärkedifferenzen der Panning-Plugins auszuschließen, habe ich weißes Rauschen mit dem MultiEncoder und dem ReaSurroundPan auf vier verschiedene Positionen im Raum gelegt und die Pegeldifferenz mit dem NTI XL2 gemessen. Diese lag bei teilweise 7 dB, was ich im 7.1.4-Setup in dem MultiEncoder kompensiert habe.

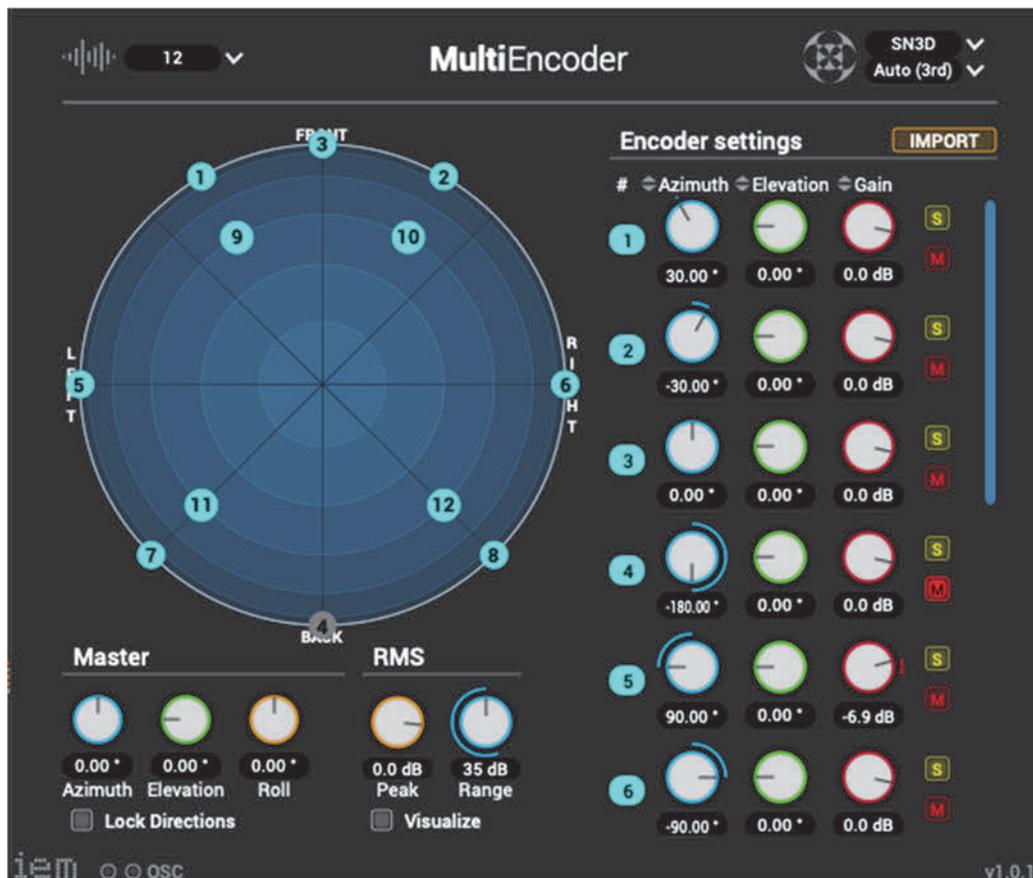


**Abbildung 7.2:** Lautsprecheranordnung aus Regie 3 im AIIRADecoder

Vor Beginn der Arbeit habe ich die Lautsprecher mit dem Smart Active Monitor (SAM™) System der Firma Genelec eingemessen und kalibriert. Um vom Tisch verursachte Reflexionen zu vermeiden, vor allem durch die unteren Lautsprecher, habe ich mir eine Art schalltransparenten Tisch aus einem 1 m \* 1 m-Eisengitter und vier Lautsprecherstativen gebaut.

Für die Bearbeitung der Audiospuren, vor allem für die En- und Decodierung, diente mir zum Großteil die IEM Plugin-Suite des Institute of Electronic Music and Acoustics, die kostenlos im Netz zur Verfügung steht. Sie ermöglicht es, Audiotracks mit bis zu 64 Kanälen zu bearbeiten. Mit deren MultiEncoder habe ich mir virtuelle Lautsprecher für das 7.1.4-Setup nach ITU-Norm (ITU, 2018) angelegt. Ich habe also technisch gesehen kein kanalbasiertes 7.1.4-Setup in der Produktion verwendet, sondern bin den Umweg über virtuelle Lautspre-

cher mittels Ambisonics gegangen. Anlass dafür waren die Interviews in Kapitel 11, da ich so deutlich flexibler war und über Knopfdruck den Rezipient:innen beide Mischungen hintereinander vorführen konnte, ohne das Studio umbauen und neu einmessen zu müssen.



**Abbildung 7.3:** Anordnung der virtuellen 7.1.4-Lautsprecheranordnung im MultiEncoder nach ITU-Norm

Das Material aus der Folge *Die drei ??? und das Dorf der Teufel* haben mir freundlicherweise Pierre Brand (Primetimestudio) und Maike Müller (Sony Deutschland) zur Verfügung gestellt. Der Fall der Ermittler, der aus der „3D-Audio“-Reihe stammt, wurde speziell für ausgewählte Planetarien von der gleichnamigen Graphic Novel adaptiert (*Die drei ???, o.D. 3D Audio*). Um vom Original nicht beeinflusst zu werden, habe ich mir das Stück nicht angehört. Als Ausgangsmaterial erhielt ich die Sprachspuren von Morton, Justus, Peter und Bob als AAF, die Musik als Stereofile, außerdem einige SFX-Spuren von Steinen, Geröll und Klamotten. Für das Sounddesign habe ich mich aus der Datenbank *Soundly* bedient, aber auch eigene Klänge aufgenommen. Die Sprachspuren habe ich kaum bearbeitet, um den Eindruck der Natürlichkeit zu wahren. Lediglich die Stimme des Erzählers habe ich mit einem Deesser und einem Kompressor bearbeitet, um die unnatürliche nichtdiegetische Gestalt markanter hervorzuheben.

## 8 Gestalterischer Prozess

### 8.1 Perspektive

Bevor ich mit dem eigentlichen Sounddesign beginnen konnte, musste ich mir erst einmal darüber klarwerden, welche Perspektive man als Hörer:in einnimmt. Die Perspektive wird akustisch durch Anordnung der Lautsprecher und durch die Klangobjekte vermittelt. Das True-Spatial-System bietet ein schier endloses Spektrum möglicher Perspektiven: von einer gottgleichen Gestalt, die den Schnitten des Films ähnelt und nicht an die physikalischen Regeln der fiktiven Welt gebunden ist, bis hin zur Besitzergreifung einer der Figuren des Hörspiels, die diegetisch partizipiert. Eine meiner ersten Ideen war es, als vierte der drei Figuren zu fungieren, in der Annahme, auf diese Weise die Immersion als psychologischen Zustand zu verstärken und Hörer:innen Teil der Geschichte werden zu lassen. Als Zuhörender soll man den drei Fragezeichen bei ihren Ermittlungen folgen und sich als Teil des Detektiv-Teams fühlen. Die Welt der Geschichte soll sich dabei an der unsrigen orientieren und denselben physikalischen Gesetzen gehorchen, ganz im Sinne der Gedanken von Field (2002), die in Kapitel 6.1.1 vorgestellt wurden. Um die Klangobjekte passend zu dieser virtuellen Realität zu gestalten, habe ich sie dinghaft angelegt (vgl. Kapitel 6.1.1). Bei der Konzeptualisierung dieser Idee habe ich Skizzen für die jeweiligen Szenen aus der Vogelperspektive angefertigt.

### 8.2 Erzähler

Bei der Platzierung des Erzählers stellen sich ähnliche Fragen: Wo platziere ich die nicht-diegetische, allwissende Figur des Erzählers, der über allem steht? Mein erster Ansatz war es, die Stimme der metaphorisch gottesgleichen Figur des Erzählers auf den gleichnamigen Lautsprecher zu schicken, dem Voice of God (VoG) direkt über den Hörer:innen. Doch dies klang zu indirekt, zu anonym und hatte nur in Ansätzen den nichtdiegetischen Effekt, den ich mir vorgestellt hatte. Daraufhin habe ich mich an Filmen orientiert, bei denen die Dialoge stets aus dem Center-Kanal kommen. Aber auch in diesem Fall sorgte die Platzierung eher für Verwirrung: In der Szene, in der der Erzähler spricht, befinden wir (die Hörer:innen) uns auf dem Rücksitz des von Chauffeur Morton gesteuerten Wagens. Dessen Stimme und die von Bob liegen virtuell bereits vorne, auf dem Fahrer- und dem Beifahrersitz. Kommt die Stimme des Erzählers nun direkt von vorne, auch bei Variation auf der Vertikalen, zerstört dies das Bild der Szenerie, das im eigenen Kopf entsteht, so dass die Wirkung der Immersion verlorengeht.

Am Ende habe ich mich für die Platzierung unter dem/der Hörer/in entschieden. Wegen der unzureichenden Fähigkeit des Menschen, auf der unteren Hemisphäre zu lokalisieren, und

wegen der Reflexionen des eigenen Rumpfes und jenen des Stuhls ist eine präzise Ermittlung des Erzähler-Ortes nicht möglich (vgl. Kapitel 11.2.6). Wie Wasser in einem See, in das man untergetaucht ist, umgibt einen der Erzähler subjektiv von allen Seiten. Doch der Ausgangsort seiner Stimme lässt sich nicht bestimmen. Das bekräftigt die nichtdiegetische Gestalt des Erzählers und lässt seine Stimme sehr intim wirken. Allerdings ist diese Wirkung nur im True-Spatial-Setup zu erzielen, da einem bei 7.1.4 einzig die obere Hemisphäre zur Gestaltung bleibt. Hier bestand die Möglichkeit, die Erzählerstimme direkt im ReaSurroundPan in den Raum zu holen. Dies lässt sich aber nur mit psychoakustischen Tricks im True-Spatial-Setup realisieren. Um denselben Effekt zu erzeugen wie in True Spatial, habe ich die Stimme des Erzählers im Kopf der Hörenden erklingen lassen. Auch auf diese Weise löst sich der Erzähler aus der Szenerie, seine Worte muten an wie die eigenen Gedanken, nicht aber wie die einer diegetischen Figur.

### 8.3 Musik

Wie der Erzähler ist auch die Musik in diesem Fall ein nichtdiegetisches Klangobjekt in der Geschichte. Insofern muss man sich auch hier wieder die Frage stellen, wie sich die Immersion aufrechterhalten lässt, trotz solcher nichtdiegetischen Klänge. Dramaturgisch spielt die Musik eine entscheidende Rolle: Gegen Ende der elften Szene ist eine Bassmelodie zu hören, als Peter das Auto des vermissten Mr. Jacobs in einer Schlucht entdeckt, was die Situation sehr spannend und fesselnd macht. Das gilt auch in der darauffolgenden Szene, in der Peter an ein Seil gebunden nach Hinweisen im Autowrack sucht, das bald darauf weiter den Hang hinabrutscht. Die Musik vermittelt keine eindeutige Antwort auf die Frage, ob Peter es rechtzeitig aus dem Wrack geschafft hat oder mit ihm am Boden zerschellt.

Um den Charakter der Nicht-Diegeese zu unterstreichen und das Gefühl der Umhüllung zu erzeugen, habe ich die Musik, die mir als Stereofile zur Verfügung gestellt wurde, mittels Halo-Upmix (NUGEN Audio) auf 7.1.4 hochgemischt. In der True-Spatial-Mischung bin ich einen Schritt weiter gegangen und habe die Spur gedoppelt, in der duplizierten Spur außerdem lediglich sechs der sieben Kanäle auf der Horizontalebene genutzt und nach unten gelegt ( $\pm 30^\circ$ ,  $\pm 135^\circ$  &  $\pm 90^\circ$  Azimut; jeweils  $-45^\circ$  Elevation).

Einen größeren Effekt auf das Gefühl der Umhüllung hatte dies allerdings nicht. Diesen Halo-Trick, bei dem ich mir nur die vier Höhenkanäle des Halo-Upmix genommen und per MultiEncoder auf den unteren Bereich der Hemisphäre gelegt habe, verwendete ich nicht nur bei der Musik, sondern auch bei einigen Atmos (Atmosphären). Obwohl das Upmixing-Tool von NUGEN einen Algorithmus besitzt, der wahrscheinlich einige Frequenzen auf eine gewisse Art für die Höhenkanäle verzerrt, funktioniert das recht gut.



**Abbildung 8.1:** Der Halo-Trick; Halo-Upmix (links) und MultiEncoder (rechts); Kanäle 5, 6, 9, 10, 11 & 12 im MultiEncoder haben  $-45^\circ$  Elevation bei jeweiliger Position; die restlichen Kanäle sind auf „mute“

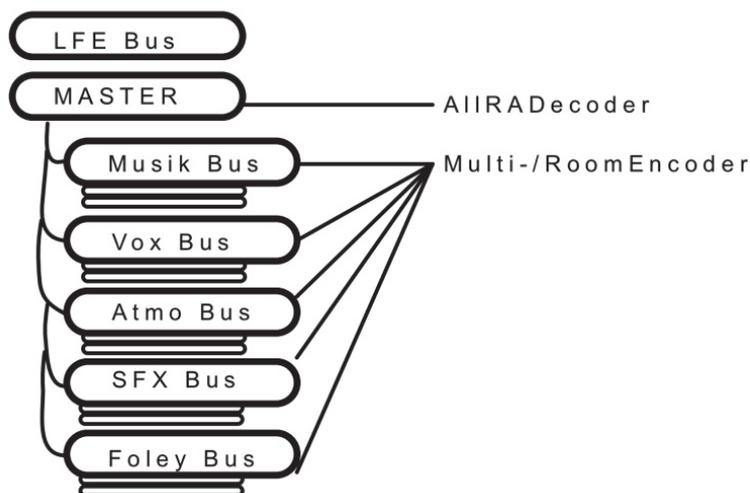
## 8.4 Einordnung der Szenen

In der 2017 erschienenen Folge *Die drei ??? und das Dorf der Teufel* machen sich die jungen Detektive Justus Jonas, Peter Shaw und Bob Andrews auf die Suche nach einem vermissten Freund ihres Chauffeurs Morton in Redwood Falls. Dieser Fall der drei ??? ist Teil der dritten Staffel von Hörspielen, die speziell für ausgewählte Planetarien geschrieben wurden (Die drei ??? Wiki, o. D.). Bei den von mir ausgewählten Szenen handelt es sich um Szene 11 und 12 des Hörspiels. Zuvor sind die drei Detektive von ihrem Chauffeur nach einer Reise in Peru vom Flughafen in Rocky Beach abgeholt worden. In dessen Auto findet Justus auf der Heimfahrt ein Amulett, das Mrs. Jacobson gehört, einer Frau, die Morton zuvor chauffiert hat. Als die vier zu dieser Frau fahren, um ihr das Amulett zu geben, kommt bei einem Gespräch mit ihr heraus, dass ihr Ehemann Daniel Jacobson, ein berühmter TV-Koch, spurlos verschwunden ist. Die drei Ermittler nehmen sich des Falls an und gelangen nach Redwood Falls, wohin die letzte Spur des verschollenen Ehemanns führt. In dem kleinen Dorf im Wald treffen sie den Bürgermeister und Vater Daniels, der berichtet, dass sein Sohn dort nie angekommen sei. Die drei Fragezeichen und Morton beschließen, in Redwood Falls zu bleiben und vor Ort zu ermitteln. Sie übernachten in einer Pension, wo sie nachts von Geräuschen der Dorfbewohner geweckt werden, die als Teufel verkleidet einer aus Stroh gefertigten Teufelsfigur huldigen und den drei Fragezeichen und Morton Böses wünschen. Am nächsten Tag beschließen die vier, das Dorf unter dem Eindruck der Ereignisse der vergangenen Nacht zu verlassen. Sie setzen sich in Mortons Wagen und machen sich auf den Weg. Ab hier beginnt Szene 11 und somit das Sounddesign.

## 9 True Spatial

Da ich der erste Student meiner Hochschule bin, der in einem von mir als True Spatial definierten vollsphärischen System ein Hörspiel produziert hat und es dazu kaum Literatur oder einschlägige Arbeiten gibt, musste ich vieles ausprobieren, und dies manchmal mehrfach. Mein Ziel war stets, den Effekt der Immersion hervorzurufen, wie ihn Agrawal et al. (2019) definiert haben (siehe oben, Kapitel 5.1.2).

Der Signalfluss bei der Bearbeitung in True Spatial sah wie folgt aus: Ich habe fünf Bus-Spuren angelegt, jeweils für Musik, Sprache (Vox), Atmosphären (Atmo), Soundeffekte (SFX) und Foley. Auf den einzelnen Spuren der jeweiligen Busse liegt der Multi- beziehungsweise der RoomEncoder. Die 16 Spuren, die ich nach der Encodierung bekomme, gehen an die jeweiligen Busse und von dort auf den Master, auf dem über den AllRADecoder die Informationen decodiert und auf die Lautsprecher beziehungsweise an Dante geleitet werden. Wenn ich bestimmte Spuren auf den LFE geben wollte, war dies über die Send-Funktion einfach zu realisieren. Der LFE Bus ging direkt über Kanal 16 in Dante auf den Subwoofer.



**Abbildung 9.1:** Spurenverteilung in der True-Spatial-Mischung mit jeweiligem En-/Decoder

### 9.1 Szene 11

#### 9.1.1 Autofahrt

Zu Beginn der elften Szene befinden wir uns im Wagen des Chauffeurs Morton und fahren langsam aus Redwood Falls los, verabschiedet von den Einheimischen.

### 9.1.1.1 *Das Auto*

Um die Atmosphäre und das Gefühl zu vermitteln, in einem fahrenden Wagen zu sitzen, habe ich auf ein hyperreales Sounddesign mit dinghaften Klängen gesetzt. Nach dem Skript fährt Morton einen Rolls-Royce, welches Modell genau, wird nicht gesagt. In jedem Fall ist davon auszugehen, dass es im Inneren eines Wagens dieser Luxus-Marke relativ still ist. Nach Berichten ist es in einem Rolls-Royce Ghost so still, dass einem physisch geradezu übel wird (Elliott, 2020). Um dennoch die Illusion des Autofahrens zu erzeugen, habe ich mich metaphorischer Auto-Klänge bedient und diese hyperreal dargestellt. Streng genommen ist also kein Rolls-Royce zu hören, sondern ein älterer Viertürer.

Die Geräusche im Auto während der Fahrt bestehen bei mir aus sechs Klängen: hochfrequente Fahrtgeräusche aus dem Innern eines Zastava Yugo Baujahr 1980, jeweils zwei Aufnahmen aus einem langsam fahrenden SUV und drei Aufnahmen aus Fahrzeugen, die über einen Schotterweg fahren. Mit den Spuren bin ich sehr unterschiedlich umgegangen.

Das hochfrequente Audiofile aus dem Fahrzeuginnern habe ich besonders wegen des Quietschens gewählt. In der Aufnahme ist zu hören, wie das Automobil bei jedem Überfahren einer Bodenwelle quietscht, was bei der Verbildlichung des Autos sehr hilft. Dieses Stereofile habe ich mit dem Halo-Upmix auf der oberen Hemisphäre im 7.1.4-Stil verteilt.

Die beiden tiefer klingenden Aufnahmen aus den SUVs habe ich gewählt, da hier sehr schön ein Motor zu hören ist und die Gänge immer wieder hoch- und runtergeschaltet werden, was die Fahrt lebhafter gestaltet und nicht so monoton wirken lässt. Beide Spuren sind mit dem StereoEncoder bearbeitet worden, indem ich sie einmal nach breit vorne unten (Azimuth  $0^\circ$ , Elevation  $-25^\circ$ , Width  $132^\circ$ ) und einmal nach hinten unten (Azimuth  $0^\circ$ , Elevation  $-170^\circ$ , Width  $125^\circ$ ) gelegt habe. Zusätzlich habe ich beide Spuren auf den LFE mit  $-6$  dB geschickt, um den Motor auch spürbar zu machen und die tiefen Frequenzen zu verstärken. Bei der Anordnung habe ich mich wieder an einem realen Auto orientiert, bei dem der Motor vorne liegt, die Bewegung des Fahrzeugs aber die Karosserie zum Vibrieren und Brummen bringt.

Auf Seite 23 des Skripts gibt Morton Auskunft über den Zustand der örtlichen Straßen, die „nicht unbedingt im besten Zustand [sind]“ (Beckmann & Menger, 2019; im Folgenden zitiert als „Skript“, hier S. 23). Entsprechend habe ich das Sounddesign gestaltet: Durch die geschlossenen Fenster des Wagens sind Geräusche zu hören, die bei der Fahrt über einen sandigen Schotterweg entstehen. Auch hier war mir die untere Hemisphäre teilweise von Nutzen. Neben Motorgeräuschen habe ich das Rollen des Wagens auf jenem Schotterweg mit dem Halo-Trick auf die untere Hemisphäre gelegt. Da sich der Klang der rollenden Reifen auf dem Weg aber nicht nur im unteren Teil des Wagens breitmacht, habe ich zwei Spuren des Rollens noch quadrafonisch und in die Höhe ( $37^\circ$  Elevation) im Raum verteilt.

Dies sollte das akustische Grundgerüst für meinen Wagen sein. Dabei musste ich allerdings die hinten liegenden Klänge deutlich lauter machen (6–8 dB) als die vorderen, damit subjektiv ein Gefühl der Umhüllung auftrat.

Um die Fahrt lebhafter zu gestalten, habe ich an zwei Stellen Autos vorbeifahren lassen. Da wir uns in den USA befinden, wo Rechtsverkehr herrscht, sind die Autos jeweils links vorbeifahren. Ähnlich den Geräuschen einer Menschenmenge habe ich sie in der Ferne schmal gehalten und während des Vorbeifahrens breiter gezogen, um den Effekt realitätsnah zu gestalten. Damit dies auch bei geschlossenen Fenstern zu hören ist, wurde auch hier wieder bei etwa 10 kHz mit 12 dB/oct gefiltert. Was die Höhe angeht, bin ich hier auf 0° Elevation geblieben. Die vorbeifahrenden Autos kamen im weiteren Verlauf der Szene erneut zum Einsatz, nachdem der Rolls-Royce am Straßenrand angehalten hat und die Insassen ausgestiegen sind. Dabei machen die vorbeifahrenden Autos die Szene nicht nur lebhafter, sie dienen auch der Orientierung. Da in meiner Version des Hörspiels die visuelle Wahrnehmung fehlt, habe ich einige Klangobjekte für die räumliche Orientierung genutzt, darunter eben auch die Autos. Während sie bei der Fahrt links an uns vorbeifahren, sind sie nach dem Ausstieg hinter uns: Folglich haben wir uns, beziehungsweise die Szenerie, um 90° nach rechts gedreht.

Während der Fahrt entdecken die Detektive auf einem Handyfoto, dass sich jemand „am Rolls-Royce zu schaffen [macht]“ (Skript, S. 39). Die Gefahr, in der sich die drei Fragezeichen nun zu befinden scheinen, wird durch bedrohlich klingende Streicher musikalisch unterlegt. Diese sind wie oben beschrieben im Raum platziert. Im Skript legt Morton als Reaktion eine Vollbremsung hin. Hier habe ich mich gegen die Vorgaben des Skripts entschieden: Statt eine Vollbremsung auszuführen, fährt Morton rechts ran. Zu Beginn habe ich mich sehr genau am Skript orientiert und eine Vollbremsung akustisch nachempfunden. Nach meinem subjektiven Eindruck fehlte dabei aber ein Element, das jene Kräfte vermittelt, die bei einer Vollbremsung auf einen Körper einwirken. Die Vollbremsung wirkte insofern also rein akustisch völlig unglaubhaft. So entschied ich mich dafür, Morton statt einer Vollbremsung den Blinker setzen und an den Straßenrand fahren zu lassen, wo der Wagen zum Halten kommt.

Den Klang des Blinkers habe ich sehr schmal nach vorne links gelegt, dahin, wo der Fahrer sitzt. Sehr auffällig war dabei die subjektive Lokalisation des Blinkers. Um den Blinker vorne links zu verorten, musste ich ihn im Encoder sehr nah an die Medianebene (6° Azimut) heranführen und um etwa -15° Elevation herabbringen. Gleiches gilt für die Handbremse, die Morton nach dem Rechtsranfahren zieht. Da diese Bremse in einem Wagen normalerweise in der Mitte unten liegt, habe ich sie dort auch platziert. Allerdings passte die Platzierung bei 0° Azimut nicht mit der wahrgenommenen Position überein, weshalb ich hier etwa -6° Azimut variiert habe, damit die Handbremse in der Mitte wahrzunehmen ist. Für die subjektiv richtige Verortung musste ich hier ebenfalls auf die untere Hemisphäre zurückgreifen und das Ziehen der Handbremse auf -32° Elevation legen.

Die Kernklänge des Wageninnern ändern sich, nachdem Morton den Wagen an den Straßenrand gefahren hat. Was wir nun hören, sind insgesamt vier Stereospuren: den Bremsvorgang eines PKWs auf Schotter und drei Klänge eines Motors im Leerlauf. Hier musste ich auch wieder experimentieren, um herauszufinden, wohin ich welche Klänge legen soll, damit ein immersives Hörerlebnis entsteht.

Das am meisten befriedigende Hörerlebnis des Bremsvorgangs entstand, als ich die Spur schmal ( $5^\circ$  Width) in etwa  $90^\circ$  Elevation legte. Was bei dem Audiomaterial für das Sounddesign entscheidend war, ist das Abbremsen der Reifen auf Schotter.

Für die Erzeugung der Klänge des Standmotors habe ich drei verschiedene Motoren ausgewählt, die in der Summe einen sehr real klingenden Motor ergeben haben. Spezifisch für die einzelnen Aufnahmen war die jeweilige Klangfärbung: Der erste Motor hat nicht viel Körper und klingt sehr dünn, der zweite Motor hat einen sehr schönen Klang in den Mitten und der dritte Motor ist sehr bass-lastig. Den bass-lastigen und den mitte-betonenden Motor habe ich wieder mit Halo auf der oberen Hemisphäre im 7.1.4-Schema verteilt. Aus den drei Spuren wurde aber erst dann ein authentisch klingender und umhüllend wirkender Motor, als ich den klein klingenden dünnen Motor sehr breit ( $130^\circ$  Width) auf  $-22^\circ$  Elevation verteilt habe. Da die Aufnahmen aus dem Fahrzeuginnern stammen, war eine weitere Bearbeitung durch einen EQ nicht nötig.

Kurz vor dem Ausstiegsvorgang habe ich mich erneut vom Skript gelöst und Morton den Motor ausschalten lassen, bevor alle aus dem Wagen aussteigen. Ich hatte auch hier ausprobiert, wie es wirkt, wenn der Wagen weiterhin läuft, während die drei Detektive, Morton und wir (die Hörer:innen) um den Wagen herumstehen. Allerdings wirkte der laufende Motor in Kombination mit Teilen der Außenatmosphäre bei der Erkennung von Klangobjekten irritierend. Darum habe ich Morton den Zündschlüssel ziehen lassen, bevor er den Wagen verlässt. Da sich das Zündschloss typischerweise unterhalb des Kopfes befindet, habe ich auch hier den unteren Teil der Hemisphäre genutzt und den Klang bei etwa  $-13^\circ$  Elevation platziert.

Für den Wechsel dieser beiden Kernklänge des Wagens habe ich zwei Bus-Spuren angelegt, jeweils mit dem fahrenden und dem stehenden Auto; diese Spuren habe ich anschließend in einer gemischt, so dass die Illusion eines an den Straßenrand fahrenden Autos entsteht.

#### 9.1.1.2 *Die Charaktere*

Um die Akustik des Wagens zu simulieren, habe ich den RoomEncoder (IEM Plugin-Suite) verwendet. Zur Erläuterung: Bei der Wiedergabe über Lautsprecher hören wir für gewöhnlich zwei Räume: den Aufnahmeraum und denjenigen, in dem das Audiomaterial wiedergegeben wird. Da ich den Hörer:innen den Eindruck einer akustisch virtuellen Welt vermitteln wollte,

um die Immersion zu steigern beziehungsweise zu erleichtern, reichten diese beiden Räume aber nicht aus, da sie nicht die akustischen Attribute eines realen Wagens haben.

Der RoomEncoder ermöglicht es, das Klangobjekt in diesen virtuellen Raum zu legen, der die akustischen Eigenschaften besitzt, die ich dem Raum gebe, und eine räumliche Distanz zwischen Klangobjekt und Hörendem herzustellen. Der Algorithmus des Plugins berechnet dabei bis zu 236 Schallreflexionen. Problematisch ist die Dokumentation des Plugins. Weder in der Bedienoberfläche noch auf der Website finden sich Informationen darüber, welche Eigenschaften diese virtuellen Wände haben, im Speziellen der Absorptionsgrad. Durch den Equalizer in der Mitte des Plugins lassen sich zwar die Reflexionen färben und die Nachhallzeit ändern, einen Einfluss auf den Grad der Absorption gewinnt man auf diese Weise aber nicht. Außerdem lässt sich der virtuelle Raum nur rechteckig anlegen. Ein ovaler oder komplexerer Raum mit mehr Ecken und Winkeln kann nicht generiert werden. Der MultiEncoder arbeitet grundverschieden: Mit ihm lassen sich Klangobjekte ausschließlich auf den Rand der Sphäre legen, die der Decoder berechnet.

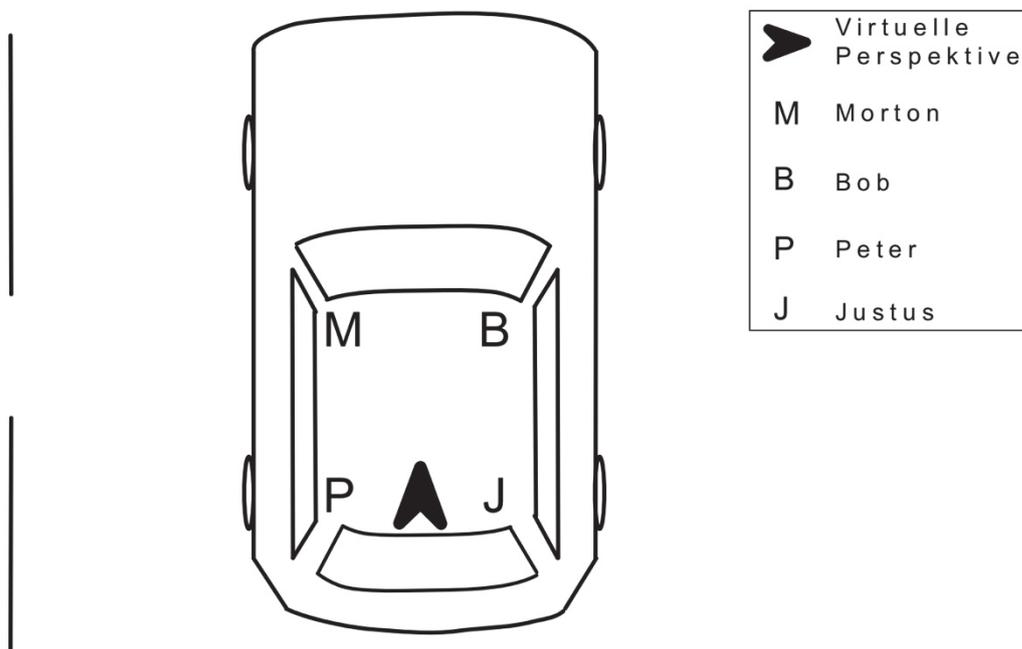
Um einen Klang dennoch ins Innere der Sphäre zu bekommen, muss man sich einiger Tricks aus dem Arsenal der Psychoakustik bedienen, wie zum Beispiel des Nahbesprechungseffekts. Das Azimut besitzt auf der rechten Seite der Hemisphäre negative Winkel, auf der linken Seite positive. Bei der Elevation verhält es sich so, wie man es erwartet: Der obere Teil der Hemisphäre besitzt positive Winkel, während der untere Teil durch negative Winkel beschrieben wird.



**Abbildung 9.2:** IEM RoomEncoder (links) und der IEM MultiEncoder (rechts)

Da ich stets versucht habe, die Verhältnisse und Gegebenheiten der realen Welt zu simulieren, um die Immersionswirkung zu steigern, habe ich mich meist auch an deren Abmessungen gehalten, um in der Vorstellung der Rezipient:innen nicht surreale oder imaginäre Welten zu schaffen. Mit diesem Tool habe ich mir im ersten Teil der Szene das Auto virtuell mit den Maßen  $1.5\text{ m} \times 1.5\text{ m} \times 1\text{ m}$  nachgebaut.

Wie erwähnt habe ich den perspektivischen Ansatz gewählt, dass wir (die Hörer:innen) als unsichtbare vierte Figur der drei Fragezeichen fungieren und Teil der diegetischen Welt sind. Dabei geben wird in der Geschichte allerdings keinerlei Geräusche von uns, sondern sind ausschließlich stille Beobachter:innen. Wir sitzen auf der Rückbank zwischen Peter zu unserer Linken und Justus zu unserer Rechten. Auf dem Fahrersitz vorne links sitzt unser Chauffeur Morton. Zu dessen Rechten auf dem Beifahrersitz befindet sich Bob. Die Personen habe ich jeweils auf derselben Höhe platziert, auf der sich unser Kopf befindet. Wichtig ist, dass sich die virtuelle Position des Kopfes der Charaktere nicht verändert. Sie sprechen stets nach vorn. Um diese statische Anordnung zu brechen und den Stimmen Körper zu verleihen, habe ich sie sich bewegen lassen. Wer in einem Auto sitzt, macht mit seinem Körper und seinen Kleidern durch Reibung etwa auf dem Sitz Geräusche, die ebenso als Folge der Bewegung des Fahrzeugs entstehen. Im Fall der drei Fragezeichen habe ich mich für knarzendes Leder entschieden, da so ein Stück weit die edle Ausstattung des Luxuswagens zur Geltung kommt. Das ständig knarrende Leder, vor allem zu Beginn der Szene, verweist auf die Unsicherheit und das Unbehagen der drei Ermittler, die sie erleben, als sie Redwood Falls frühzeitig verlassen. Das wollte ich noch deutlicher machen, indem ich das Knarzen des Leders etwa 4 dB lauter gemacht und die Insassen merklich häufiger in Bewegung versetzt habe als gegen Ende der Autofahrt.



**Abbildung 9.3:** Darstellung der virtuellen Umgebung in der ersten Hälfte von Szene 11

Für die Geräusche „winkender Menschen“ (Skript, S. 38), welche die drei Fragezeichen verabschieden, konnte ich mich keiner Datenbank für Klänge bedienen. Darum habe ich am

17. Mai 2022 sechs Kommiliton:innen im Aufnahmeraum 1 der HAW DMI mikrofoniert und aufgenommen. Dabei habe ich einige Verabschiedungsphrasen vorgeschlagen, die ich dann von jedem und jeder Einzelnen aufgenommen habe. Um das Scharmgefühl zu dämpfen und die Sprecher:innen authentisch klingen zu lassen, haben wir uns im Kreis aufgestellt und die Verabschiedungsformeln als Gruppe ausgesprochen.

Um die Illusion einer Menschenmenge zu erzeugen, habe ich die sechs Spuren zweimal dupliziert, gegeneinander verschoben und im virtuellen Raum verteilt. So sind keine Individuen mehr zu hören, der Eindruck einer Menschenmenge entsteht. Da die Menschen am Rand der Straße stehen, auf der wir entlangfahren, musste jede einzelne Spur hinsichtlich des Panning automatisiert werden. Dabei bin ich stets im positiven Elevationswinkel geblieben, da die stehenden Menschen höher verortet sind als wir, die im Wagen sitzen. Ansonsten reichte das für die Illusion. Im Skript heißt es zudem, dass die Einwohner „durch die Scheibe“ (Skript, S. 38) zu hören sind. Um diesen Effekt zu erzielen, habe ich die Aufnahmen mit dem MultiEq (IEM Plugin-Suite) bearbeitet, die Höhen mit einem 12 dB/oct Tiefpassfilter bei 300 Hz gefiltert und den tieffrequenten Anteil der Signale bei 100 Hz mit 12 dB/oct weggefiltert. Das Verlassen des Dorfes habe ich durch Lautstärke-Automation der von den Einwohner:innen verursachten Geräusche simuliert. Mit zunehmender Entfernung des Wagens habe ich die Verabschiedungsäußerungen stetig leiser gemacht, von einem bestimmten Punkt an kommen die Stimmen nur noch von hinten, wobei deren Winkel mit fortschreitender Entfernung schmaler wird.

### 9.1.1.3 *Das Wetter*

Abgesehen von den Klängen des Wagens ist im Skript von „drohendem Gewittergrollen“ (Skript, S. 38) die Rede. Vereinzelt Donnergrollen habe ich mit dem MultiEncoder auf die obere Hemisphäre gelegt und die Bewegung des Donners am Himmel automatisiert. Zusätzlich ging die Spur mit den Donnern auf den LFE, um das drohende Gewitter spürbar zu machen und dessen Bedrohlichkeit zu steigern. Da wir immer noch im Innern des Rolls-Royce sind, habe ich die Höhen des Donnerns etwas gefiltert. Die Aufnahmen aus der Datenbank waren ohnehin recht tieffrequent, so dass hier ein 12 dB Tiefpass bei 15 kHz ausreichte. Während der Autofahrt ist sonst keinerlei Naturgeräusch zu hören.

Neben den vorbeifahrenden Autos dient vor allem ein bestimmtes Klangobjekt der Orientierung im virtuellen Raum: der Wasserfall. Mein akustischer Wasserfall besteht aus insgesamt sechs Spuren. Drei davon sind Aufnahmen realer Wasserfälle, zwei sind Aufnahmen eines Flusses, die sechste Spur besteht aus weißem Rauschen. Ähnlich den diversen Motoren besitzt jede der drei Aufnahmen von Wasserfällen eine eigene Klangfarbe. Durch Kombination der Aufnahmen habe ich mir meinen eigenen Wasserfall geschaffen. Bei der Platzierung habe ich die Spuren nach Klangfarbe aufgeteilt: die hellere, dünner klingende Aufnahme

nach oben bei etwa 30° Elevation, die mittenlastige Aufnahme bei 0° Elevation und den bassigen Teil auf -35° Elevation.

Beim Hören fiel mir allerdings auf, dass diese vertikale Aufteilung kaum Auswirkung auf die Lokalisation hatte, was den Feststellungen Blauerts zu breitbandigen Klängen in Kapitel 4.3 entspricht. Um den Workflow zu erleichtern, habe ich mich daraufhin entschieden, die Stereospuren auf eine Bus-Spur zu routen und dort als Summe mit dem StereoEncoder zu verschieben. Die fehlende Trennung der Spuren ist nicht aufgefallen, allerdings musste ich die Summe auf die untere Hemisphäre schieben (-31° Elevation), um den Wasserfall realistisch wirken zu lassen. Zuvor war das Rauschen nur oben zu verorten.

Beim ersten Durchlesen des Skripts hatte ich das Bild eines großen Wasserfalls vor Augen, der Hunderte von Metern in die Tiefe stürzt. Entsprechend klingt der von mir designte akustische Wasserfall: sehr rauschig und vor allem tiefenlastig. Das Problem bei Klängen wie dem des Wasserfalls ist die spektrale Verteilung. Ähnlich dem weißen Rauschen lassen sich verschiedene Objekte und Bilder damit erzeugen. So kann weißes Rauschen hervorragend für die Gestaltung des Meers genutzt werden, ebenso gut gelingt es aber auch für Wind. Da wir uns in diesem Fall allein auf die Ohren verlassen müssen und die optische Wahrnehmung entfällt, musste ich mit metaphorischen Klängen arbeiten, um die Illusion des Wasserfalls zu schaffen. Deshalb habe ich zusätzlich Aufnahmen von zwei kleinen Flüssen genutzt, in denen plätscherndes Wasser zu hören ist. Durch dieses Plätschern wollte ich den Eindruck des Mediums Wasser transportieren. Das Plätschern liegt dabei auf der unteren Hemisphäre bei jeweils -30° Elevation und -35° Elevation.

Während der Autofahrt überqueren wir laut Skript eine Brücke, hinter der wir auch anhalten (siehe Skript, S. 39). Diese Brücke habe ich lediglich durch den Wasserfall hörbar gemacht, an dem wir vorbeifahren. Wie bei den Autos heben auch die Geräusche des Wasserfalls leise und schmal vorne links an und werden beim Vorbeifahren breiter und lauter, während sie anschließend hinten links wieder schmaler und leiser werden. Um auch hier den Effekt der geschlossenen Fenster wahrnehmbar zu machen, habe ich einen 12 dB/oct Tiefpassfilter mit der Grenzfrequenz bei etwa 1.100 Hz über den Wasserfall gelegt.

### 9.1.2 Ausstieg/Draußen

Nachdem Morton an den Straßenrand gefahren ist, entscheiden sich die drei Ermittler und der Chauffeur, den Wagen zu verlassen, um ihn auf Manipulationen hin zu untersuchen; hatten sie doch auf einem Foto gesehen, dass sich jemand am Wagen zu schaffen gemacht hat. Zunächst habe ich vier Türen öffnen lassen, die Fahrer- und Beifahrertür vorne links/rechts sowie die beiden hinteren Türen, an denen Peter und Justus sitzen. Dafür verwendet habe ich vier Aufnahmen von sich öffnenden und schließenden Autotüren aus Soundly, die ich mit

dem MultiEncoder auf die gerade beschriebenen Positionen im Raum gelegt habe. Für die wahrgenommene korrekte Lokalisation war ich erneut gezwungen, negative Elevationswinkel zwischen  $-10^\circ$  bis  $-20^\circ$  zu nutzen. Bei den vorderen Türen musste ich die Klänge wieder recht nah an der Medianebene platzieren, um die virtuellen Begrenzungen des Autos nicht zu zerstören. Um das Öffnen der Türen zu vermitteln, habe ich in einer kurzen Fahrt die Grenzfrequenz des Tiefpassfilters, der auf dem Wasserfall liegt, so lange erhöht, bis keinerlei Filterung mehr vorhanden war. Zusätzlich sind nun Naturgeräusche zu hören, die ebenfalls mit einer kurzen Tiefpass-Filterfahrt unterlegt sind.

Uns als Hörer:innen, die in der Mitte der Rückbank sitzen, habe ich aus der hinteren linken Tür aussteigen lassen, durch die auch Peter den Wagen verlässt. Prinzipiell wird der Eindruck von Bewegung unseres virtuellen Charakters durch die physische Bewegung der Rezipient:innen und durch die Verschiebung der Klangobjekte vermittelt. Für die Darstellung der virtuellen Blickrichtung unserer Figur musste ich verschiedene Ansätze ausprobieren. Da ich mich stets an der realen Welt orientiert habe, war auch hier die erste Überlegung, wie man denn tatsächlich einen Wagen verlässt und wohin man dabei schaut. Nach eigener Beobachtung wendet man sich zunächst in die Richtung, in die man aussteigt, während man sich im anschließenden Prozess des Aussteigens um die eigene Körperachse dreht. Genau das war im Sounddesign mein erster Ansatz: Ich habe die Charaktere, die Geräusche von sich geben, während sie aussteigen, sowie den Wasserfall sich drehen lassen, um die Kopfbewegung unseres fiktiven Charakters zu simulieren. So wanderte zum Beispiel Morton, der in Fahrtrichtung vorne links zu hören war, auf einmal nach rechts. Justus, der rechts neben uns sitzt, würde aus der anderen Tür aussteigen und sich nun hinter uns befinden. Doch recht schnell kam ich zu der Erkenntnis, dass diese Drehung der kompletten Szene bei der Wiedergabe für große Verwirrung sorgt. Daher entschloss ich mich dazu, die Blickrichtung während des Aussteigens weiterhin vorne durch die Frontscheibe des Wagens zu belassen und nur die Drehung um die eigene Körperachse nach dem Ausstieg zu simulieren. So blieben die Charaktere erst einmal auf ihren virtuellen Positionen, erst nach dem Ausstieg begaben sie sich auf ihre neue Position, die ich ihnen zugewiesen habe. Diese Bewegungen habe ich mit aufgenommenen Schrittgeräuschen für jeden einzelnen Charakter und Tönen sich bewegnender Kleidung unterlegt. Die Schritte liegen dabei auf der unteren Hemisphäre bei etwa  $-45^\circ$  Elevation, über den Azimut-Potentiometer habe ich die Figuren auf ihre neue Position geführt.

### 9.1.2.1 *Charaktere*

Wir stehen nun neben dem Wagen, die Straße in unserem Rücken, der Blick richtet sich über Mortons Rolls-Royce. Peter steht rechts neben uns, Justus uns gegenüber auf der anderen Seite des Wagens. Morton befindet sich zu unserer Linken, ihm gegenüber steht Bob. Mit



stärke und durch etwas EQing, zum Beispiel Absenkung der Tiefen, die Figuren näher in die Sphäre oder außerhalb der Sphäre platzieren. So geschehen bei Justus und Bob, die in diesem Teil der elften Szene weiter entfernt von uns stehen, als es Morton und Peter tun.

#### 9.1.2.2 *Die Atmosphäre*

Mit dem Öffnen der Türen beginnt man nun auch die Natur und die Umgebung zu hören. Da sich Redwood Falls in einem Wald mit Fluss und Wasserfall befindet, habe ich das Sounddesign für den Ort des Ausstiegs ähnlich gewählt. Das mir vorschwebende Bild ist eine zweispurige Straße gewesen, die vorbei an dem Wasserfall durch einen idyllischen Wald führt. Zu hören ist immer noch das entfernte Gewittergrollen am Himmel, dabei ist der Boden aber trocken und es regnet nicht. Weil wir uns nun draußen befinden, habe ich den Tiefpass auf den Donner-Spuren komplett entfernt. Die Spuren gehen trotzdem noch auf den LFE und liegen auf der oberen Hemisphäre zwischen 50° bis 90° Elevation.

Das akustische Gerüst der Außenatmosphäre besteht aus fünf Spuren. Als Erstes haben wir eine Stereoaufnahme von leichtem Wind, der durch Blattwerk weht. Diese habe ich wiederum mit dem Halo-Upmix-Tool auf der oberen Hemisphäre im 7.1.4-Schema verteilt. Dasselbe habe ich mit einer eigenen Aufnahme von mäßigem Wind ohne Vögel, Verkehr, Menschen oder Ähnlichem getan, um dem Ganzen mehr Struktur zu geben. Auf der dritten Spur findet sich eine Aufnahme aus einem Wald, der von böigem Wind durchweht wird; diese Spur habe ich leicht dazugemischt, um mehr Dynamik in das Geschehen zu bringen. Bei dieser Spur habe ich allerdings den StereoEncoder verwendet und die Spur auf 65° Width gebracht, während ich das Panning gefahren bin. Dabei habe ich mich ausschließlich auf dem oberen Teil der Sphäre bewegt.

Wie nutze ich aber nun die untere Sphäre bei der Gestaltung der Außenatmo? Was für Klänge lege ich nach unten, und was hören wir überhaupt von unten, speziell an so einem Ort? Um das herauszufinden, habe ich verschiedene Aufnahmen aus Wäldern auf die untere Hemisphäre gelegt. Doch das von unten kommende Rauschen der Blätter wirkte irritierend. Deshalb habe ich zwei von mir aufgenommene Wind-Spuren, mit jeweils unterschiedlichen Klangfarben, mittels des Halo-Tricks auf den unteren Teil der Hemisphäre gelegt. Und das hat außerordentlich gut funktioniert. Beim Hören lässt sich nicht genau identifizieren, was nun von unten kommt, doch der Einsatz dieses Mittels rundet die Illusion ab, in einem Wald zu stehen.

Um auch den Wald lebendig zu gestalten, habe ich verschiedene Aufnahmen von Vögeln herangezogen und deren Stimmen ab und zu erklingen lassen. Auch hier bin ich stets auf dem oberen Teil der Hemisphäre geblieben. Bei der Positionierung habe ich variiert und bin vereinzelt das Panning gefahren, so zum Beispiel bei der Krähe.

### 9.1.2.3 Die Schlucht

Nachdem Morton den Wagen inspiziert hat, entdeckt Peter ein Autowrack auf einem Felsvorsprung im Abhang einer Schlucht direkt neben der Straße, an der sie angehalten haben. Musikalisch sind Bass-Flächen zu hören, die dramaturgisch auf den Zusammenhang mit dem vermissten Mr. Jacobs hinweisen. Dieser Bass ist wie oben beschrieben auf der Sphäre verteilt zu hören. Um den Hörenden ein erstes Bild dieser Schlucht zu geben, tritt Peter einen Stein hinab. Dieser Klang des hinabstürzenden Steins besteht in meinem Fall aus vier Klängen. Zunächst haben wir einen sehr transienten-lastigen Impuls, der durch einen fallengelassenen Stein erzeugt wird. Dieser löst wiederum weitere Steine vom Rand der Klippe und reißt sie mit hinab. Da Peter den Stein wie erwähnt hinuntertritt, liegen alle vier Spuren auf der unteren Hemisphäre bei etwa  $-60^\circ$  Elevation. Um den Fall der Steine zu simulieren, habe ich sie im Verlauf des Falls noch weiter nach unten geschwenkt, auf bis zu  $-80^\circ$  Elevation. Außerdem werden die Geräusche während des Falls nach und nach leiser. Zusätzlich habe ich einen Hall auf die Spuren gegeben, um dem Raum der Schlucht Struktur und Grenzen zu geben. Bei der Nachhallzeit habe ich allerlei experimentiert, bis ich das am meisten zufriedenstellende Ergebnis bei 5 Sekunden und einer Dry/Wet Einstellung von 0.55 erzielte.

## 9.2 Szene 12

Die drei Fragezeichen haben sich entschieden, Peter an einem Abschleppseil in die Schlucht abzulassen, wo er das im Abhang hängen gebliebene Autowrack, das dem verschwundenen Mr. Jacobson zu gehören scheint, nach Spuren absucht.

### 9.2.1 Die Schlucht

Wenn ich an eine Schlucht denke, kommt als Erstes das Bild eines steilen Abhangs auf, der räumlich nicht begrenzt ist. Das Sounddesign würde sich demnach atmosphärisch allerdings kaum von der vorherigen Szene abheben. Deshalb habe ich in der Klanggestaltung statt einer Schlucht, wie ich sie mir vorstelle, eine Höhle ausgewählt. Der Wechsel von Szene 11 zu 12 versetzt einen auf diese Weise in eine von Grund auf andere Umgebung als zuvor.

### 9.2.2 Die Blende

Wie geht man nun aber mit einem solch drastischen Szenenwechsel um, wo wir als diegetische Figur doch an die Regeln der virtuellen Welt gebunden sind? Der Orts- und Zeitwechsel wird über die Musik vermittelt. Während die Musik zu spielen anfängt, blende ich in den neuen virtuellen Raum der Höhle, indem ich die zwei verschiedenen Atmo-Busse der Szene 11 und

12 ineinander übergehen lasse. Für diesen kurzen Moment von ungefähr acht Sekunden sind wir als Hörende gleichsam körperlos und springen direkt zu unserer nächsten virtuellen Position. Für diesen Sprung habe ich die komplette Szenerie sich drehen lassen, um so die neue Position zu verdeutlichen. Als Orientierungspunkt dient uns wieder der Wasserfall, der sich nun zu unserer Linken befindet. Dieser wandert während der Blende langsam hinter uns von der rechten auf die linke Seite, gemäß der Aufteilung der von mir erschaffenen Welt.

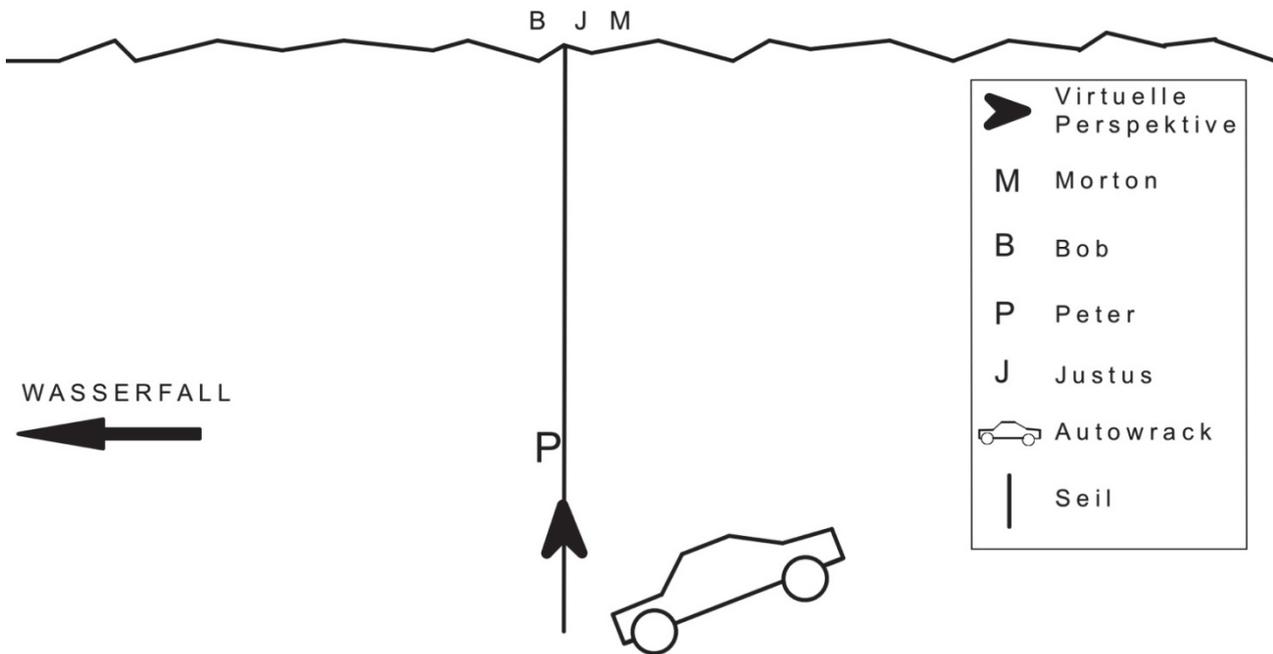
### 9.2.3 Die Charaktere

Da ich die Perspektive eines vierten Ermittlenden gewählt habe, standen mir in Szene 12 zwei schlüssige Optionen für dessen Platzierung als virtueller Charakter zur Verfügung: Entweder diese Figur steht oben bei Justus, Bob und Morton, oder sie seilt sich mit Peter in die Schlucht beziehungsweise Höhle hinab.

Die erste Option habe ich zuerst ausprobiert. Dabei standen wir zwischen Justus und Bob, neben dem noch Morton stand. Der virtuelle Blickwinkel war dabei stets nach unten gerichtet. Folglich habe ich die zuvor oben liegenden Vögel auf die hintere Seite der Hemisphäre bei etwa 20° Elevation gelegt. Schnell wurde klar, dass dieser Ansatz für Verwirrung sorgt, weshalb ich an der Stelle abgebrochen und Option zwei versucht habe.

In diesem zweiten Szenario befinden wir uns hinter Peter, der sich gerade an dem Abschleppseil in die Höhle hinablässt. Morton, Justus und Bob stehen oben am Hang und blicken auf uns herab. Für die Platzierung Peters habe ich zunächst den RoomEncoder verwendet; leitend war die Idee, die Höhle als virtuellen Raum in dem Plugin zu entwerfen. Dort würden mir die Reflexionen berechnet werden, ich müsste mich auch nicht mehr um die akustische Gestaltung kümmern. Leider sind die Maße der Seiten auf 30 m \* 30 m \* 20 m beschränkt, so dass sich nicht die Räumlichkeit erzielen ließ, die meiner Vorstellung von einer Höhle entsprach. Also griff ich auf den MultiEncoder zurück und stellte die akustischen Eigenschaften der Höhle über den FdnReverb, den DualDelay und den MultiEQ (Plugins der IEM Plugin-Suite) dar.

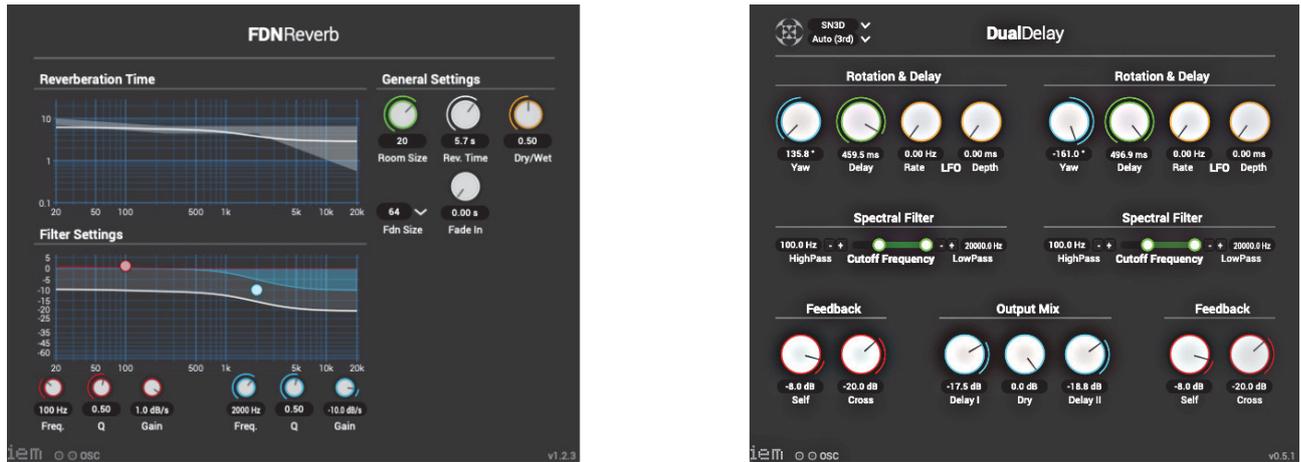
Da sich Peter direkt vor uns befinden sollte, liegt er im MultiEncoder auf 0° Azimut. Für die korrekte Darstellung musste ich ihn allerdings um -22° Elevation hinabversetzen, sonst wäre die Stimme eher von oben wahrzunehmen. Während wir uns als Figur die letzten Meter noch hinabseilen müssen, ist Peter bereits auf dem Felsvorsprung angekommen und geht in Richtung Wrack. Um das zu erreichen, wird Peter auf -41° Elevation hinabgelassen und bewegt sich auf -55° Azimut rechts von uns. Peters Bewegungen werden durch Geräusche seiner Kleider untermalt.



**Abbildung 9.5:** Darstellung der virtuellen Umgebung in Szene 12

Justus und Bob, die oben an der Klippe stehen, reden unterdessen mit uns. Entsprechend habe ich beide Detektive auf 75° Elevation gebracht. Bei höheren Winkeln würden sie nicht mehr oben vorne, sondern direkt über uns wahrgenommen werden, womit das imaginierte Bild der Szenerie zerstört würde. Um die Entfernung zu verdeutlichen und den Klang der Höhle zu gestalten, bediente ich mich der oben genannten Hall- und Delay-Plugins und des MultiEq. Das beste Ergebnis bekam ich bei einer Nachhallzeit von 5,7 Sekunden, bei einer Raumgröße von 20 m. Da ich mir keinen Hall-Bus gebaut habe, lag auf jeder Sprachspur ein solcher Hall. Justus und Bob wurde dieser Hall zu 50% zugemischt, während ich Peter nur 8% seines Hallsignals beimischte. Auch dadurch wurden die Perspektive und die virtuellen Positionen beibehalten.

Abgesehen vom Hall habe ich versucht, die Größe der Höhle auch durch einen Delay wahrnehmbar zu machen. Die Idee war, ein Echo zu erzeugen, das sich nicht nur statisch an einem Punkt im Raum befindet, sondern das Sprachsignal von den verschiedenen virtuellen Wänden und Steinen der Höhle reflektiert. Das DualDelay der IEM Plugin-Suite bietet dafür zwei individuell einstellbare Delays, die sich über den „Yaw“-Potentiometer im Raum verteilen lassen. Um der Dimension der Höhle gerecht zu werden, entschied ich mich für Zeiten um das Maximum des Plugins von 500 ms. Die beiden Delays bekamen jeweils unterschiedliche Verzögerungszeiten von 450 ms und 496 ms, um der Höhle mehr Struktur zu verleihen. Von den reflektierten Signalen Justus' und Bobs habe ich -18 dB hinzugemischt. Bei Peter habe ich nur -30 dB hinzugemischt, da wir uns näher bei ihm befinden und er sich mit zu viel Delay weiter von uns wegbewegen würde. Zum Signalfloss ist zu sagen, dass ich das beste Ergebnis erzielte, als ich den Delay in Serie vor den Hall gelegt habe.

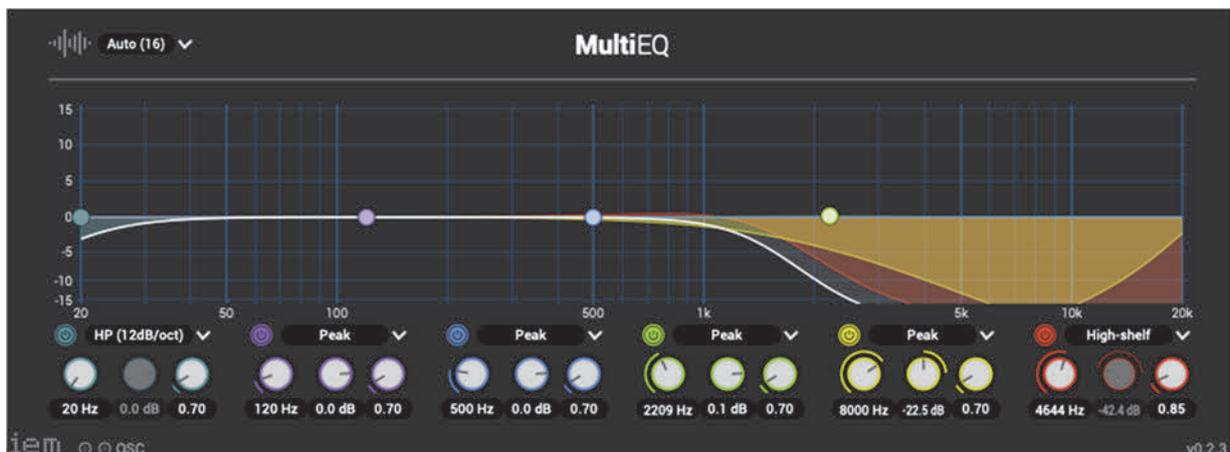


**Abbildung 9.6:** IEM FdnReverb (links) und IEM DualDelay (rechts) mit den Einstellungen für Justus und Bob aus Szene 12

Um Bob und Justus etliche Meter von uns entfernt zu platzieren, habe ich ihre Spuren noch mit dem MultiEq bearbeitet. Mein dabei gewählter Ansatz kommt aus Kapitel 4.7 beziehungsweise aus der in der Nachrichtentechnik bekannten Gesetzlichkeit, nach der mit steigender Frequenz die Freiraumdämpfung  $\alpha_0$  zunimmt.

$$(\alpha_0 = \left(\frac{4 \cdot \pi \cdot r}{\lambda}\right)^2; \lambda = \frac{c_0}{f}).$$

Folglich können tiefere Frequenzen größere Distanzen überwinden als höhere Frequenzen. Um dieses Gefühl von Entfernung zu erzeugen, habe ich mit einem Highshelf-Filter ab 4,5 kHz eingegriffen. Da mir die Flankensteilheit des Filters zu gering war, habe ich einen zweiten EQ-Punkt verwendet und bei 8 kHz um -22,5 dB abgesenkt. Bei der Grenzfrequenz des Highshelf-Filters habe ich darauf geachtet, dass Justus und Bob trotzdem noch zu verstehen sind.



**Abbildung 9.7:** IEM MultiEQ auf dem Kanal von Justus und Bob aus Szene 12

#### 9.2.4 Atmosphäre

Um den örtlichen Zusammenhang der Szene 11 und 12 zu wahren, ist in der Höhle auch das Rauschen des Windes und der Blätter zu hören. Im Sinne der Authentizität habe ich die Spur mit einem Tiefpassfilter bearbeitet, damit die Perspektive nicht vertauscht wird. Die Wind-Aufnahme habe ich mittels Halo auf die obere Hemisphäre gelegt. Dort teilt sie sich den Platz mit einer Stereoaufnahme aus dem Innern einer Höhle, bei der vor allem der Wind zu hören ist, der an deren Öffnung vorbeiweht. In der Mischung ist die Atmos aus Szene 11 6 dB leiser als die Höhlenaufnahme. Diese Spur habe ich dupliziert und mit dem Halo-Trick nach unten gelegt. Der Wasserfall ist, wie oben beschrieben, auf unsere linke Seite gewandert. Um auch hier den physikalischen Gesetzen meiner fiktiven Welt zu entsprechen, habe ich den Wasserfall mit einem 12 dB/oct Tiefpassfilter, bei einer Grenzfrequenz von 3,1 kHz, dumpfer gemacht. Hall habe ich nicht auf den Wasserfall gerechnet, da dieser schon zu diffus ist. Durch diese Filterung ist es noch schwieriger geworden als ohnehin schon, das Klangobjekt als Wasserfall zu erkennen. Um dem Problem abzuhelpfen, habe ich mich einer Höhlenaufnahme bedient, in der ein Fluss fließt. Von dessen plätschernden Klängen hatte ich mir erhofft, dass sie das imaginierte Bild des Wasserfalls festigen. Da sich ein solches Gewässer unter uns befinden würde, habe ich auch hier wieder zum Halo-Trick gegriffen und das nun hochgemischte Soundscape auf der unteren Hemisphäre verteilt. Um den Vorgang des Hinabseilens lebhafter zu gestalten und etwas mehr Dynamik und Spannung in die Szene zu bringen, habe ich immer wieder Steine virtuell an uns vorbei hinabstürzen lassen, wobei die untere Hemisphäre sehr hilfreich war. Dabei bin ich wie in 9.1.2.3 vorgegangen.

#### 9.2.5 Das Autowrack

Nachdem Peter und wir als virtuell Mit-Ermittelnde am Autowrack angekommen sind, ist zu hören, wie Peter die Tür des Autos öffnet. Dafür habe ich mir die Aufnahme eben einer sich öffnenden Autotür herausgesucht und diese dann mit dem Geräusch quietschenden Metalls überlagert, um den defekten Zustand des Autos zu vermitteln. Da wir uns noch über Peter befinden, liegt das Geräusch der sich öffnenden Tür bei  $-21^\circ$  Elevation. Sobald wir das Innere des Wagens betreten, nimmt die zuvor gehörte Atmo um etwa 6 dB ab. So gelingt es deutlich besser, den Eindruck des Hineinsteigens zu erwecken. Zur Verbildlichung des Hineinsteigens habe ich dem Autowrack eine eigene Atmosphäre gegeben. Um die Situation zu überspitzen und dramatischer wirken zu lassen, habe ich hierfür Aufnahmen aus einem sich bewegenden Schiffscontainer genutzt, die ich mittels Halo-Upmix sowohl auf die obere als auch auf die untere Sphäre platziert habe. Zusätzlich habe ich zwei weitere Tonaufnahmen von sich bewegendem Metall herangezogen und diese dynamisch auf der kompletten Sphäre

verteilt. Was also eigentlich das Autowrack eines mutmaßlichen Viertürers ist, habe ich in einen riesigen Metallkoloss verwandelt. Da ich zwei dieser Metall-Spuren auf den LFE geroutet habe, wird dieser Effekt dadurch nochmals verstärkt. Aufgrund der physischen Nähe zu den Lautsprechern und der in R3 kleinen Sphäre wirkt dieses Metallungetüm akustisch aber bei weitem nicht so groß.

Im Wageninnern befindet sich Peter rechts von uns. Mit eigenen Foley-Aufnahmen und mittels fertiger Klänge habe ich den Vorgang darzustellen versucht, wie das Fahrzeug durchsucht und abgetastet wird. Weil davon auszugehen ist, dass dieses Durchsuchen unterhalb unseres Kopfes stattfindet, habe ich die Klänge zwischen  $-4^\circ$  bis  $-41^\circ$  Elevation platziert und sie sich dynamisch mit Peters Stimme bewegen lassen.

Die Spannung findet ihren Höhepunkt, als das im Abhang festhängende Wrack weiter hinabzurutschen droht. Die einsetzende Abrutschbewegung habe ich durch vier kombinierte Klänge zum Ausdruck gebracht. Zu hören ist, wie links unten hinter uns, bei etwa  $-27^\circ$  Elevation, Metall unkontrollierbar zu rütteln anfängt. Dieses stetig zunehmende Rütteln endet mit drei übereinandergelegten Aufnahmen von einschlagendem Metall. Dieser Einschlag ist die akustische Übertreibung des Geräuschs, welches das Autowrack erzeugt, während es sich vom Felsen löst. Um dieses Geschehen noch deutlicher zu untermalen, habe ich die gestapelten Spuren auf den LFE geroutet.

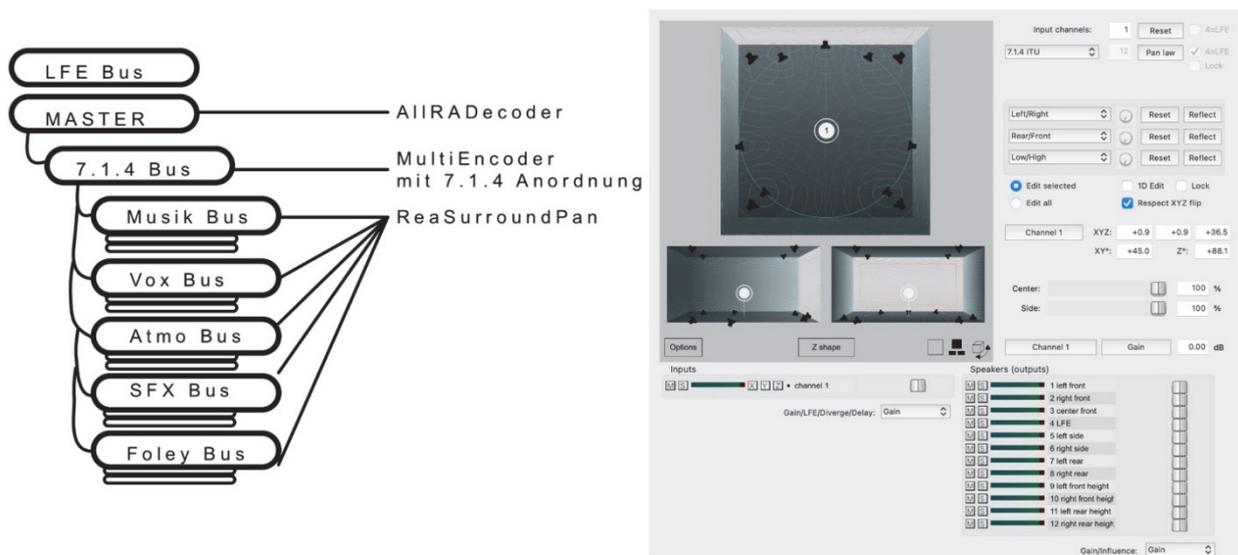
In diesem Moment setzt die Musik ein, die ich – wie oben beschrieben – umhüllend auf die Sphäre gelegt habe. Peter schreit rechts neben uns auf und fleht Justus und Bob an, ihn wieder hinaufzuziehen. Um auch diesen Vorgang zu verbildlichen, habe ich das Spannen eines Taus direkt über uns bei  $85^\circ$  Elevation aufklingen lassen. Dadurch wird metaphorisch vermittelt, dass Peter nur noch an dem Seil hängt, das ihn vor dem Hinabstürzen bewahrt.

## 10 7.1.4

Die Klänge des Sounddesigns und deren Platzierungen haben sich für das 7.1.4-Setup nicht verändert. Im Folgenden werde ich erläutern, inwiefern ich anders mit den Klängen umgegangen bin; denn schließlich stand mir bei der Wiedergabe über 7.1.4 nur die obere Hemisphäre zur Verfügung. Um ein ähnliches Ergebnis zu erzielen, habe ich mir das gerenderte File der True-Spatial-Mischung in die 7.1.4-Session geladen, was mir einen direkten Vergleich ermöglichte.

Für das Panning habe ich den ReaSurroundPan der Firma Cockos verwendet. Dieses Panning-Tool verfügt über einen integrierten virtuellen Aufbau des 7.1.4-Wiedergabesetups nach ITU-Norm. Die Anzahl der Kanäle kann man in der oberen rechten Ecke beliebig hoch ansetzen (nach 10.000 habe ich aufgehört zu zählen), was auch die Arbeit mit Stereo, Quadro, 5.1 oder höherkanaligen Aufnahmen zulässt. Wie bereits erwähnt lässt sich das Klangobjekt hier auch ganz einfach in den virtuellen Raum legen.

Der Signalfluss ist ähnlich dem des True-Spatial-Setups. Den entscheidenden Unterschied macht der 7.1.4-Bus. Hier liegt der MultiEncoder, mit dem ich die virtuellen Lautsprecher in dem nach ITU genormten Setup angelegt habe. Bis auf den LFE gehen alle Spuren auf diesen 7.1.4-Bus, dieser wiederum auf den Master und von dort über den AllRADecoder auf die Ausgänge.



**Abbildung 10.1:** Spurenverteilung in der 7.1.4-Mischung mit jeweiligem En-/Decoder (links) und der ReaSurroundPan der Firma Cockos (rechts)

## 10.1 Szene 11

### 10.1.1 Autofahrt

#### 10.1.1.1 *Das Auto*

Für den quietschenden Klang, der die Bewegung der Karosserie vermitteln sollte, habe ich denselben Ansatz wie in der True-Spatial-Mischung gewählt. Dabei habe ich die Stereoaufnahme mit dem Halo-Upmix im 7.1.4-Schema auf der oberen Hemisphäre verteilt. Die beiden SUV-Motoren-Aufnahmen lösten bei mir in der True-Spatial-Mischung ein subjektiv immersives Gefühl aus, als ich beide Spuren auf der unteren Hemisphäre verteilt habe. Dies war nun aber keine Option mehr. Um trotzdem größtmögliche Tiefe zu erzielen, habe ich die Motoren bei 0° Elevation jeweils bei  $\pm 135^\circ$  und  $\pm 70^\circ$  Azimut platziert. Zusätzlich gingen die beiden Spuren mit -6 dB auf den LFE. Angesichts der unterentwickelten Fähigkeit des Menschen, tiefe Frequenzen zu lokalisieren, hatte ich mir hier umso mehr erhofft, dass das Fehlen der unteren Hemisphäre über den LFE kompensiert werden könnte.

Für die drei Geräusche der Reifen, die über den Schotterweg rollen, hatte ich einen ähnlichen Ansatz verfolgt: Kann ich die Klänge nicht auf die untere Sphäre legen, platziere ich sie so tief, wie es mir das Setup erlaubt. So geht eine der Spuren auf  $\pm 130^\circ$  Azimut bei 0° Elevation, während die zweite Spur bei  $\pm 70^\circ$  Azimut ebenfalls auf 0° Elevation liegt. Um die Verbreitung des Schalls im Wageninnern dazustellen, habe ich die dritte Spur mit dem Halo-Upmix-Tool auf der oberen Hemisphäre verteilt und bei 4,5 kHz einen 12 dB/oct Tiefpassfilter gelegt.

Auch in der 7.1.4-Mischung kommen uns während der Fahrt zwei Autos entgegen. Die Filterung der beiden Spuren erfolgt genauso wie in der True-Spatial-Variante mit einem 12 dB/oct Tiefpass bei 10 kHz. Beim Panning habe ich zunächst versucht, die Autos auf 0° Elevation zu legen, was der Höhe unserer Ohren entspricht. Allerdings hörte ich die Wagen räumlich deutlich tiefer, auf der unteren Hemisphäre. Dem versuchte ich entgegenzuwirken, indem ich mich allmählich höher tastete, bis ich bei etwa 30° Elevation den räumlich-akustischen Eindruck hatte, die Autos würden uns auf gleicher Höhe passieren. Um das Vorbeifahren wirklichkeitsgetreu zu simulieren, habe ich auch hier die Autos mit kleiner werdendem Abstand im ReaSurroundPan breiter werden lassen.

Bei den Blinkern, die Morton setzt, nachdem sie das Foto auf Bobs Handy genauer betrachtet haben, stellte sich dieselbe Problematik wie in der True-Spatial-Mischung: Um das Geräusch des Blinkers vorne links wahrzunehmen, musste ich sehr nah an die Medianebene (10° Azimut) rücken. Da mir die untere Hemisphäre fehlte, habe ich die Blinker bei 0° Elevation platziert, wobei man sie trotzdem noch weiter oben empfunden hat. Dasselbe Problem trat

auch bei der Handbremse und dem Zündschlüssel auf, die ich ebenfalls so tief wie möglich bei 0° Elevation platziert habe.

Mit den Kernklängen des nun stehenden Wagens bin ich ähnlich verfahren wie in der True-Spatial-Variante. Der Bremsvorgang beziehungsweise das Abbremsen der Reifen liegt hier ebenfalls relativ schmal bei 90° Elevation. Das Gleiche gilt für den bass-lastigen und den mit-tenbetonten Motor, die ich über Halo auf der oberen Hemisphäre im 7.1.4-Schema verteilt habe. Für den dünn klingenden Motor musste ich mir etwas Neues einfallen lassen. Nach einigen Versuchen bekam ich das am intensivsten umhüllend wirkende Ergebnis, als ich den linken Kanal auf den vorderen linken Lautsprecher und den rechten Kanal auf den Lautsprecher zu meiner rechten Seite gelegt habe. Dabei blieb ich bei 0° Elevation. Für einen sauberen Übergang der beiden unterschiedlichen Innenatmos des Autos habe ich mir auch hier zwei Atmo-Busse angelegt, die ich ineinander gemischt habe.

#### 10.1.1.2 Charaktere

In der 7.1.4-Variante herrscht die gleiche Sitzanordnung, wie ich sie oben erläutert habe: Justus sitzt zu unserer Rechten, Peter zu unserer Linken, Bob auf dem Beifahrersitz und Morton hinter dem Steuer (vgl. Abbildung 9.3). Für die Platzierung habe ich, bis auf Morton, alle Charaktere bei 0° Elevation platziert. Mortons Stimme empfand ich als tiefer liegend als jene der drei Fragezeichen, weshalb er in etwa 11,5° Elevation liegt.

Erst im Nachhinein kam mir die Idee, dass ich für die halbsphärische Mischung in 7.1.4 auch den RoomEncoder hätte benutzen können. Da ich das während der Mischung aber nicht im Sinn hatte, habe ich die Akustik des Wageninnern mit dem FdnReverb simuliert. Hierfür habe ich mich für eine Nachhallzeit von 200 ms entschieden, wobei der Nachhall zu 50% den trockenen Sprachspuren hinzugemischt wurde. Um die virtuelle Distanz zwischen uns und den einzelnen Charakteren zu wahren, habe ich diese durch Pegel zu simulieren versucht.

Bei den Geräuschen der Sitze bin ich, was die Pegel angeht, genauso vorgegangen wie in der True-Spatial-Mischung und habe sie zu Beginn etwa 4 dB lauter gepegelt, um die Nervosität zu verdeutlichen. Bei der Elevation blieb ich bei 0°. Das hatte zur Folge, dass ich das Herumrutschen auf den Sitzen auf derselben Höhe wahrgenommen habe wie die Stimmen.

Für die Dorfbewohner, die sich zu Beginn der elften Szene von den drei Ermittlern und ihrem Chauffeur Morton verabschieden, habe ich dieselben Aufnahmen vom 17. Mai aus A1 des Tonlabors der HAW DMI verwendet. Ebenso habe ich die Stimmen in etwa 20° Elevation belassen und sie links wie rechts am Wagen vorbei ertönen lassen. Um darzustellen, dass sich die Bewohner außerhalb des Wagens befinden, habe ich auch hier einen 12 dB/oct Tiefpassfilter mit der Grenzfrequenz von 100 Hz gewählt. Für die virtuelle Distanzierung des Rolls-Royce vom Dorf habe ich die Stimmen immer leiser und hinter uns enger werden lassen.

### 10.1.1.3 Das Wetter

Für das „drohende Gewittergrollen“ (Skript, S. 38) habe ich mich derselben Klänge aus der True-Spatial-Mischung bedient und diese mit dem ReaSurroundPan auf der oberen Hemisphäre automatisiert. Um die Breite und Richtung des Signals wahrnehmbar zu machen, habe ich auch hier versucht, einen sich ausbreitenden Donner am Himmel zu simulieren. Dem Donner habe ich wiederum dadurch mehr Lebendigkeit zu vermitteln versucht, dass ich die Spuren ebenfalls auf den LFE mit -6 dB geroutet habe. Da wir das Donnern im Innern des Wagens durch verschlossene Fenster hören, habe ich die Spuren mit einem 12 dB/oct Tiefpassfilter mit der Grenzfrequenz 15 kHz bearbeitet.

Auch in der 7.1.4-Mischung dient uns vor allem der Wasserfall zur Orientierung. Dieser besteht aus denselben sechs Aufnahmen und Spuren wie in der True-Spatial-Variante. In 7.1.4 blieb mir für die Gestaltung eines real wirkenden Wasserfalls nur die obere Hemisphäre. Da ich schon in der True-Spatial-Mischung gemerkt habe, dass die Aufteilung des Wasserfalls in verschiedene Layer kaum die erhoffte Wirkung hatte, habe ich mir auch in 7.1.4 einen Wasserfalls-Bus angelegt, über den ich die Summe der sechs Spuren im Raum verteilt habe. Leider trat auch hier das Problem auf, dass ich das Rauschen des Wasserfalls subjektiv von zu weit oben wahrnahm. Um dem so weit wie möglich gegenzusteuern, habe ich den Wasserfall-Bus auf 0° Elevation platziert, was die empfundene Höhe kaum veränderte.

Für die Überfahrt der Brücke habe ich mit dem ReaSurroundPan eine Automation geschrieben, die das Vorbeifahren simulieren soll. Um den Vorgang möglichst detailgetreu darzustellen, habe ich wieder auf einen 12 dB/oct Tiefpassfilter zurückgegriffen, der eine Grenzfrequenz von 1,1 kHz hat.

### 10.1.2 Ausstieg/Draußen

Für das Aussteigen der drei Fragezeichen und Mortons aus dem Auto habe ich die gleichen Aufnahmen von sich öffnenden und schließenden Türen wie in der ersten Mischung genutzt und diese so in unserem virtuellen Raum verteilt, dass die Illusion eines Autos nicht durch irritierende Dimensionen zerstört wird. Anders als beim MultiEncoder hat die Platzierung im ReaSurroundPan bei jeweils 20° Elevation zur korrekten Wahrnehmung geführt.

Um den Übergang der zwei Räume, vom Wageninnern in die Szenerie außerhalb des Autos, zu gestalten, habe ich mich ebenfalls entschieden, eine Filterfahrt des Tiefpassfilters einzusetzen, der auf dem Wasserfall liegt, und die Außenatmos einzublenden.

Die Illusion des Aussteigens habe ich dadurch zu steigern versucht, dass ich die Aufnahmen von Schritten wieder in das Projekt geladen habe. Die in der True-Spatial-Mischung unten liegenden Schritte liegen nun bei 0° Elevation. Das ruiniert allerdings subjektiv, wenn

auch nur kurzfristig, die wechselseitigen Beziehungen der einzelnen Klangobjekte, in diesem Fall der Stimmen und der Schritte. Trotzdem habe ich auch hier über eine kurze Fahrt des Azimut-Reglers die Charaktere auf ihre neuen Positionen begleitet.

#### 10.1.2.1 Charaktere

Im zweiten Teil der Szene, der außerhalb des Wagens stattfindet, habe ich den FdnReverb auf den Sprachspuren deaktiviert und die Stimmen wie in Abbildung 9.4 dargestellt platziert.

Morton soll hier zu Beginn unterhalb des Wagens nach Manipulationen durch Dritte Ausschau halten. In True Spatial konnte ich das einfach durch die Platzierung auf der unteren Hemisphäre zu unserer Linken darstellen. In 7.1.4 war dieselbe Maßnahme mit weniger Erfolg gekrönt. Ich habe Morton linksseitig neben uns platziert und durch ein leichtes Anheben der Elevation von  $0^\circ$  auf  $10^\circ$  und des Pegels um 3 dB versucht, das Hervorkrabbeln und Wiederaufstehen zu simulieren. Genauso habe ich es auch mit der Foley-Spur getan, in der zu hören ist, wie Morton sich von Staub befreit, indem er seine Kleider abklopft. Für das Geräusch der Schritte, die Morton vollzieht, als er hinter uns zum Kofferraum geht, um das Abschleppseil zu holen, habe ich das ähnlich gehandhabt. Die Schrittgeräusche habe ich bei  $0^\circ$  Elevation im ReaSurroundPan halbkreisförmig hinter uns gefahren. Durch leichte Anpassungen des Pegels um etwa 2 dB habe ich die Distanz, die Morton zurücklegt, etwas gesteigert.

Um die Distanzen zwischen uns und jedem Einzelnen der drei Fragezeichen zu verdeutlichen, habe ich deren Pegel entsprechend angepasst. Da bei der 7.1.4-Wiedergabe derselbe Lautsprecheraufbau wie in True Spatial verwendet wurde, befinden wir uns in derselben kleinen Sphäre aus Lautsprechern. Dabei habe ich versucht, den virtuellen Raum dadurch größer wirken zu lassen, dass ich mittels leichter Absenkungen der Mitten und Höhen im EQ die Figuren und Klangobjekte, in dem Fall die drei Ermittler, außerhalb der physisch begrenzten Sphäre platzierte.

#### 10.1.2.2 Die Atmosphäre

Die Anzahl der von mir verwendeten Spuren für die Grundgestaltung der Außenatmosphäre hat sich im True-Spatial-Mix von fünf auf drei Spuren verringert. Die zwei verschieden klingenden Aufnahmen von Winden, die ich in True Spatial auf die untere Hemisphäre gelegt habe, fielen nun komplett weg.

Mit den restlichen drei Atmos musste ich in dieser Mischung anders umgehen. Zunächst habe ich getreu der Arbeitsweise der True-Spatial-Mischung die drei Aufnahmen über Halo auf der oberen Hemisphäre im 7.1.4-Schema verteilt. In der Summe habe ich das subjektiv anders wahrgenommen, als dies in True Spatial der Fall war. Die Atmos klangen zu rauschig, mit großen Löchern auf der hinteren Seite. Nach einigen Versuchen kam ich zu folgender

Lösung: Die aufgenommenen Geräusche des Windes, der durch Bäume und Blattwerk weht, blieben im 7.1.4-Schema auf der oberen Hemisphäre. Die eigene Aufnahme von leichtem Wind ohne weitere Klänge habe ich dupliziert und jeweils bei 27° Elevation auf  $\pm 45^\circ$  Azimut und  $\pm 132^\circ$  Azimut platziert. Die Atmosphäre hatte in der Folge deutlich mehr Klarheit und Struktur. Um auch in die Windgeräusche mehr Dynamik zu bringen, bin ich die Aufnahme des böigen Winds live gefahren und habe von 25° Elevation aufwärts variiert.

Bei den Vogelklängen bin ich genau wie in der True-Spatial-Variante vorgegangen und habe diese auf der oberen Hemisphäre bei +60° Elevation verteilt.

### 10.1.2.3 Die Schlucht

Das Fallen des Steines, den Peter bei der Entdeckung des Autowracks versehentlich in Bewegung versetzt, musste ich grundverschieden darstellen. Wo zuvor die untere Hemisphäre in hervorragender Weise zum Einsatz kam, musste ich mir nun etwas einfallen lassen, wie ich die Immersion und die Dimensionen des virtuellen Raums aufrechterhalte. Zuerst habe ich alle vier Spuren auf 0° Elevation gelegt und anschließend mit Hall, Eq und Pegeln versucht, die Steine tiefer als 0° Elevation fallen zu lassen. Das Fallen des Steins fängt auf dem vorderen Teil der Hemisphäre an der Horizontalebene an, setzt sich dann langsam in das Innere der Sphäre fort, bis es kurz vor unserem Kopf verstummt. So wollte ich vorerst die Bewegung darstellen. Da sich der Stein auf diese Weise aber wahrnehmbar nähert, habe ich durch leiser werdende Pegel, verstärktes Zumischen des Hallsignals und Reduzieren des Höhenanteils des Signals versucht, den Stein akustisch tiefer fallen zu lassen, als es mir das Setup ermöglicht. Das Hallsignal wird dabei über einen Zeitraum von ungefähr drei Sekunden von 25% auf 100% hinzugemischt. Die Nachhallzeit ist dieselbe wie in True Spatial, sie beträgt fünf Sekunden. Die Höhen habe ich mit einem 12 dB/oct Tiefpassfilter über denselben Zeitraum auf bis 8 kHz weggefiltert. Für die kurze Pegelfahrt habe ich eine lineare Automationskurve geschrieben, die bei -60 dB endet.

## 10.2 Szene 12

### 10.2.1 Die Blende

Da das in True Spatial erprobte Verfahren für den Übergang von Szene 11 nach 12 subjektiv gut funktionierte, bin ich in 7.1.4 ebenso vorgegangen und habe uns als Hörende kurzzeitig körperlos zu Peter hinabgelassen. Der Wasserfall, der hinter uns auf unsere linke Seite wandert, dient uns wieder zur Orientierung.

### 10.2.2 Die Charaktere

Wo Peter sich in True Spatial noch unter uns befunden hat, hängt er nun direkt vor uns am Seil. Ich hatte zuvor versucht, ihn – ähnlich wie bei dem herabfallenden Stein am Ende der elften Szene – über Eingriffe in Pegel, Hall und EQ unter uns zu platzieren. Doch das gelang leider nicht. Durch den vermehrten Einsatz von Hall und EQ war die Sprache im virtuellen Raum kaum noch verständlich und klang diffus. Außerdem war die Folge, dass die Stimme Peters nicht mehr mit dem virtuellen Klang der Höhle zusammenpasste, was sich sehr verwirrend anhörte. Da diese Mittel also nicht funktionierten, habe ich unsere Perspektive leicht angepasst und uns stetig hinter Peter herlaufen lassen.

Justus, Bob und Morton befinden sich oben am Rand der Schlucht, den ich akustisch als Höhleneingang gestaltet habe. Bei einem direkten Vergleich mit der True-Spatial-Mischung habe ich die Charaktere analog dazu platziert, was sehr gut funktionierte. Um auch hier die virtuelle Distanz und die akustische Charakteristik der Höhle zu simulieren, habe ich mir die Einstellungen des FdnReverbs, des MultiEQs und des DualDelays aus dem True-Spatial-Projekt kopiert.

### 10.2.3 Die Atmosphäre

Der Kern meiner virtuellen Höhle besteht auch in dieser Mischung aus vier Spuren. Über Halo habe ich die Aufnahme des durch Bäume streichenden Windes aus Szene 11 auf der oberen Hemisphäre gleichmäßig verteilt. Mit Einsatz eines Tiefpassfilters konnte ich den örtlichen Zusammenhang der beiden Szenen und die virtuellen Maße wie in True Spatial bewahren. Um vor allem den örtlichen Zusammenhang aufrechtzuerhalten, ist die Krähe aus Szene 11 auch unten in der Höhle kurzzeitig über uns zu hören. Da ich mich in dieser Mischung auf die obere Hemisphäre beschränken musste, habe ich die plätschernden Klänge der Höhlenaufnahme quadrofonisch bei 0° Elevation platziert. Das Ergebnis waren wahrnehmbare Löcher im akustischen Raum, woraufhin ich die Spur dupliziert und die Stereoaufnahme bei  $\pm 90^\circ$  Azimut platziert habe. Dies sorgte für eine deutliche Steigerung der Umhüllungswirkung. Der Wasserfall gehört auch hier zum festen Bestandteil unserer virtuellen Szenerie. Die Geräusche herabstürzenden Wassers sind wie in True Spatial mit einem Tiefpassfilter bearbeitet worden, um die Illusion der Höhle aufrechtzuerhalten.

### 10.2.4 Das Autowrack

Wie beim Abseilen musste ich auch hier die Perspektive ein wenig dem technischen Setup anpassen. In True Spatial ist Peter vorausgegangen, während wir noch am Seil hängen, und

hat sich vor uns Zugang zum Autowrack verschafft. Da ich wiederholt mit diversen Versuchen gescheitert bin, Peter unterhalb von uns wahrnehmbar werden zu lassen, habe ich mich für die Option entschieden, dass wir uns ebenfalls direkt hinter Peter in das Innere des Autowracks begeben. Dementsprechend liegen auch die Geräusche, die Peter verursacht, während er das Innere des Wracks nach Spuren durchsucht, bei 0° Elevation. Bei meinen Versuchen habe ich die in 9.1.2.3 erläuterten Tricks mit Hall, EQ und Pegeln angewendet.

Für die klangliche Gestaltung des Wageninnern habe ich drei der zuvor vier verwendeten Spuren genutzt; zwei davon habe ich wie in True Spatial mit -6 dB auf den LFE geroutet. Die auf der zuvor unteren Hemisphäre platzierte Atmo-Spur habe ich in 7.1.4 schlicht weggelassen. Um die Spannung zu steigern und leichte Bewegungen des Autowracks zu simulieren, habe ich ebenfalls zwei Aufnahmen von sich bewegendem Metall an verschiedenen Stellen der Szene auf der oberen Sphäre verteilt.

Das drohende Herabstürzen des Autowracks, in dem wir uns zusammen mit Peter befinden, wird mit den in 3.2.5 beschriebenen Klängen vernehmbar gemacht. Das vorausgehende metaphorisch akustische Abbrechen des vorspringenden Felsplateaus, auf dem sich das Autowrack befindet, ist in 7.1.4 ebenfalls hinten links bei etwa -159° Azimut wahrzunehmen, allerdings bei nun 0° Elevation. Das Knarzen des Seiles, mit dem vermittelt werden soll, wie Justus und Bob es auffangen, musste ich etwas höher platzieren, bei 89° Elevation, um es an derselben Stelle zu lokalisieren wie in der True-Spatial-Mischung.

## 11 Diskussion

Für meine Arbeit in dem von mir als True Spatial definierten System konnte ich keine vergleichbaren Arbeiten heranziehen, da es sie zum aktuellen Zeitpunkt noch nicht gibt. Daher gehe ich im Folgenden auf einige Punkte und Fragen ein, die mir im Verlauf der Produktion und speziell bei der Wahrnehmung in den Sinn gekommen sind. Die Wahrnehmung hat einen direkten Einfluss auf die kreative Gestaltung der Produktion. Eine klare Trennung von Wahrnehmung und Gestaltung ist nicht möglich. Deshalb habe ich beide unter einen Punkt zusammengefasst.

### 11.1 Interviews

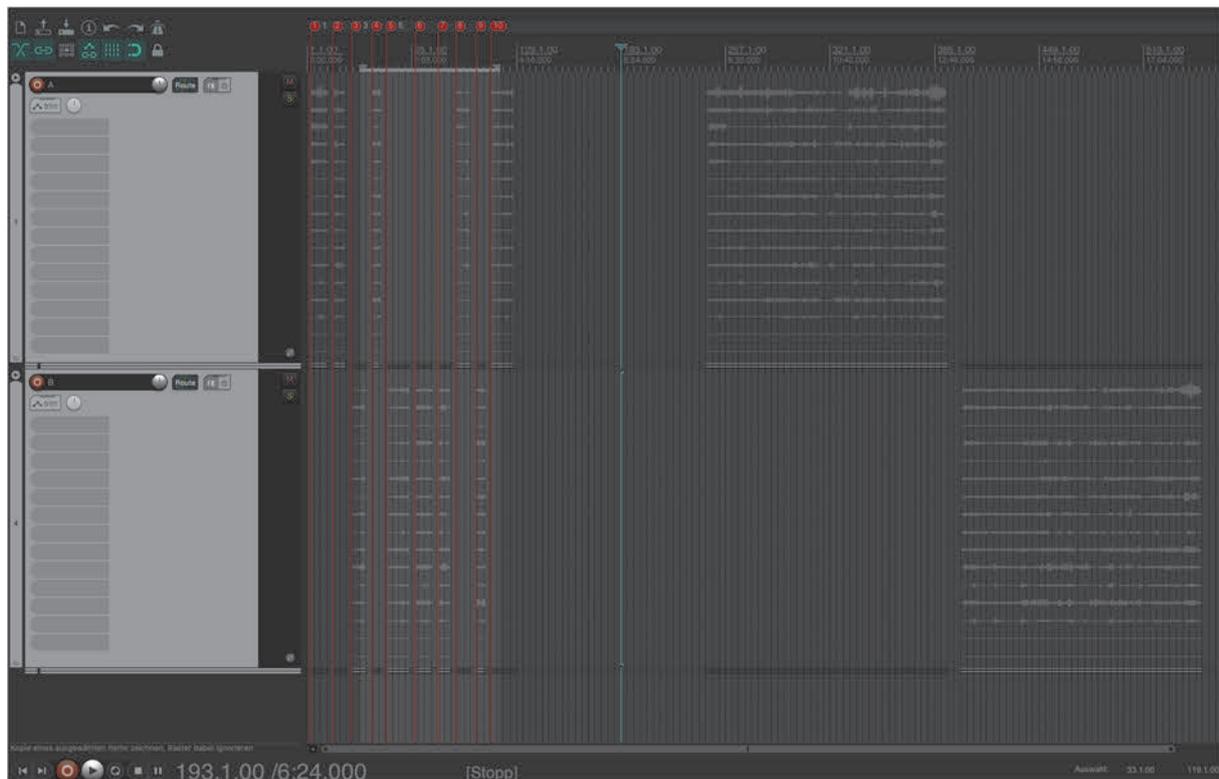
Um nicht ausschließlich mit meinen eigenen subjektiven Eindrücken zu argumentieren, habe ich vom 20. bis 22. Juni 2022 kurze Interviews mit vier audioaffinen und vier nicht-audioaffinen Proband:innen durchgeführt. Die insgesamt acht Rezipient:innen sind zwischen 22 und 50 Jahre alt, die vier audioaffinen Rezipient:innen sind Medientechnik Studierende, die vier nicht-affinen Freunde und Verwandte.

In Regie 3 des Tonlabors der HAW DMI habe ich ihnen zunächst zwei Mal fünf Samples aus bestimmten Szenen vorgeführt, um eine Diskussion über die Lokalisation, Erkennung und Relevanz einiger Klangobjekte zu ermöglichen. Damit sie einen möglichst unvoreingenommenen Eindruck gewinnen und sich nicht an eines der beiden Formate gewöhnen, habe ich bei der Reihenfolge der Samples nach dem Zufallsprinzip zwischen 7.1.4 und True Spatial variiert. Beim zweiten Durchgang habe ich das jeweils andere Format gewählt. Die beiden Durchgänge folgten unmittelbar hintereinander.

Um in den Pausen, während denen ich mich mit den Proband:innen unterhalten habe, auch keine visuelle Information aus dem Projekt zu liefern, habe ich die Samples jeweils sowohl in True Spatial als auch in 7.1.4 angelegt und das jeweils andere Paar stumm geschaltet. Auf diese Weise wollte ich verhindern, dass durch das Aufleuchten des Solo-Schalters auf eine der beiden Varianten geschlossen werden kann. Außerdem habe ich keinerlei Auskunft über den Hörversuch und die zu hörende Geschichte beziehungsweise die Szenerie gegeben.

Um die Proband:innen so wenig wie möglich zu beeinflussen, habe ich sie auch nicht direkt auf die mich interessierenden Themen angesprochen, sondern bin durch Fragen beispielsweise nach der klanglichen Gestaltung, der Narration und der Szenerie in ein Gespräch gekommen. Für die Ermittlung, ob und wie einzelne Klangobjekte, wie etwa der Wasserfall, wahrgenommen wurden, habe ich darum nach der Beschreibung der Szenerie gefragt oder im Falle des Erzählers nach der Narration.

Um den Wiedergaberaum und beeinträchtigende Einflüsse (vgl. Kapitel 4.10) auszublenden, habe ich die Rezipient:innen gebeten, während des Hörens die Augen zu schließen und sich gänzlich auf ihren akustischen Sinn zu konzentrieren. Dadurch sollte die Wechselwirkung der Wahrnehmungen auf ein Minimum reduziert werden. Während des Hörens war es ihnen gestattet, den Kopf zu bewegen, um sich im Sinne der in Kapitel 4.5 beschriebenen Drehtheorie zu orientieren.



**Abbildung 11.1:** Beispiel-Projekt der Interviews. Obere Gruppe (A): True Spatial; untere Gruppe (B): 7.1.4; linksseitig die zehn Samples, rechtsseitig die beiden Mischungen

Schließlich habe ich den Rezipient:innen die kompletten zwei Szenen sowohl in True Spatial (A) als auch in 7.1.4 (B) vorgespielt. Dabei habe ich sie gefragt, in welchem der beiden Versionen (A oder B) sie gerne das komplette Hörspiel hören würden. Bei der Wiedergabe habe ich darauf geachtet, dass zwei der beiden Gruppen zuerst die True-Spatial-Mischung und zwei zuerst die 7.1.4-Mischung hören. Den Rezipient:innen habe ich vorher nicht verraten, welche Version sie hören. Besonders wichtig war es, darauf zu achten, dass die beiden Mischungen gleich laut sind, um eine sachgemäße Einschätzung der Proband:innen zu ermöglichen. Sowohl die kurzen Samples als auch die beiden Szenen habe ich nur einmal vorgeführt. Damit wollte ich die reale Nutzungssituation eines mehrdimensionalen Hörspiels simulieren, bei der man sich üblicherweise einzelne Szene beziehungsweise das komplette Stück auch nur einmal anhört und dabei Objekte und Positionen umgehend erkennen soll.

Da die Interviews selbst nicht im Fokus meiner Forschungsarbeit standen und sie lediglich als Input für meine Diskussion dienen, habe ich in Absprache mit meinem Erstprüfer auf eine Transkription verzichtet.

## 11.2 Auffälligkeiten in Produktion und Wahrnehmung

### 11.2.1 Umgang mit der akustischen Perspektive

Wie in Kapitel 8.1 erläutert, musste ich vor Beginn des Sounddesigns die Frage klären, welche Rolle beziehungsweise welche Perspektive man als Hörende/r einnimmt. Diese Frage stellt sich bei jeder Hörspielproduktion. Der Raum ist bei mehrdimensionalen Audioformaten und im Fall von True Spatial um eine zusätzliche Ebene erweitert, die untere Hemisphäre. Das Skript gibt über die virtuelle Perspektive keinerlei Auskunft. Nur an manchen Stellen werden darin Orte oder lokale Gegebenheiten erwähnt, an denen wir uns als Rezipient:innen befinden. So heißt es in einer Anmerkung der Autoren zum Beispiel: „SFX: Schnitt zurück ins Auto“ (Skript, S. 5). Oder: „Wir hören den WASSERFALL durch das geschlossene Fenster“ (Skript, S. 39). Doch diese Anmerkungen scheinen sich eher an filmischen Regieanweisungen zu orientieren und geben für das Sounddesign keine klare Richtung vor, welche akustische Perspektive geschaffen werden soll. Das gilt erst recht für ein vollsphärisches Wiedergabesetup wie True Spatial. Insofern gibt es zig Möglichkeiten, Mischung und Sounddesign mit entsprechenden Perspektiven zu gestalten. In einem Video, in dem die beiden Autoren John Beckmann und Ivar Leon Menger über die Entstehung des Drehbuchs sprechen, erwähnen sie mit keinem Wort die Perspektive (Die drei ???, 2021).

Nachdem ich mit der Mischung fertig war, habe ich den Toningenieur Pierre Brand, der die drei Fragezeichen in „3D Audio“ (Die drei ???, o.D. *3D Audio*) mischt, in einem Telefonat am 25. Mai 2022 befragt und wollte wissen, wie er mit dem Problem der Perspektive umgeht. Sein Ansatz ähnelt jenem, den ich gewählt habe: Er versucht die Hörenden als stille Beobachter:innen in das Geschehen einzubeziehen. Ob sie dabei als diegetische oder nichtdiegetische Figur partizipieren, hängt von der Geschichte und der Szene ab. Es wäre sicherlich interessant, sich an Schnitten aus Filmen zu orientieren und diese in einem solchen mehrkanaligen Audiosystem akustisch zu übersetzen oder einen völlig anderen Ansatz als den meinen zu wählen.

Während meiner Sounddesign-Arbeit an dem Fall der drei Fragezeichen fiel mir auf, dass sich die Autoren keine Gedanken über die Frage der akustischen Perspektive gemacht haben. Pierre Brand bestätigte mir in dem Telefonat, dass dies generell so sei. Am Ende müsse man eine Lösung finden, die nicht nur einem selbst, sondern auch den Autoren gefällt. Es scheint ein ähnliches Problem zu sein wie in den 1970er Jahren bei den Kunstkopf-

Hörspielen (vgl. Kapitel 2.11). Der Regisseur Ulrich Gerhardt hatte damals geäußert: Es gibt dafür keine Autoren. Jedenfalls keine, die sich mit dem Wiedergabemedium und allen damit verbundenen Fragen und Herausforderungen beschäftigen. Meines Erachtens gibt es zwei Lösungen für dieses Problem: Entweder die Autoren befassen sich mehr mit dem Wiedergabesystem und dessen Möglichkeiten oder die beiden Gewerke (Hörspielautor:innen und klangliche Gestalter:innen) arbeiten enger und vor allem früher zusammen. So könnte man zum Beispiel das Skript während des Schreibprozesses immer wieder von dem/der zuständigen Sounddesigner:in gegenlesen lassen und wechselseitig zu klären versuchen, wie Perspektive und räumliche Gegebenheiten klanglich gestaltet werden könnten.

Während meiner praktischen Arbeit habe ich erfahren, dass die originale Planetariumsvariante des Hörspiels mit Bewegtbildern unterlegt ist, die auf die Kuppel der Planetarien projiziert werden. Dabei handelt es sich nicht um einen klassischen Film, sondern um leicht dynamisierte Bilder, die als „illustrierte immersive Untermalung“ (Die drei ???, 12. November 2021, 10:02–10:05) dienen und ein „Gefühl für die Szene“ (ebd., 11:29–11:31) vermitteln sollen. Auf diese Weise wird die virtuelle Perspektive, die ich mit meinem Sounddesign rein akustisch zu vermitteln versucht habe, auch durch die visuelle Wahrnehmung bestimmt. Dank dieser Wechselwirkung der akustischen und optischen Wahrnehmung müssen die Autoren sich nicht so stark auf eine bestimmte Perspektive festlegen. In meinem Fall fehlt aber das Element der visuellen Wahrnehmung. Für rein akustische mehrdimensionale Hörspiele müsste sich, wie vorhin erläutert, an der Arbeit der Autoren beziehungsweise an der Zusammenarbeit etwas ändern.

### 11.2.2 Umgang mit nichtdiegetischen Elementen

Im Hinblick auf die Immersion und die Möglichkeiten eines dreidimensionalen Wiedergabesystems stellt sich die Frage, wie man mit nichtdiegetischen Klangobjekten umgehen soll.

Ist zum Beispiel noch ein Erzähler notwendig, der Zusammenhänge, Orte, Blenden, Zeit- oder Ortssprünge erklärt? Ich habe nur zwei der insgesamt zwanzig Szenen designt und gemischt, in denen es auch nur zu Beginn einen kurzen Erzähler-Anteil gibt. Für eine allgemeingültigere Aussage wären mehr Zeit und Arbeit erforderlich. Nach meiner Ansicht reit einen die nichtdiegetische Figur des Erzählers und vor allem die Art und Weise, wie ich sie angelegt habe, aus der Immersion heraus. Der Ursprung der von unten kommenden Stimme lässt sich kaum lokalisieren. Beim Hören versucht man gleichwohl herauszufinden, von woher die Stimme des Erzählers nun kommt. Indem man aber den Fokus auf den Erzähler richtet und in der physischen Welt seinen Standort zu lokalisieren versucht, verliert man den Bezug zur virtuellen Welt und wendet seine Sinne der physischen Welt zu.

Bei größeren Zeit- oder Ortsprüngen hätte ein/eine Erzähler:in wahrscheinlich eine andere Bedeutung. Wenn die optische Wahrnehmung fehlt, wird es deutlich schwieriger, vor allem Zeitsprünge markant darzustellen. Da ein solcher Fall in meiner Arbeit aber nicht vorkam, kann ich auch hier nur mutmaßen. Die Videokuppel in den Planetarien, wo das Original gezeigt wird, ist dabei gewiss von Vorteil. Für dieses Problem könnte ich mir zwei Lösungsansätze vorstellen: entweder einen angepassten Dialog, in dem die Figuren beiläufig erläutern, zu welcher Zeit sie sich wo befinden, oder die Besetzung der Rolle eines Erzählenden. Ungeachtet dessen wäre mein Präferenz für ein längeres Hörspiel in diesem Format, auf die nichtdiegetische Figur des Erzählers zu verzichten. Denn durch den stetig mit dem Ertönen der Erzählerstimme wiederhergestellten Bezug zur physischen Welt würde meines Erachtens die Immersion leiden. Die Welt, mit der die Rezipient:innen gedanklich verschmelzen, nachdem sie vergessen haben, dass sie sich in einem Studio befinden (vgl. Holman, Kapitel 4.9), und von dem Gefühl durchdrungen sind, selbst Teil der Geschichte zu sein (vgl. Fritz Peter Buch, Kapitel 2.4), würde auf diese Weise kollabieren. Die für die Immersion notwendige Verschiebung der Aufmerksamkeit und das Abkoppeln der Wahrnehmung von der physischen Welt (vgl. Agrawal et al., Kapitel 5.1.2) würden so beeinträchtigt.

Dieselbe Frage stellt sich bei der Musik, die in den zwei von mir bearbeiteten Szenen nichtdiegetisch und somit nicht Teil der erzählten Welt ist. Leider konnte mir Sony nicht die Stems der Musikstücke zukommen lassen, so dass ich ein wenig eingeschränkt war bei der Verteilung der Musik und dem Umgang mit ihr. Zuvor hatte ich vergeblich versucht, die Stems mit dem Plugin Music Rebalance der Firma iZotope zu generieren. Dieses Plugin scheint aber für die Verwendung im Pop- und Rockbereich entwickelt worden zu sein, da sie es lediglich ermöglicht, Stimmen (Vocals), Bass, perkussive Elemente (Percussion) und andere (others) neu zu balancieren. Die von mir verwendete Musik wurde aber vor allem von Streichinstrumenten erzeugt, so dass es mir nicht möglich war, Stems zu extrahieren.

Nach meinem Empfinden gilt für die nichtdiegetische Gestalt der Musik dasselbe wie für den Erzähler hinsichtlich der Immersion: Wenn wir mit der neu geschaffenen akustischen Realität verschmelzen, fühlen wir uns in dieser derart präsent, dass sie von nichtdiegetischen Elementen – sei es ein Erzähler oder Musik als nichtdiegetisches Mittel – zerstört wird. Die nichtdiegetische Musik hilft aber wiederum beim Spannungsaufbau. So sind bei der Entdeckung des Autowracks in der Schlucht tieffrequente Bassflächen zu hören, welche die Bedrohlichkeit und Anspannung der Situation untermalen und steigern sollen. Der gewünschte Effekt ließe sich allerdings auch zum Beispiel durch Klänge der Natur erzielen. Der Wind könnte beispielsweise auffrischen, das Blattwerk mehr und mehr zu rascheln beginnen. Eine solche Lösung könnte der Immersion förderlich sein, solange einen solche diegetischen Klänge, überspitzt oder hyperreal dargestellt, nicht aus der akustisch virtuellen Welt abziehen. Mit den entsprechenden Stems ließen sich bei der Verteilung der Musik neue Wege

erproben und ermitteln, ob die Immersion oder auch die Dramaturgie auf diese Weise verbessert werden könnte.

Bei den Interviews stellte sich allerdings heraus, dass keinen der acht Proband:innen die Musik oder der Erzähler gestört hat. Fünf von ihnen gefiel die Musik sogar. Ihnen hatte ich als Sample den Übergang von Szene 11 zu 12 vorgeführt. Eine der Befragten sagte, dass die Musik gut zur Szenerie passe. Ein weiterer äußerte, dass die Musik umhüllend und atmosphärisch klinge. Wiederum eine andere sagte, dass die Musik deshalb gut passe, weil sich der Rest um sie gedreht und die Musik insofern als orientierender Anker gedient habe. Drei der fünf waren audioaffin, die beiden anderen nicht-audioaffin.

Um eine allgemeingültigere Aussage über die Verwendung nichtdiegetischer Elemente in mehrdimensionalen Audioformaten treffen zu können, bedarf es sicherlich weiterer Forschung. Immerhin lässt sich festhalten, dass ein Großteil der Rezipient:innen weder die Musik noch den Erzähler als störend empfand oder geäußert haben, dass diese sie aus der erzählten Welt herausgerissen hätten. Mehr als der Hälfte gefiel der Einsatz solcher nichtdiegetischen Elemente sogar, weil sie dadurch ein besseres Hörerlebnis hatten.

### 11.2.3 Neutrale Mischung/Sounddesign

Nachdem ich mir das Skript wieder und wieder durchgelesen und mit der klanglichen Gestaltung begonnen hatte, formte sich ein Bild der Szenerie in meinem Kopf. Dass wir im ersten Teil der elften Szene in einem Auto sitzen, danach aussteigen und uns in einem Waldstück am Rande einer Höhle befinden, in die wir in Szene 12 hinabsteigen, stand für mich ab einem gewissen Zeitpunkt nicht mehr in Frage. Nun muss aber auch jemand, der das Stück zum ersten Mal hört, ohne das Skript gelesen zu haben, sofort verstehen, wo wir uns befinden, und ebenso Klangobjekte unmittelbar identifizieren können.

Vor allem bei Klangobjekten, die ein virtuelles Objekt darstellen sollen, wurde mir im Nachhinein bewusst, dass ich mit viel mehr akustischen Metaphern hätte arbeiten müssen. Speziell der Wasserfall war eine echte Herausforderung. Da ich das Skript gelesen hatte, war in meiner Wahrnehmung klar, dass das zu hörende Rauschen, das hin und wieder an- oder abschwilt, eben jener Wasserfall ist. Für jemanden, der das Stück zum ersten Mal hört, war dies allerdings ganz und gar nicht klar. Ich hatte im Verlauf meiner klanglichen Gestaltung einigen Personen meine Mischung vorgeführt, um Feedback zu bekommen<sup>4</sup>. Keiner dieser Personen erkannte den Wasserfall als solchen, weder während der Autofahrt noch

<sup>4</sup> Fünf dieser Personen waren Studierende und Lehrende des Studiengangs Medientechnik. Hinzu kamen rund zehn Music-Technology-Student:innen und ein Dozent der Capital University Ohio. Da diese kein Deutsch verstehen, sind ihre Meinungsäußerungen zu bestimmten Klangobjekten mit Vorsicht zu betrachten, da der deutsche Erzähler und die Charaktere selbst einiges vorwegnehmen.

außerhalb des Wagens. Dem wollte ich in den Interviews auf den Grund kommen. Dafür habe ich zunächst die Stelle aus dem Hörspiel vorgespielt, in dem der Wasserfall zuerst auftaucht: Szene 11 bei Minute 1:28–1:38.

Sieben der acht Rezipient:innen identifizierten meinen Wasserfall als ein oder mehrere vorbeifahrende Autos. Eine weitere Rezipientin dachte, dass es sich um einen Zug handeln würde. Im weiteren Verlauf des Hörspiels, als wir uns im zweiten Teil der elften Szene in der Natur befinden, erkannten immerhin drei Rezipient:innen (zwei Laien und ein Experte), dass es sich um einen Wasserfall handeln soll. Eine weitere der audioaffinen Rezipient:innen hielt das Rauschen draußen für das eines Bachs. Die restlichen vier blieben bei der Version eines oder mehrerer Autos auf einer befahrenen Straße.

Wegen der Begrenzung auf den akustischen Sinn scheint die Verwendung metaphorischer Klänge um ein Vielfaches wichtiger zu sein. Zunächst hatte ich mich an dem Zitat von Tobias Falke aus Kapitel 4.6 orientiert und versucht, ein möglichst realitätsnahes Schallfeld zu erzeugen. Wie aber in Kapitel 9.1.1.3 erläutert, führte das keineswegs zum gewünschten Erfolg. Falke ging es in seiner Arbeit allerdings nicht um die Wahrnehmung, sondern um die rein technische Reproduktion. Zwar hatte ich den Wasserfall daraufhin mit Aufnahmen von plätschernden Flüssen und kleinen Bächen unterlegt, um metaphorisch einen Wasserfall zu evozieren; dabei schien ich aber zu zaghaft gewesen zu sein. Für einige Klangobjekte scheint auch eher die Devise *weniger ist mehr* zu gelten. Mein Wasserfall besteht aus sechs Spuren. Davon sind drei Aufnahmen solche von echten Wasserfällen. In der Summe scheint sich ein verrauschter Klang zu ergeben, jedenfalls wenn als Korrektiv die optische Wahrnehmung fehlt. Dies bestätigte mir auch Pierre Brand. Bei einem Wasserfall, sagte er, solle man sich auf einen Kernklang beschränken, der dann mit metaphorischen Klängen zu unterlegen sei. Ein sehr großdimensioniert klingender Wasserfall scheint also die falsche Wahl gewesen zu sein. Besser hätte ich einen kleinen, eher plätschernden Wasserfall akustisch designen sollen, bei dem sofort das Bild eines Wasserfalls in den Köpfen der Rezipient:innen entstanden wäre.

Wie bereits erwähnt ist es ebenso wichtig, sich während der Mischung allein auf den akustischen Sinn zu fokussieren. Laut Scherzer (2010), den ich in Kapitel 4.10 zitiert habe, spielt der physische Raum eine große Rolle. Die ständige visuelle Präsenz der mich umgebenden Lautsprecher hat mich dazu verleitet, diesen Lautsprechern Klangobjekte zuzuordnen. Das Ergebnis waren nicht nur meinem subjektiven Eindruck nach akustische Lücken zwischen den Lautsprechern, es war auch kontraproduktiv für die neutrale Mischung, die jemandem, der das Stück zum ersten Mal hört, die Erfahrung vermitteln soll, die ich mit meinem Sounddesign vermitteln wollte. Um dies zu ermöglichen, muss man während des Hörens, wie Scherzer schon sagte, den Wiedergaberaum für die Wahrnehmung ausblenden.

Das habe ich durch das Verschließen der Augen zu lösen versucht, was sich aber als nicht optimal erwiesen hat.

Wie in 11.2.1 erwähnt, ist das Originalstück mit Bildern unterlegt. Auch hier wurde im Sinne Scherzers gehandelt und der physische Raum mitinszeniert. Die Bilder und die sich mit ihnen ergebende Wechselwirkung des visuellen und des akustischen Sinnes machen es meines Erachtens deutlich einfacher, die Szenerie und bestimmte Objekte in der Weise zu vermitteln, wie man sich das für die akustische Gestaltung vorgenommen hat. Wahrscheinlich würde das Problem der falschen Identifizierung von Klangobjekten, wie in meinem Fall des Wasserfalls, nicht auftreten, wenn man die Töne mit Bildern untermalt.

Schließlich musste ich auch noch feststellen, dass Pausen besonders wichtig sind. Bei der langen Arbeit unter Dauerbeschallung von oft über 90 dB ermüdet das Gehör sehr schnell, was zu dem von Blauert (1974) in Kapitel 4.4 beschriebenen Phänomen führt, dass sich Hörereignisse näher an die Medianebene verschieben. Eben dies dürfte erklären, warum ich einige Klangobjekte, wie zum Beispiel den Blinker oder das Ziehen der Handbremse, so nah an die Medianebene gelegt habe.

#### 11.2.4 Die untere Hemisphäre in True Spatial

Was das von mir als True Spatial definierte Setup ausmacht, ist die Möglichkeit, vollsphärisch auch mit unteren Lautsprechern zu arbeiten. Gleich zu Anfang stellte sich mir die Frage, was für Audiomaterial sich für den unteren Teil der Hemisphäre in Hörspielen überhaupt nutzen ließ und auf welche Weise – und würde diese Erweiterung überhaupt etwas bringen?

Wie in 8.2 erläutert, habe ich herausgefunden, dass sich die untere Hemisphäre hervorragend für die Platzierung von nichtdiegetischen Elementen eignet, insbesondere von Sprache. Das bedeutet allerdings nicht, dass Sprache quasi automatisch nichtdiegetisch wirkt, sobald sie auf der unteren Hemisphäre liegt. Dieser Effekt trat bei mir nur auf, wenn ich die Stimme des Erzählers direkt unter mich gelegt habe und sie dort statisch blieb. Bei Morton im zweiten Teil der elften Szene, wo dieser unter dem Wagen liegt und sich anschließend aufrichtet, war das nicht der Fall. Weil der Ursprung des Klangs im Falle des Erzählers undefinierbar ist, kommt es subjektiv zu einer Abkopplung der Klangobjekte von der erzählten Welt. Dass dies nicht unbedingt förderlich für die Immersion ist, habe ich in 11.1.2 dargelegt.

Das wirft natürlich die Frage auf, ob die untere Hemisphäre für ein Hörspiel überhaupt notwendig ist, das immersiv wirken soll. Notwendig ist die untere Hemisphäre für den Effekt der Immersion meines Erachtens nicht. Nach der Definition von Agrawal et al. (2019) aus Kapitel 5.1.2, an der ich mich orientiert habe, bräuchte es keinerlei sensorische Stimulation, um das Gefühl der Immersion zu erzeugen, also mental mit der erzählten Welt zu verschmelzen und dabei die Wahrnehmung der physischen Welt abzukoppeln, wie in dem Beispiel von

Tagträumen. Folglich könnte auch ein Hörspiel in Mono immersiv wirken. Was True Spatial aber bietet, ähnlich wie der Übergang von Mono zu Stereo (vgl. Kapitel 2.10) oder die Einführung von 5.1 (vgl. Kapitel 3.8), ist ein vergrößerter Spielraum und eine neue künstliche Raumakustik. Diese Akustik lässt sich für die wirklichkeitsgetreue Gestaltung und die Narration nutzen und bietet größeren künstlerischen Freiraum. Bewegungsabläufe wie die unseres Chauffeurs Morton im zweiten Teil der elften Szene sind beispielsweise nur in True Spatial darstellbar. Hier liegt er akustisch tatsächlich unter dem Auto, was auch subjektiv so empfunden wird, und richtet sich anschließend virtuell auf. Sechs der Rezipient:innen (drei audio-affin, drei nicht-affin) erkannten diesen Bewegungsablauf dann auch in True Spatial.

Der Handlungsraum für solche Bewegungen ist zum Beispiel in 7.1.4 von deutlich kleineren Dimensionen. Ein virtuelles Aufrichten Mortons ist hier kaum realisierbar, da die tiefstmögliche Ebene auf Ohrhöhe liegt. Da wir virtuell neben Morton stehen, befindet er sich somit schon auf unserer Höhe. Ein weiteres Hinaufschwenken würde die von mir gewählte Perspektive zunichtemachen. Man könnte die Perspektive neu wählen und uns virtuell schrumpfen lassen, um Freiraum zu gewinnen, doch dies engt wiederum den Spielraum in den Höhen ein. Dass sich dadurch das Hörerlebnis steigern ließe, bezweifle ich stark.

Neben solchen Bewegungsabläufen lässt sich die untere Hemisphäre sehr gut für SFX nutzen, die subjektiv – wie die Untersuchungen von Tsingos et al. (2011) aus Kapitel 3.9 für die Erweiterung durch Höhenkanäle zeigten – ein besseres Gefühl für den akustischen virtuellen Raum vermitteln. Die Erweiterung auf eine vollsphärische Darstellung potenziert dieses Gefühl meines Erachtens noch etwas mehr. Sämtliche Schritte liegen dabei im unteren Teil der Hemisphäre. Dadurch setzen sich die Schritte virtuell von der Sprache ab und im Kopf entsteht ein subjektiv wirklichkeitsgetreueres Bild. Das Gleiche gilt für den ersten Teil der zwölften Szene, wo wir uns mit Peter zum Autowrack hinabseilen. Zu hören ist, wie ständig Steine an uns vorbei hinunterfallen. In 7.1.4 stoppt deren Fall nach subjektivem Eindruck auf Ohrhöhe, was die Illusion des virtuellen Raums meines Erachtens ruiniert. Ich war zuvor aber auch die True-Spatial-Mischung gewohnt, bei der die Steine eben tiefer fallen. Beim Wechsel zu 7.1.4 habe ich die für mich zu jenem Zeitpunkt gewohnte untere Ebene vermisst.

Da es sich bei meinen Effekten vor allem um breitbandige impulshafte Klänge handelt, wäre in weiteren Untersuchungen zu klären, inwieweit die spektrale Beschaffenheit und die Dauer der Klänge Auswirkung auf die Wahrnehmung von unten haben. Außerdem müsste im Fall von Sprache ermittelt werden, bis zu welchem Grad die Lokalisation auf der unteren Hemisphäre verlässlich funktioniert. Bei Morton, der im zweiten Teil der elften Szene in True Spatial bei  $-70^\circ$  Elevation unter dem Auto liegt, erkannten sechs der acht Rezipient:innen (drei Audio-Experten & drei Audio-Laien) dessen virtuelles Aufrichten. Bei der Verortung des Erzählers, der bei  $-90^\circ$  Elevation liegt, gelang dies keinem von ihnen (siehe Kapitel 11.2.6).

Auch unauffälligeren Klängen wie den Atmosphären kommt die untere Ebene zugute. Indem man nun auch Atmo unter sich hat, steigert sich das subjektive Gefühl der Umhüllung, und wie bei den SFX wird so auch das Gefühl für den virtuellen Raum zusätzlich gesteigert. Wichtig ist allerdings die Wahl der Klänge für die untere Hemisphäre. Mein Ansatz war der von Field (2002) für die Surround-Kanäle aus Kapitel 4.9, bei dem man dekorrelierte und vor allem verschiedene, aber gleich klingende Aufnahmen verwendet. Nach den Erfahrungen, die ich bei meiner Arbeit gesammelt habe, würde ich auch in diesem Fall sagen, dass weniger mehr ist. Auf die untere Hemisphäre habe ich in der Hauptmischung keine sehr dynamische oder aktionsreiche Atmos gelegt. Bei vorausgehenden Versuchen mit sehr dynamischen Klangflächen oder Effekten, wie etwa längerem Geraschel von Büschen, kam es zu Sprüngen in der Wahrnehmung und zu falscher Verortung. Es reichte vollkommen aus, ein bis zwei dezente Aufnahmen leicht hinzuzumischen, um die Illusion einer vollständigen Umhüllung zu erwecken. Vor allem in der zwölften, der Höhlen-Szene kommt das besonders gut zum Ausdruck. Hier hat man subjektiv das Gefühl, sich tatsächlich in einer Höhle zu befinden und von dieser umhüllt zu sein.

Grund dafür könnte das von Thomas Görne in Kapitel 4.9 beschriebene Phänomen sein, dass wir die Umhüllung erst dann bewusst wahrnehmen, wenn sie außergewöhnlich klingt. Und genau das tut sie in dieser Szene. Das entspricht auch den Erkenntnissen von Hamasaki et al. (2004) aus Kapitel 3.10 über die Verwendung von Höhenkanälen in Surround-Konfigurationen. Ähnlich wie bei den Höhenkanälen kommt es auch hier auf die verwendeten Klangszenen an. Die untere Hemisphäre ist mithin nicht zwingend notwendig für die Erzeugung eines Gefühls von Immersion, sie erleichtert aber den Übergang in die virtuelle Welt und steigert das Gefühl der Umhüllung und der virtuellen Räumlichkeit. Die Rolle der unteren Hemisphäre, vor allem in Bezug auf die Umhüllung, wird einem erst bewusst, wenn man von der True-Spatial-Mischung auf die 7.1.4 wechselt.

Meine Meinung bestätigte sich in den Interviews vom 20. bis 22. Juni: Mit einer Ausnahme präferierten alle Rezipient:innen die Mischung in True Spatial. Als Gründe wurden ein besseres und natürlicheres Hörerlebnis, größerer Realismus oder ein intensiveres Gefühl der Umhüllung genannt. Bei der 7.1.4-Variante sorgten beispielsweise die Schritte in drei Fällen (zwei Audio-Laien & ein Audio-Experte) für Verwirrung, da sie sich auf derselben Höhe befanden wie die Stimmen. Dass sich das Aufstehen Mortons unter dem Wagen nach dem Ausstieg nicht vernehmen ließ, sorgte in 7.1.4 ebenfalls bei drei weiteren Rezipient:innen für Verwirrung. Aber nicht nur die Umhüllungswirkung wurde durch Erweiterung um die untere Hemisphäre gesteigert; nach Anhören der True-Spatial-Mischung sagten vier Rezipient:innen, dass sie sich gefühlt hätten, als seien sie vor Ort gewesen. Zwei gaben sogar an, während der letzten, der zwölften Szene in True Spatial eine Gänsehaut bekommen zu haben (ein Audio-Laie & ein Experte). Als Begründung wurden die intensiver wirkende Umhüllung und

eine realer klingende Welt genannt. Eine Rezipientin (Audio-Laiin) gab nach Anhören der 7.1.4-Mischung an, dass sie sich fühle, als befände sie sich unterhalb der erzählten Welt. Bei der einen Ausnahme, die keine der beiden Mischungen mochte und nach eigener Aussage lieber in Stereo hören würde, sorgte vor allem die Perspektive für Verwirrung. Das Autoinnere und etliche andere Klangobjekte wurden von diesem Rezipienten nicht als solche erkannt, der Ausstieg mit dem wandernden Wasserfall ergab für ihn ebenso wenig Sinn wie die Blende in Szene 12. Alles würde nur rauschen und unnatürlich klingen. Bei dem nicht audioaffinen Probanden sorgte das nach eigener Aussage für grundsätzliche Verwirrung.

Die Meinungen sind bei den Audio-Experten und den Laien recht ähnlich, lediglich in ihren Begründungen unterscheiden sich die beiden Gruppen. Hier kommt die Erfahrung und das Wissen um das Audio-Medium den vier audioaffinen Rezipient:innen zugute. Sie artikulieren Sätze wie: „Version A gibt mir ein besseres Raumgefühl mit einer deutlich besseren Lokalisation, was das Ganze viel immersiver macht.“ Während die nicht audioaffinen Rezipient:innen oft angeben, ein besseres Hörerlebnis gehabt zu haben, das „mehr Spaß“ macht.

Das Hörerlebnis scheint bei True Spatial demnach definitiv besser zu sein. Die Gründe dafür sind ein höherer Grad an Realismus, natürlichere Bewegungsabläufe und mehr Umhüllungswirkung. Wie erwähnt ist die Erweiterung um die untere Hemisphäre nicht notwendig für die Erzeugung eines immersiven Gefühls. Ungeachtet dessen scheint sich meine These, dass sich der Übergang in die virtuelle erzählte Welt und die Loslösung vom physischen Raum auf diesem Wege schneller erreichen lassen, in den Interviews bestätigt zu haben.

#### 11.2.5 Arbeit in True Spatial vs. 7.1.4

Hätte ich nicht mit True Spatial, sondern mit 7.1.4 angefangen, wäre wahrscheinlich ein völlig anderes Sounddesign entstanden. Da ich dasselbe Sounddesign für beide Formate verwendet habe, um besser vergleichen zu können, blieb mir in der zweiten, der 7.1.4-Mischung, bei der Gestaltung deutlich weniger kreativer Freiraum. Einige Entscheidungen waren für die 7.1.4-Mischung also schon getroffen. Es blieb mir lediglich, sie so gut es ging in dem neuen Setup zu adaptieren. Mit 7.1.4 als Auftakt wäre die Schlucht eventuell eine Schlucht geblieben. Angesichts der Möglichkeiten einer vollsphärischen Gestaltung gab mir eine Schlucht akustisch aber zu wenig her, weshalb ich mich für eine Höhle entschieden habe, die im krassen akustischen Gegensatz zu der vorherigen Szene steht. Deshalb sind die Ausführungen in Kapitel 10 auch so viel knapper gehalten als in Kapitel 9.

Meine Arbeit in 7.1.4 unterschied sich deshalb kaum von jener in True Spatial. Dank des ReaSurroundPan entfielen lediglich psychoakustische Tricks, die ansonsten genutzt werden, um Objekte in die Sphäre zu holen, da hier die Platzierung von Klangobjekten im Innern der Sphäre auf dem Interface vollzogen wird. Die Möglichkeit, Objekte einfacher in die Sphäre

und näher an die Zuhörer:innen zu bringen, nutzte ich wegen der Orientierung an der True-Spatial-Mischung kaum aus.

Außerdem bin ich den Umweg über Ambisonics und virtuelle Lautsprecher gegangen und habe in der 7.1.4-Mischung dieselben Plugins verwendet, bis auf den ReaSurroundPan. Auch das ist ein Grund, warum sich die Arbeit in 7.1.4 kaum von der in True Spatial unterschieden hat. Ich konnte die meisten Einstellungen der Plugins direkt übernehmen und musste die Parameter nur an wenigen Stellen anpassen.

### 11.2.6 Lokalisation

Auffällig bei der Wiedergabe ist die Lokalisation einiger Klangobjekte gewesen. In extremer Form gilt dies, wie in 8.2 und 11.1.2 erwähnt, für den Erzähler in True Spatial. Das von mir beschriebene Phänomen, dass sich dessen Ursprung nicht auf einen Punkt festlegen ließ, wodurch ein Meer aus Sprache entstand, spiegelte sich in den Interviews wider.

Ohne Ausnahme verorteten alle den Erzähler über sich. Drei Rezipient:innen gaben an, die Lokalisation sei kaum möglich, der Sprecher sei aber „irgendwo oben“. Die anderen erklärten, dass sich der Erzähler über ihnen befinde, in der Tendenz etwas links. Ein audio-affiner Proband kam der Lösung am nächsten. Nachdem er zunächst geäußert hatte, dass sich der Erzähler über ihm befände, korrigierte er sich kurz darauf und sagte nun, dass der Erzähler aus seiner Körpermitte käme. Doch wollte er sich nicht festlegen. Nach den Interviews habe ich die Rezipient:innen aufgeklärt, worin der Unterschied zwischen den Samples und den zwei Mischungen bestanden hat. Als ich ihnen dann eröffnet hatte, dass sich der Erzähler in True Spatial eigentlich direkt unter ihnen befinde, wollten sechs (vier Audio-Experten, zwei Laien) das Sample noch einmal hören. Beim erneuten Anhören gaben sie an, das nun auch gehört zu haben, da sie ja nun wüssten, dass sich die Stimme unter ihnen befindet. Das entspricht den Erkenntnissen aus Kapitel 4.6 zur Lokalisation auf der vertikalen Ebene, nach denen die Lokalisation bis zu einem gewissen Anteil über Lernprozesse und Erfahrungen erfolgt.

In 7.1.4 wurde der Erzähler von allen acht Proband:innen direkt über ihren Köpfen lokalisiert. Fünf (drei Audio-Experten und zwei Audio-Laien) gaben an, dass er sich direkter und näher anhöre, was meiner Platzierung aus Kapitel 8.2 nach schlüssig ist.

Um die Frage der Lokalisationsschärfe auf der unteren Hemisphäre zu klären, bedarf es folglich weiterer Untersuchungen.

Eine weitere Auffälligkeit bei der Lokalisation ist mit Peter verbunden. Während der Autofahrt im ersten Teil der elften Szene befindet er sich virtuell zu unserer Linken. Nach meiner subjektiven Wahrnehmung, und dies auch schon während der Produktion, springt Peter in True Spatial unregelmäßig zwischen dem oberen und unteren Lautsprecher. Keiner der Rezi-

ipient:innen sprach von einem ähnlichen Eindruck, jedoch ist auffällig, dass drei Peter links oben (ein Laien/zwei Experten), vier aber links unten (zwei Laien/zwei Experten) lokalisierten. Lediglich ein Rezipient empfand, dass Peter sich auf seiner Ohrhöhe befinden würde.

Zunächst dachte ich, das würde an der Aufteilung der seitlichen Lautsprecher liegen, da zwischen beiden eine relativ große Lücke genau auf Ohrhöhe besteht. Mir kam auch in den Sinn, was Falke et al. in ihren Arbeiten mit HOA von 2014 herausgefunden hatten, nämlich dass mehr Lautsprecher auf der Vertikalen nötig sind, um ein solches Springen zu vermeiden (vgl. Kapitel 4.6). Da die Aufstellung auf der rechten Seite aber identisch ist, konnte es daran nicht liegen.

Mein zweiter Erklärungsansatz war, dass man sich eher nach links orientiert, da das Bodengitter, unter dem sich ein weiterer Lautsprecher befindet, nicht symmetrisch im Raum verbaut, sondern etwas weiter links im Raum platziert ist. Bei den Interviews habe ich aber darauf geachtet, dass der Stuhl stets an dem Messpunkt der Kalibrierung steht. Dieser Messpunkt, am Sweetspot der Lautsprecherkonfiguration gelegen, ist besonders schmal. Martin (2002) hatte darauf hingewiesen (vgl. Kapitel 4.6), dass der Sweetspot durch mehr Lautsprecher schmaler wird. Besonders wenn sie sich so nah an den Hörenden befinden, wie das in Regie 3 der Fall ist. Schon mit kleinen Kopfbewegungen verlässt man hier schnell den Sweetspot.

Außerdem ist es wichtig, die Ambisonics-Ordnung zu betrachten. Ich habe in dritter Ordnung mit lediglich 14 Lautsprechern gearbeitet. Durch die Aufstockung der Anzahl an Lautsprechern würde sich in höherer Ordnung arbeiten lassen. Die Folge wären mehr sphärische Harmonische, die Richtungsinformationen übertragen, was nach Erkenntnissen etwa von Boren (2017) für eine größere räumliche Präzision sorgt (vgl. Kapitel 3.12). Aber auch das erklärt nicht, warum nur Peter zu springen scheint beziehungsweise oben oder unten lokalisiert wird. Ein weiterer Versuch der Erklärung wäre, dass das Problem in der spektralen Zusammensetzung seiner Stimme liegt. Peters Stimme ist im Vergleich zu der von Morton, Justus oder Bob eher hoch und etwas quietschig. Auch wenn Blauert seine Erkenntnisse zu den richtungsbestimmenden Bändern später etwas relativierte, wäre dies doch eine mögliche Erklärung für dieses Phänomen. Für eine fundiertere Aussage wären auch hier weitere Versuche und Experimente nötig.

### 11.2.7 Narrative Fehler/Diegetische Inkonsistenz

Die Fehler des Sounddesigns und das Festhalten an der von mir gewählten akustischen Perspektive betrafen nicht nur die Gestaltung des Wasserfalls. Bei der Entdeckung der Schlucht, die ich zu einer Höhle umgewandelt habe, bewegt sich keine der Figuren in der Szene. Die Aufstellung hatte ich wie in Abbildung 9.4 beschrieben gewählt. Peter müsste

sich eigentlich von uns weg in Richtung Abgrund bewegen. Akustisch tut er das aber nicht. Diese narrative Verdichtung scheint allerdings nicht weiter schlimm zu sein, da dieser Fehler – beziehungsweise der elastische Gebrauch der Zeit – nur einem der Rezipient:innen vom 20. bis 22. Juni 2022 aufgefallen ist, und zwar sowohl in True Spatial als auch in 7.1.4.

Begründen würde ich das mit der Immersion. Wir fühlen uns so sehr mit der Geschichte verschmolzen und als Teil von ihr, dass solche Inkonsistenzen und die narrative Verdichtung nicht weiter aufzufallen scheinen. Das gilt auch für die Blende von Szene 11 in Szene 12. In diesem Teil des Hörspiels verlassen wir die diegetische Rolle unserer Figur und lassen uns zu Peter in die Höhle hinabseilen. Aber auch hier scheint das nicht weiter zu stören. Sechs der acht Proband:innen empfanden die Blende nicht als Beeinträchtigung des Immersions-effekts und konnten mir im Nachhinein ein ähnliches Bild der Szenerie beschreiben, wie ich es aufgemalt hatte. Wie in 11.2.2 erwähnt, gaben zwei Proband:innen an, dass sie das Drehen der Szene gestört und verwirrt habe, während die Musik einer der Proband:innen nach eigener Aussage half, weil sie wie ein Anker wirkte, um den herum sich alles dreht.

#### 11.2.8 Zeitaufwand/Tools

Entscheidende Faktoren für die Frage nach der etwaigen Zukunft solcher Hörspiele und der Arbeit daran für den kommerziellen Markt sind der Zeitaufwand und die damit verbundenen Kosten. Laut Ulrich Gerhardt (vgl. Kapitel 2.11) war einer der Gründe für das Scheitern der Kunstkopf-Hörspiele der Zeitaufwand, der mindestens doppelt so hoch war wie der von Stereostücken. Zwar gilt seine Aussage für eine Zeit, in der man noch mit Tonbändern arbeitete. Doch musste ich während meiner Arbeit in beiden mehrdimensionalen Formaten Ähnliches feststellen. Für zwei Szenen von etwa fünf Minuten Länge habe ich 1 ½ Wochen gebraucht. Dafür gibt es meiner Meinung nach zwei Gründe: ein detailreicheres Sounddesign und die Bedienoberflächen.

Wo in Stereo vielleicht zwei Layer an Atmos ausreichen würden, braucht es für ein umhüllendes Hörerlebnis sowohl in 7.1.4 als auch in einem True-Spatial-Setup mehr Klänge, da es subjektiv sonst nicht dicht und real genug klingt. Allein für die Darstellung der oberen wie unteren Hemisphäre in True Spatial sind mehr Klänge nötig. Dabei ist ebenfalls wichtig, was für Klänge man verwendet, speziell für die untere Hemisphäre.

Ein weiteres Problem schließt sich direkt an: das Hochmischen von zweikanaligen Aufnahmen. Ein Großteil der Atmo-Aufnahmen aus der von mir verwendeten Online-Datenbank Soundly sind in Stereo. Hin und wieder finden sich auch Aufnahmen in 5.1, doch die Mehrheit ist zweikanalig. Solche Stereoaufnahmen sind für den konsumerfreundlichen Anwendungsbereich des Fernsehens völlig ausreichend, für die Arbeit in True Spatial oder 7.1.4 aber keineswegs.

Es gibt zwei Möglichkeiten, damit umzugehen. Entweder zieht man die Stereoaufnahmen künstlich in die Breite und layert sie mit ähnlich klingenden Aufnahmen, um so ein Gefühl der Umhüllung zu schaffen und den Raum zu füllen. Oder man benutzt Plugins wie Halo, die in der Lage sind, die zweikanalige Aufnahme auf bis zu 7.1.4 hochzumischen. Ersteres funktionierte in meinem Fall sehr gut, ist aber insofern zeitaufwändig, als man ähnlich klingende Aufnahmen finden muss. Bei der zweiten Methode über ein Upmixing-Plugin spart man zwar Zeit, man ist auch aber weniger flexibel bei der Platzierung und der Aufteilung des Signals. Abgesehen davon ist die Auswahl an solchen Plugins für 3D-Audio zum jetzigen Zeitpunkt sehr überschaubar. Neben dem von NUGEN Audio ist mir nur ein Plugin von Blue Ripple Sound bekannt. Ein Upmixing-Tool wie das von mir verwendete Halo ist jedenfalls auch kein Problemlöser. Es genügt häufig nicht, eine atmosphärische Stereoaufnahme zu nehmen und hochzumischen, da man auf diese Weise vielfach noch akustische Löcher hat. Das kann allerdings auch auf die Ordnung des von mir verwendeten Setups zurückzuführen sein. In einem solchen Setup sind mehr Details nötig, über die 7.1.4 als auch True Spatial verfügen.

Beim Vergleich der beiden Methoden, Atmosphären zu gestalten, musste ich feststellen, dass die erste Methode mit dem Layering nach meiner subjektiven Wahrnehmung ein dichter und realer klingendes Schallfeld ergibt, das mehr Tiefe hat. Über die Musik kann ich kein eindeutiges Urteil fällen, da mir wie erwähnt die Stems nicht zur Verfügung standen. Interessant wäre ein Vergleich der beiden von mir genannten Methoden für Musik.

Der zweite und für mich entscheidendere Grund für den hohen Zeitaufwand ist die Bedienoberfläche beziehungsweise die Tatsache, dass sie nicht haptisch angelegt ist. Bei der Arbeit in Stereo haben wir zum Beispiel Potentiometer (Potis), mit denen sich das Panning haptisch ausführen lässt. Dabei können auch mehrere Potis parallel bewegt und somit Automationen geschrieben werden. Das Problem bei den von mir verwendeten Plugins für das Panning ist, dass ein solcher haptischer Controller fehlt. Sowohl in dem Multi-/Stereo-/RoomEncoder als auch in dem ReaSurroundPan stehen lediglich digitale Potis oder ein virtueller, das Klangobjekt repräsentierender Punkt in einem virtuellen Raum zur Verfügung. Zwar kann man damit sehr genau arbeiten, möglich ist aber stets nur die Bearbeitung eines einzelnen Signals und von maximal drei Parametern (X-, Y-, Z-Achse lassen sich parallel über den virtuellen Punkt regeln). Für die Bearbeitung weiterer Parameter, zum Beispiel der Breite des Signals, muss man dieselbe Szene erneut fahren, um den neuen Parameter zu automatisieren. Hat man wie in meinem Fall knapp 120 Spuren, ziehen sich Automatisierung und Platzierung in die Länge. Hat man außerdem beispielsweise eine Automation geschrieben und möchte sie unterdessen bearbeiten (Automationsmodus: Touch/Latch), fällt es sehr schwer, den sich nun bewegendem virtuellen Punkt mit der Maus zu fassen.

Für das zweite Problem, die fehlende haptische Bedienbarkeit, ließe sich womöglich eine Lösung finden. Für das Fehlen mehrkanaliger Atmo-Aufnahmen müsste man deutlich mehr

Töne und Geräusche selbst aufnehmen oder eine eigene Datenbank mit selbstdesignten Atmosphären anlegen.

### 11.2.9 Size matters

Nachdem ich mit dem Sounddesign und den beiden Mischungen fertig geworden war, habe ich mir die beiden Szenen für den Lautsprecherdom in Aufnahmeraum 1 (A1) der HAW DMI gerechnet. Hier haben wir auf der oberen Hemisphäre 33 Lautsprecher, verteilt auf fünf Ringe mit einer durchschnittlichen Lautsprecherentfernung von 3,8 m (Chevalier et al., 2018). Der Unterschied in meiner subjektiven Wahrnehmung war gewaltig. So empfand ich das Innere des kleinen Viertürers, den ich für die erste Hälfte der elften Szene akustisch gestaltet habe, eher als das eines riesigen LKWs. Die Szene funktionierte auf diese Weise überhaupt nicht, da die Größenverhältnisse überdimensioniert waren. Ähnlich verhielt es sich mit der Höhle, die ich plötzlich als um ein Vielfaches größer empfand. Abstände zwischen mir und den akustischen Klangobjekten vervielfachten sich enorm. Die Hörerfahrung war, wie Scherzer (2010) schon vermuten ließ (vgl. Kapitel 4.10), eine völlig andere. Hätte ich dasselbe Hörspiel für A1 gestaltet, wäre vermutlich ein von Grund auf verschiedener virtueller Raum entstanden, mit anders klingenden Klangobjekten und einer womöglich anderen Szenerie.

Ähnlich verhielt es sich bei der binauralen Wiedergabe. Hierfür habe ich den Binaural-Decoder auf den Masterbus gelegt und die Beyerdynamic DT 770 Pro mit 250Ω als Preset ausgewählt. Hier geschah das genaue Gegenteil. Alles wirkte noch näher und intimer als in Regie 3. Mangels Subwoofer wirkte der Wasserfall noch weniger glaubhaft. Auch das drohende Gewitter verlor so an Stärke. Die Umhüllung leidet nicht an der binauralen Wiedergabe. Ein entscheidender Unterschied ist aber, dass die Objekte nicht an ihren Positionen verbleiben, wenn man den Kopf bewegt. Bewege ich den Kopf während der binauralen Wiedergabe, ändere ich auch zum Beispiel im ersten Teil der elften Szene die Bewegungsrichtung des Wagens. Bewege ich meinen Kopf nach links, fährt das Auto nach links. Über Lautsprecher ist das nicht so. Hier ist es möglich, sich in der virtuellen Szene umzuhören, ohne dass die virtuelle Welt ihre Ausrichtung ändert.

Bei Ambisonics als szenenbasiertem Audio scheint die Produktion also keineswegs im selben Maße von dem Wiedergabesystem entkoppelt zu sein wie in der ITU-R BS.2051-2 (2018) (vgl. Kapitel 3.12) beschrieben. Das Wiedergabesystem und dessen Größe spielen meines Erachtens eine erhebliche Rolle sowohl bei der künstlerischen Gestaltung wie bei der Wahrnehmung.

## 12 Fazit

Ziel dieser Arbeit war es, ein Hörspiel in einer vollsphärischen Lautsprecher-Konfiguration akustisch zu gestalten und dabei den Effekt der Immersion zu erzielen. Da ich der Erste meiner Hochschule gewesen bin, der in einer solchen Umgebung eine Produktion erarbeitet hat, kamen während des Sounddesigns/der Mischung und in den Interviews Fragen auf und waren Auffälligkeiten zu beobachten.

Festhalten lässt sich, dass die Erweiterung um eine untere Hemisphäre der klanglichen Gestaltung einen größeren Spielraum bietet, der sich für bestimmte Bewegungsabläufe, SFX sowie Atmos in einer Weise nutzen lässt, wie das bei keinem anderen Format der Fall ist. Dieser erweiterte Spielraum bietet nicht nur neue Möglichkeiten für die virtuelle Raumakustik, er sorgt in der Wahrnehmung auch für ein wirklichkeitsgetreueres Hörerlebnis als in einem Mehrkanalsystem, das sich einzig um Höhenkanäle erweitern lässt. Zwar ist diese Erweiterung keinesfalls notwendig, um das Gefühl von Immersion zu erzeugen, wie Agrawal et al. (2019) sie definiert haben. Sie fördert aber den Übergang in die erzählte virtuelle Welt und die Abkopplung vom physischen Raum. Bei der Gestaltung geht es nicht darum, das Schallfeld von Objekten zu reproduzieren, sondern darum, das Bild dieser Objekte in den Köpfen der Rezipient:innen zu erschaffen. Dafür ist die wirklichkeitsgetreue Reproduktion nicht unbedingt förderlich. Die vollsphärische Beschallung steigert maßgeblich das Gefühl der Umhüllung, wodurch die Partizipation und der Eindruck einer erzählten Welt deutlich besser vermittelt werden.

Die Klänge sollten dabei behutsam ausgewählt, der Einsatz der unteren Hemisphäre wohlüberlegt sein. Bis zu einem gewissen Grad lassen sich Klangobjekte hier gut lokalisieren und ermöglichen die Darstellung virtueller Bewegungsabläufe, wie etwa das Abseilen eines Menschen, an dem vorbei ein Stein fällt, oder das Aufrichten einer Person. Überspannt man diese Möglichkeit aber, kommt es bei bestimmten Klängen zu massiven Fehllokalisationen und einer veränderten und unscharfen Wahrnehmung, so dass Oben und Unten verwechselt werden kann. Von welchen Parametern eine solche Verwechslung abhängt und wie genau die Schärfe der Lokalisation von unten ist, muss in weiteren Untersuchungen erörtert werden.

Für eine immersive Gestaltung in mehrkanaligen Audioformaten muss sich außerdem die Arbeit der Skript-Autoren ändern. In einer frühen Phase der Produktion sollte ein Plan für die akustische Perspektive aufgestellt werden. Rezipient:innen sind hier nicht mehr nur Zuhörer:innen, sie werden selbst Teil der Geschichte. Geschichten spielen sich nicht nur zweidimensional vor ihnen ab, die erzählte Welt spielt sich vielmehr um sie herum ab. Welche Rolle die Rezipient:innen in dieser virtuellen Welt spielen, müssen die Autoren festlegen. Da man von Klängen umhüllt ist, ist die gute klangliche Gestaltung besonders wichtig. Die virtu-

elle Welt muss in sich schlüssig sein und Sinn ergeben. Hören wir etwas, das nach unserer Erfahrung keinen Sinn hat, verlassen wir die virtuelle und suchen nach Gründen in der physischen Welt, und das ist nicht im Sinne der Immersion. Darum ist es ebenso wichtig, Klangobjekte so dinghaft zu gestalten, dass beim ersten Hören sofort zu erkennen ist, um was für ein Objekt es sich handelt. Auch bei der Verwendung nichtdiegetischer Elemente muss man sich im Klaren sein, dass sie einen womöglich in die physische Welt zurückholen.

Als Sounddesigner würde ich in künftigen mehrkanaligen immersiven Hörspielproduktionen ein solches True-Spatial-Setup stets einem 7.1.4-Setup vorziehen, da es mir mehr künstlerischen Freiraum bietet. Aber nicht nur den Klangkünstler:innen kommt die vollsphärische Kulisse zugute, auch Hörer:innen scheinen in der Mehrzahl diese Art der Wiedergabe höher zu schätzen.

## 13 Danksagung

Zum Schluss möchte ich mich bei einigen Menschen bedanken, ohne die diese Arbeit nicht möglich gewesen wäre. Mein erster Dank gilt Pierre Brand von Primetimestudio, der mich in seine Arbeit und den Umgang mit 3D Audio einbezogen hat und für meine Fragen stets ein offenes Ohr hatte. Maike Müller, stellvertretend für Sony Germany, hat mir freundlicherweise erlaubt, die Aufnahmen der Episode der *Drei ??? und das Dorf der Teufel* und das Skript zu verwenden. Den Kontakt zu Maike Müller verdanke ich wiederum Pierre Brand. Herzlich möchte ich mich auch bei NUGEN Audio bedanken, die mir ihr Halo-Upmix-Plugin inklusive der 3D-Erweiterung für die Zeit des Sounddesigns und der Mischung kostenlos zur Verfügung gestellt haben. Großer Dank gebührt meinen Kommiliton:innen Paula Gnoth, Ben Arndt, Oliver Hermes und Julian Fischer, die mir ihre Stimmen für die Verabschiedungsszene zu Beginn der elften Szene geliehen haben. Tobias Falke, wissenschaftlicher Mitarbeiter der HAW, hat mich während meiner Arbeit tatkräftig unterstützt und meine Fragen stets geduldig und gewohnt kompetent beantwortet. Nicht zuletzt hat er mir einige seiner optisch verbesserten Grafiken aus Blauerts Buch von 1974 zur Verfügung gestellt. Stefan Troschka hat mir geholfen, die Regie 3 im Tonlabor der HAW DMI auszumessen und einzurichten. Zusätzlich überließ er mir den Grundriss aus Abbildung 7.1. Ohne ihn wäre diese Forschungsarbeit nicht möglich gewesen. Ebenso hat mich Thomas Görne in besonderer Weise tatkräftig unterstützt und mir bei der Themenfindung ebenso geholfen wie bei der Suche nach Antworten auf meine vielen Fragen. Ich bedanke mich bei allen Rezipient:innen, die mir vom 20. bis 22. Juni 2022 Interviews gegeben haben. Dank gebührt zu guter Letzt meinen Eltern Susanne und Christoph Baron, vor allem meinem Vater, der die Arbeit gelesen und mir stilistische Verbesserungshinweise gegeben hat.

## 14 Anhang

### 14.1 Literaturverzeichnis

- Adams, E. & Rollings, A. (2006). *Fundamentals of Game Design*. Upper Saddle River, N.J.: Pearson Prentice Hall
- Agrawal, S., Simon, A., Bech, S., Bærentsen, K. & Forchhammer, S. (2019). Defining Immersion: Literature Review and Implications for Research on Immersive Audiovisual Experience. *147<sup>th</sup> Audio Engineering Society Convention*
- Akstinat, B. (2008). *Das ABC der drei Fragezeichen* (1. Aufl.). Hannover: Humboldt [entnommen aus Gänßle & Kuchinke, 2020]
- Arnold, A. (Zusammenstellung) (2013). 40 Jahre Kunstkopf. Ein noch nicht abgeschlossenes Kapitel der Hörspielgeschichte [Audio-Podcast]. In *Artmix.Galerie*. Bayern 2. Abgerufen am 14. April 2022 von <https://www.ardaudiothek.de/episode/artmix-galerie/40-jahre-kunstkopf-ein-noch-nicht-abgeschlossenes-kapitel-der-hoerspielgeschichte/bayern-2/78742760/>
- Arsenault, D. (2005). Dark Waters: Spotlight on Immersion. *EUROSIS Game-On North America 2005 Conference*
- Bannasch, B. (2020). Neue Sprechweisen: das Nachkriegshörspiel von Eich bis Bachmann. In N. Binczek & U. Wirth (Hrsg.), *Handbuch Literatur & Audiokultur*. Berlin & Boston: De Gruyter. <https://doi.org/10.1515/9783110340631-028>
- Beckmann, J. & Menger, I. L. (02. Dezember 2019). *Die drei ??? und das Dorf der Teufel. Version 04 – Sprachaufnahme Hauptcast*. © John Beckmann und Ivar Leon [im Archiv des Autors – im Text als „Skript“]
- Berg, J. (2009). The Contrasting and Conflicting Definitions of Envelopment. *126<sup>th</sup> Audio Engineering Society Convention*
- Biocca, F. & Delaney, B. (1995). Immersive Virtual Reality Technology. In F. Biocca & B. Delaney (Hrsg.), *Communication in the Age of Virtual Reality* (S. 57–124). Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc. [entnommen aus Agrawal et al., 2019]
- Blaes, R. & Kraft, H. (Hrsg.). (2002). *Geschichten, die das Medium schrieb. Schriftsteller über 80 Jahre Radio*. Berlin: Vistas [entnommen aus Krug, 2008]
- Blauert, J. (1974). *Räumliches Hören* (1. Aufl.). Stuttgart: S. Hirzel Verlag
- Blauert, J. (1996). *Spatial Hearing. The Psychophysics of Human Sound Localization*. Cambridge, Mass.: MIT Press

- Blauert, J. & Braasch, J. (2008). Räumliches Hören. In S. Weinzierl (Hrsg.), *Handbuch der Audiotechnik* (9. Aufl.) (S. 87–122). Berlin/Heidelberg: Springer
- Blum, A.-C. (2019, 30. Juni). *Hörbuch oder Hörspiel – was ist der Unterschied?* Anomalia. Abgerufen am 14. April 2022 von <https://anomaliam-das-hoerspiel.de/index.php/hoerbuch-oder-hoerspiel-was-ist-der-unterschied/>
- Blumlein, A. D. (1931, 14. Dezember). *Improvements in and relating to sound-transmission sound-recording and sound-reproducing systems*. British Patent Spec. 394.325
- Boren, B. (2017). History of 3D Sound. In A. Roginska & P. Geluso (Hrsg.), *Immersive Sound. The Art and Science of Binaural and Multi-Channel Audio*. London: Routledge (e-book)
- Brecht, B. (1967). Radiotheorie 1927–1932 [fünf Aufsätze]. In *Gesammelte Werke* (Schriften 2, S. 119–134), Frankfurt am Main: Suhrkamp [entnommen aus Kälin, 1991]
- Breh, K. (1973). Was ist Quadrofonie? *Hifi-Stereophonie Testjahrbuch '73*. Abgerufen am 30. März 2002 von <http://www.hifimuseum.de/1973-was-ist-quadrofonie.html>
- Buch, F. P. (1930). Die Kunst des Hörspiels. *Südwestdeutsche Rundfunk-Zeitung (SRZ)*, 51 [entnommen aus Krug, 2008]
- Bülow, R. (2013, 31. August). *Vor 40 Jahren: Ein Kunstkopf für binaurale Stereophonie*. Heise. Abgerufen am 14. April 2022 von <https://www.heise.de/newsticker/meldung/Vor-40-Jahren-Ein-Kunstkopf-fuer-binaurale-Stereophonie-1946286.html>
- Bundestag, Wissenschaftliche Dienste (2016, 05. August). *Kaufkraftvergleiche historischer Geldbeträge* [Sachstand, WD 4 – 3000 – 096/16]. Abgerufen am 14. April 2022 von <https://www.bundestag.de/resource/blob/459032/1d7e8de03e170f59d7cea9bbf0f08e5c/wd-4-096-16-pdf-data.pdf>
- Bung, S. & Schrödl, J. (Hrsg.). (2017). *Phänomen Hörbuch. Interdisziplinäre Perspektiven und medialer Wandel*. Bielefeld: transcript Verlag
- Chaffey, R. & Shirley, B. G. (2007). Up-Mixing and Localisation – Localisation Performance of Up-Mixed Consumer Multichannel Formats. *122<sup>th</sup> Audio Engineering Society Convention*
- Chevalier, N. P., Majdak, P., Wilk, E. & Görne, T. (2018). Rapid HRTF measurement in a loudspeaker dome. *145<sup>th</sup> AES Convention*
- Cusy, P. & Germinet, G. (1994). Maremoto: A radio play (1924). *Réseaux. The French journal of communication*, 2 (2), 251–265. Abgerufen am 14. April 2022 von [https://www.persee.fr/doc/reso\\_0969-9864\\_1994\\_num\\_2\\_2\\_3281](https://www.persee.fr/doc/reso_0969-9864_1994_num_2_2_3281)
- Davis, M. F. (2003, 15. Juni). History of Spatial Coding. *Journal of the Audio Engineering Society*, 51, 554–569
- Dickreiter, M. (2008). Schallwahrnehmung. In M. Dickreiter et al. (Hrsg.), *Handbuch der Tonstudioteknik* (Band 1, 7. Aufl.) (S. 95–112). München: K. G. Saur

- Die drei ??? (o. D.). *Geschichte und Live-Hörspiel-Historie*. Abgerufen am 24. März 2022 von <https://dreifragezeichen.de/bobs-archiv/geschichte>
- Die drei ??? (o. D.). *3D Audio*. Abgerufen am 7. Juni 2022 von <https://www.dreifragezeichen.de/3d-audio>
- Die drei ??? Wiki (o. D.). *Das Dorf der Teufel*. Abgerufen am 25. Mai 2022 von [https://diedreifragezeichen.fandom.com/wiki/Das\\_Dorf\\_der\\_Teufel](https://diedreifragezeichen.fandom.com/wiki/Das_Dorf_der_Teufel)
- Die drei ??? (12. November 2021). *Die drei ??? – „Das Dorf der Teufel“*. *Planetariums-Hörspiel | Spezial-Feature*. YouTube. Abgerufen am 15. Juni 2022 von <https://www.youtube.com/watch?v=ipM9HE-Zzoo>
- Dietze, E. R. (1954, 25. November). *Hans Bredow: Der Weg zum Rundfunk*. Interview. SWR2. Abgerufen am 14. April 2022 von <https://www.swr.de/swr2/wissen/archivradio/hans-bredow-der-weg-zum-rundfunk-100.html>
- Döhl, R. (1992). *Das Hörspiel zur NS-Zeit. Geschichte und Typologie des Hörspiels*. Darmstadt: Wissenschaftliche Buchgesellschaft
- Dudenredaktion (o. D.). Hörspiel. In *Duden online*. Abgerufen am 22. März 2022 von <https://www.duden.de/rechtschreibung/Hoerspiel#close-cite>
- Dudenredaktion (o. D.). Originalton. In *Duden online*. Abgerufen am 23. März 2022 von <https://www.duden.de/rechtschreibung/Originalton>
- Elliott, H. (2020, 16. Oktober). Rolls-Royce's New Car Was So Quiet at First, It Nauseated Drivers. Bloomberg.com. Abgerufen am 27. Mai 2022 von <https://www.bloomberg.com/news/articles/2020-10-16/rolls-royce-s-new-car-was-so-quiet-at-first-it-made-drivers-nauseous>
- Falke, T., Batke, J.-M. & Görne, T. (2014). Remixing a Historic Film in Higher Order Ambisonics 3-D Audio – Workflow and Technical Solutions. *136<sup>th</sup> Audio Engineering Society Convention*
- Falke, T. (2016). *Dreidimensionale Lokalisierbarkeit und Differenzierbarkeit von Hörereignissen* (Masterarbeit, Hochschule für Angewandte Wissenschaften Hamburg). Abgerufen am 9. April 2022 von [https://www.researchgate.net/publication/314632374\\_Dreidimensionale\\_Lokalisierbarkeit\\_und\\_Differenzierbarkeit\\_von\\_Horeignissen](https://www.researchgate.net/publication/314632374_Dreidimensionale_Lokalisierbarkeit_und_Differenzierbarkeit_von_Horeignissen)
- Field, A. (2002). Breaking and making the rules: sound design in 6.1. *113<sup>th</sup> Audio Engineering Society Convention*
- Fletcher, H. & Sivian, L. J. (1927). Binaural telephone system. US Patent 1,624,486. [entnommen aus Boren, 2017]
- Fletcher, H. (1933). An Acoustic Illusion Telephonically Achieved. *Bell Laboratories Record*, 11 (10), 286–289 [entnommen aus Boren, 2017]
- Flückiger, B. (2006). *Sound Design: Die virtuelle Klangwelt des Films* (3. Aufl.). Marburg: Schüren [entnommen aus Scherzer, 2010]

- Fuchs, C. (2018). *3D Audio. Wie 3D Audio im Vergleich zu Surround Sound die Wahrnehmung und den Fokus eines Spielers in Virtual Reality-Anwendungen auf Basis von Kopfhörern beeinflussen und lenken kann* (Bachelorarbeit, Hochschule der Medien Stuttgart). Abgerufen am 5. Juni 2022 von <https://curdt.home.hdm-stuttgart.de/PDF/Fuchs.pdf>
- Gänßle, S. & Kuchinke, B. (2020). *Die Hörspielserie „Die drei ???“ – Der wirtschaftliche Erfolg und Gründe dafür*. (Ilmenau Economics Discussion Papers, No. 134). Ilmenau: Technische Universität Ilmenau, Institut für Volkswirtschaftslehre. Abgerufen am 14. April 2022 von <https://www.econstor.eu/bitstream/10419/210651/1/1686650299.pdf>
- Gerzon, M. A. (1973). Periphony: With-Hight Sound Reproduction. *Journal of the Audio Engineering Society*, 21 (1), 2–10
- Gisselbrecht, J. (o. D.). *Hörspiel und Hörbuch – Unterschiede*. hoerspiel.com. Abgerufen am 14. April 2022 von <http://www.hoerspiel.com/geschichte-zeitlauf-hoerspiel/definition-hoerspiel/>
- Goldstein, E. B. (2015). *Wahrnehmungspsychologie* (9. Aufl.) (aus dem Englischen übersetzt von Katharina Neuser-von Oettingen & Guido Plata). Berlin/Heidelberg: Springer
- Görne, T. (2015). *Tontechnik* (4. Aufl.). München: Hanser
- Görne, T. (2017). *Sounddesign. Klang, Wahrnehmung, Emotion*. München: Hanser
- Hamasaki, K., Nishiguchi, T., Hiyama, K. & Ono, K. (2004). Advanced Multichannel Audio Systems with Superior Impression of Presence and Reality. *116<sup>th</sup> Audio Engineering Society Convention*
- Hawes, R. (1991). *Radio Art*. London: Green Wood
- Heister, H. S. von (1925, 04. Oktober). Nochmals: Das Hörspiel. *Der Deutsche Rundfunk*, 3 (40). Abgerufen am 14. April 2022 von <https://hoerspiele.dra.de/pdf/heister-nochmals-hoerspiel.pdf>
- Hiyama, K., Komiyama, S. & Hamasaki, K. (2002). The Minimum Number of Loudspeakers and its Arrangement for Reproducing the Spatial Impression of Diffuse Sound Field. *113<sup>th</sup> Audio Engineering Society Convention*
- Holman, T. (2008). *Surround Sound: Up and running* (2. Aufl.). Amsterdam: Focal Press
- Höltgen, S. (2009, 17. April). *Ästhetische Immersion und Filmtheorie*. Filmforen.de. Abgerufen am 18. April 2022 von <http://www.filmforen.de/index.php?showtopic=14140>
- Hönemann, K. (2011). Die verschiedenen Systeme der Quadrophonie 1969–80. Abgerufen am 31. März 2022 von [https://www.quadrophonie.de/deutsch/quadro\\_1969\\_-\\_80.html](https://www.quadrophonie.de/deutsch/quadro_1969_-_80.html)
- Hörburger, C. (2006). *Hörspiel*. SWR. Abgerufen am 14. April 2022 von [https://www.lmz-bw.de/fileadmin/user\\_upload/Downloads/Handouts/hoerburger-hoerspiel-swr.pdf](https://www.lmz-bw.de/fileadmin/user_upload/Downloads/Handouts/hoerburger-hoerspiel-swr.pdf)
- ITU (International Telecommunication Union) (2018). Advanced sound system for programme production (Recommendation ITU-R BS.2051-2)

- ITU (International Telecommunication Union) (2019). Multichannel sound technology in home and broadcasting applications (Report ITU-R BS.2159-8)
- Kälin, S. (1991). *Die Anfänge des Hörspiels in der Weimarer Republik. Versuch einer Analyse*. Stuttgart: Verlag Hans-Dieter Heinz
- Kapeller, L. (1924, 14. November). Selbsterkenntnis des Rundfunks. Versuch einer „furchtbaren Kritik“. *Funk. Die Wochenschrift des Funkwesens*, 1 (29), 437–439. Abgerufen am 14. April 2022 von <https://soundandscience.de/text/funk-die-wochenschrift-des-funkwesens-1924-heft-18-35>
- Keckeis, H. (1973). *Das deutsche Hörspiel. 1923–1973. Ein systematischer Überblick mit kommentierter Bibliographie*. Frankfurt am Main: Athenäum Verlag [entnommen aus Köhler, 2005]
- Kim, S., Lee, Y. W. & Pulkki, V. (2010). New 10.2-Channel Vertical Surround System (10.2-VSS); Comparison Study of Perceived Audio Quality in Various Multichannel Sound Systems with Height Loudspeakers. *129<sup>th</sup> Audio Engineering Society Convention*
- Ko, S., Oh, E., Park, S. H. & Shim, H. (2010). Perceptual Evaluation of Spatial Audio Quality. *129<sup>th</sup> Audio Engineering Society Convention*
- Köhler, S. (2005). *Hörspiel und Hörbuch. Mediale Entwicklung von der Weimarer Republik bis zur Gegenwart*. Marburg: Tectum Verlag
- Krug, H.-J. (2008). *Kleine Geschichte des Hörspiels* (2. Aufl.). Konstanz: UVK Verlagsgesellschaft
- Leonhard, J.-F. (Hrsg.). (1997). *Programmgeschichte des Hörfunks in der Weimarer Republik*. (2 Bände). München: DTV
- Lombard, M., Ditton, T. B. & Weinstein, L. (2009). Measuring Presence: The Temple Presence Inventory. *Proceedings of the 12<sup>th</sup> Annual International Workshop on Presence*, 1–15 [entnommen aus Agrawal et al., 2019]
- Lund, T. (2000). Enhanced Localization in 5.1 Production. *109<sup>th</sup> Audio Engineering Society Convention*
- Magnus, K. (1928). Die Rundfunkbewegung im Jahre 1927. *Archiv für Funkrecht*, 1, 1–11
- Maier, F. (2015). *Das Hörspiel. Eine technische Kunstform?* (Dissertation, Philosophische Fakultät). Abgerufen am 5. Juni 2022 von [https://opus.bibliothek.uni-wuerzburg.de/opus4-wuerzburg/frontdoor/deliver/index/docId/13971/file/Das-Hoerspiel-Eine-technische-Kunstform\\_Frank-Maier.pdf](https://opus.bibliothek.uni-wuerzburg.de/opus4-wuerzburg/frontdoor/deliver/index/docId/13971/file/Das-Hoerspiel-Eine-technische-Kunstform_Frank-Maier.pdf)
- Martin, G. (2002). Interchannel Interference at the Listening Position in a Five-channel Loudspeaker Configuration. *113<sup>th</sup> Audio Engineering Society Convention*
- Martin, G. (2005). A New Microphone Technique For Five-Channel Recording. *118<sup>th</sup> Audio Engineering Society Convention*

- Mausfeld, R. (2005). Wahrnehmung: Geschichte und Ansätze. In J. Funke & P. A. French (Hrsg.), *Handbuch der Allgemeinen Psychologie – Kognition* (Handbuch der Psychologie, Band 5) (S. 97–107). Göttingen: Hogrefe
- McMahan, A. (2003). Immersion, Engagement, and Presence: A Method for Analyzing 3-D Video Games. In M. J. P. Wolf & B. Perron (Hrsg.), *The Video Game Theory Reader* (S. 67–86). New York/London: Routledge
- Murray, J. H. (2017). *Hamlet on the holodeck: The future of narrative in cyberspace* (updated edition). Cambridge, Mass./London: MIT Press
- Natusch, M.-C. (2014, August). *Das Comeback der Hörspiele – Analyse und Strategieempfehlungen von Online-Marketing-Maßnahmen für kleine und mittelständische Hörspiellabels in Deutschland*. (Bachelorarbeit, HAW Hamburg, Studiengang Medien und Information). Abgerufen am 5. Juni 2022 von <https://reposit.haw-hamburg.de/handle/20.500.12738/7492?&locale=en>
- Nilsson, N. C., Nordahl, R. & Serafin, S. (2016, November). Immersion Revisited: A Review of Existing Definitions of Immersion and Their Relation to Different Theories of Presence. *Human Technology*, 12(2), 108–134 [entnommen aus Agrawal et al., 2019]
- o. V. [vermutlich von Heister, H. S.] (1924, 10. August). Zur Frage der Sendespiele. *Der Deutsche Rundfunk*, 32, 1779. Abgerufen am 5. Juni 2022 von <https://hoerspiele.dra.de/pdf/nn-sendespiele.pdf>
- Paul, S. (2009). Binaural Recording Technology: A Historical Review and Possible Future Developments. *Acta Acustica united with Acustica*, 95 (5), 767–788.
- Philipson, J., Hirst, J. & Woollard, S. (2002). Investigation into a Method for Predicting the Perceived Azimuth Position of a Virtual Image. *Convention Paper 5597, presented at the 112<sup>th</sup> Audio Engineering Society Convention*
- Power, P. J. (2015). *Future Spatial Audio: Subjective Evaluation of 3D Surround Systems* (PhD thesis, University of Salford, UK). Abgerufen am 5. Juni 2022 von <http://usir.salford.ac.uk/id/eprint/34100/>
- Radio Drama at 90: Julius Caesar. Some of the landmarks in the history of Radio Drama and Shakespeare on the radio.* (o. D.). BBC. Abgerufen am 21. März 2022 von <https://www.bbc.co.uk/programmes/articles/2D0KC1p9vJMt2dN2SS7ZWng/radio-drama-at-90>
- Ratliff, P. A. (1974). *Properties of hearing related to quadrasonic reproduction*. Research Department, Engineering Division, BBC
- Ryan, M.-L. (2003). *Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media*. Baltimore, Md.: The Johns Hopkins University Press
- Schabbing, B. (2005). *Musik- und Audiot Technologien zwischen Technik, Marketing und Kundenwunsch. Praxisstudien zur Einführung von Quadrophonie, Dolby Surround, Compact Disc und Music on Demand auf dem deutschen Markt*. Osnabrück: Electronic Publi-

- shing Osnabrück. Abgerufen am 31. März 2022 von <https://www.epos.uni-osnabrueck.de/books/s/schb005/pages/iii.htm>
- Schanze, H. (Hrsg.). (2002). *Metzler-Lexikon Medientheorie, Medienwissenschaft*. Stuttgart & Weimar: J. B. Metzler
- Scherzer, J. (2010). *Der Tondramaturgische Raum als Gestaltungsmittel für die Immersive Ästhetik* (Diplomarbeit, Filmuniversität Babelsberg). Abgerufen am 5. Juni 2022 von <https://opus4.kobv.de/opus4-filmuniversitaet/frontdoor/index/index/year/2010/docId/67>
- Schöning, K. (Hrsg.). (1982). *Spuren des Neuen Hörspiels*. Frankfurt am Main: Suhrkamp [entnommen aus Krug, 2008]
- Schwitzke, H. (1963). *Das Hörspiel: Dramaturgie und Geschichte*. Köln & Berlin: Kiepenheuer und Witsch
- Sengpiel, E. (o. D.). *Richtungshören als Lokalisation beim natürlichen Hören in den drei Ebenen*. Abgerufen am 7. April 2022 von <http://www.sengpielaudio.com/RichtungshoerenInDreiEbenen.pdf>
- Sennheiser (o. D.). *Unsere Geschichte*. Sennheiser. Abgerufen am 24. März 2022 von <https://de-de.sennheiser.com/about-us-our-company-our-history>
- Skript, *siehe Beckmann & Menger, 2019*
- Slater, M. (2003, January). A note on presence terminology. *Presence connect*. Abgerufen am 26. Mai 2022 von [https://www.researchgate.net/profile/Mel-Slater/publication/242608507\\_A\\_Note\\_on\\_Presence\\_Terminology/links/0046352dc18e11dcf1000000/A-Note-on-Presence-Terminology.pdf?origin=publication\\_detail](https://www.researchgate.net/profile/Mel-Slater/publication/242608507_A_Note_on_Presence_Terminology/links/0046352dc18e11dcf1000000/A-Note-on-Presence-Terminology.pdf?origin=publication_detail)
- SSA Plugins (2017, 18. Juli). *What is... Higher Order Ambisonics?* SSA Plugins. Abgerufen am 25. Mai 2022 von <https://www.ssa-plugins.com/blog/2017/07/18/what-is-higher-order-ambisonics/>
- Steinberg, J. C. & Snow, W. B. (1934). Auditory Perspective – Physical Factors. *Electrical Engineering*, 53, 12–17
- Steinke, G. (1996). Surround Sound – The New Phase. An Overview. *100<sup>th</sup> Audio Engineering Society Convention*
- Streicher, R. & Everest, A. (1998). *The New Stereo Soundbook* (2. Aufl.). Pasadena: Audio Engineering Associates [entnommen aus Power, 2015]
- Stuart, J. R. (1996). The Psychoacoustics of Multichannel Audio. *11<sup>th</sup> Audio Engineering Society Conference* [entnommen aus Power, 2015]
- Stuhlmann, A. (2001). *Radio-Kultur und Hör-Kunst. Zwischen Avantgarde und Popularkultur 1923–2001*. Würzburg: Königshausen & Neumann [entnommen aus Krug, 2008]
- Theile, G. & Plenge, G. (1977, 1. April). Localization of Lateral Phantom Sources. *Journal of Audio Engineering Society*, 25, 196–200

- Theile, G. (1993). The New Sound Format "3/2-Stereo". *94<sup>th</sup> Audio Engineering Society Convention*
- Theile, G. (2008). Tonaufnahme und Tonwiedergabe. In M. Dickreiter et al. (Hrsg.), *Handbuch der Tonstudioteknik* (Band 1, 7. Aufl.) (S. 183–321). München: K. G. Saur
- Thon, J.-N. (2008). Immersion Revisited: On the Value of a Contested Concept. In O. Leino, H. Wirman & A. Fernandez (Hrsg.), *Extending Experiences. Structure, Analysis and Design of Computer Game Player Experience* (S. 29–43). Rovaniemi, Finland: Lapland University Press [entnommen aus Agrawal et al., 2019]
- Torick, E. (1998). Highlights in the History of Multichannel Sound. *Journal of the Audio Engineering Society*, 46, 27–31
- Troschka, S., Stephan, M., Wong, Y.-F. & Görne, T. (2021). Listening experiment on the perception of spatial sound configurations. *Conference: 2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*
- Tsingos, N., Chabanne, C., Robinson, C. & McCallus, M. (2011). Surround Sound with Height in Games Using Dolby Pro Logic IIz. *41<sup>st</sup> Audio Engineering Society International Conference: Audio for Games*
- Verstraten, P. (2007). Diëgese. *Versus*, 1, 59–70. Abgerufen am 29. Mai 2022 von <https://ugp.rug.nl/versus/article/view/30785/28085>
- Watkinson, J. (1998). *The Art of Sound Reproduction*. Oxford: Focal Press
- Wellershoff, D. (1985). *Die Arbeit des Lebens. Autobiographische Texte*. Köln: Kiepenheuer und Witsch
- Wenzlaff, O. (2013, 28. März). *Das Hörspiel und der lange Schatten des Buches*. YouTube. Abgerufen am 25. März 2022 von <https://www.youtube.com/watch?v=T-y7aQAMhqY>
- Witting-Nöthen, P. (2000). Die Einführung der Stereophonie im Rundfunk. *Geschichte im Westen (GiW)*, 15 (2), 185–195
- Wulff, H. J. (2022, 22. März). Diegese. *Lexikon der Filmbegriffe* (online). Abgerufen am 29. Mai 2022 von <https://filmlexikon.uni-kiel.de/doku.php/d:diegese-12>

## Eigenständigkeitserklärung

„Ich versichere, die vorliegende Arbeit selbständig ohne fremde Hilfe verfasst und keine anderen Quellen und Hilfsmittel als die angegebenen benutzt zu haben. Die aus anderen Werken wörtlich entnommenen Stellen oder dem Sinn nach entlehnten Passagen sind durch Quellenangaben eindeutig kenntlich gemacht.“

Hamburg, den 26.07.2022  