

Friedrich-Alexander-Universität Erlangen-Nürnberg

**Lehrstuhl für Multimediakommunikation und  
Signalverarbeitung**

Prof. Dr.-Ing. André Kaup

Bachelorarbeit

**Erstellung von Hörbeispielen für ein  
Wellenfeldsynthese-System**

von Florian Josef Rettelbach

14. Oktober 2011

Betreuer: Prof. Dr.-Ing. Rudolf Rabenstein



# Erklärung

Ich versichere, dass ich die vorliegende Arbeit ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe, und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

---

Ort, Datum

---

Unterschrift



# Inhaltsverzeichnis

<b>Kurzfassung</b>	<b>V</b>
<b>Abkürzungsverzeichnis</b>	<b>VI</b>
<b>Formelzeichen</b>	<b>VII</b>
<b>1 Einleitung</b>	<b>1</b>
<b>2 Räumliche Wiedergabeverfahren</b>	<b>3</b>
2.1 Amplitudendifferenz-Stereofonie . . . . .	3
2.1.1 Wiedergabetechnik . . . . .	4
2.1.2 Aufnahmetechnik . . . . .	6
2.1.3 Beurteilung und Erweiterung des Verfahrens . . . . .	7
2.2 Rundumklang . . . . .	9
2.2.1 Wiedergabetechnik . . . . .	9
2.2.2 Aufnahmetechnik . . . . .	10
2.2.3 Beurteilung des Verfahrens . . . . .	10
2.3 Ambisonics . . . . .	11
2.3.1 Beziehung zwischen Aufnahme- und Wiedergabetechnik . . . . .	11
2.3.2 Mathematischer Hintergrund . . . . .	12
2.3.3 Beurteilung des Verfahrens . . . . .	14
2.4 Wellenfeldsynthese . . . . .	16
2.4.1 Beschreibung des Wellenfeldsynthese-Wiedergabeverfahrens . . . . .	16

2.4.2	Mathematischer Hintergrund der Wellenfeldsynthese . . . . .	17
2.4.3	Aufnahmetechnik . . . . .	19
2.4.4	Beurteilung des Verfahrens . . . . .	19
2.5	Binaurale Wiedergabe . . . . .	21
2.5.1	Beschreibung des menschlichen Gehörs . . . . .	21
2.5.2	Messung der HRTFs und Wiedergabetechnik . . . . .	22
2.5.3	Beurteilung des Verfahrens . . . . .	24
<b>3</b>	<b>Spiegelquellen Theorie</b>	<b>25</b>
<b>4</b>	<b>Objektbasierte Wiedergabe von Audio Szenen</b>	<b>27</b>
4.1	Soundscape Renderer . . . . .	27
4.2	Audio-Szenen Beschreibungsformat . . . . .	30
<b>5</b>	<b>Vohandene Wellenfeldsynthese-Demos</b>	<b>35</b>
<b>6</b>	<b>Audio-Szenen-Portierung durch Nutzung des SSR</b>	<b>39</b>
6.1	Portierung Kanalbasierter Audio-Dateien . . . . .	41
6.2	Portierung der akustischen Gegebenheiten des Aufnahmerraums . . . . .	44
6.3	Portierung von Schallquellenbewegungen . . . . .	47
6.3.1	Lineare Bewegung . . . . .	49
6.3.2	Zirkulare Bewegung . . . . .	49
6.3.3	Spirale Bewegung . . . . .	50
6.4	Kombination von Schallquellenbewegungen und Raumakustik . . . . .	52
<b>7</b>	<b>Ausblick</b>	<b>56</b>
<b>8</b>	<b>Zusammenfassung</b>	<b>58</b>
<b>A</b>	<b>Anhang Kapitel</b>	<b>60</b>
A.1	How To SSR . . . . .	60
A.1.1	Installation . . . . .	60

---

A.1.2	Jack Audio Server . . . . .	61
A.1.3	Starten des Soundscape Renderers . . . . .	62
A.1.4	Ausführen der Python- bzw. Shell-Skripte . . . . .	63
	<b>Abbildungsverzeichnis</b>	<b>65</b>
	<b>Tabellenverzeichnis</b>	<b>67</b>
	<b>Literaturverzeichnis</b>	<b>69</b>





# Kurzfassung

Diese Arbeit zeigt auf welche Probleme bei der Portierung von kanal-basierten Audiodaten räumlicher Audioszenen zwischen verschiedenen Wiedergabesystemen bzw. Wiedergabeverfahren auftreten können. Zusätzlich werden Lösungsansätze bereit gestellt, um diese Probleme zu umgehen. Weiterhin wurden Ansätze entwickelt um Probleme dieser Art gar nicht erst entstehen zu lassen, indem eine Audioszene objekt-orientiert aufgezeichnet wird. Die neu entwickelten Ansätze wurden in einer kleinen Demonstration zusammengefasst.

# Abkürzungsverzeichnis

ALSA	Advanced Linux Sound Architecture
ASDF	Audio Scene Description Format
BRIR	Binaural Room Impulse Response
HOA	Higher Order Ambisonics
HRIR	Head Related Impulse Response
HRTF	Head Related Transfer Function
ILD	Interaural Level Difference
ITD	Interaural Time Difference
LMS	Lehrstuhl für Multimediakommunikation und Signalverarbeitung
SSR	SoundScape Renderer
WFS	Wellenfeldsynthese
VBAP	Vector Based Amplitude Panning
XML	eXtensible Markup Language

# Formelzeichen

$c_s$	Schallgeschwindigkeit
$G(\boldsymbol{x} \boldsymbol{\xi}, \omega)$	Greensche Funktion
$P, p$	Schalldruck
$\rho$	Luftdichte
$Q(x, y)$	Schallquelle
$SQ(x_s, y_s)$	Spiegelquelle
$V$	Schallschnelle
$W(x', y')$	Wand parallel zu x- bzw. y-Achse
$z_0$	Schallkennimpedanz



# Kapitel 1

## Einleitung

Der Lehrstuhl für Multimediakommuniktion und Signalverarbeitung der Friedrich-Alexander Universität Erlangen-Nürnberg besitzt für Forschungszwecke einen 48-Kanal-Lautsprecherring zur Nachbildung von Schallwellen mittels Wellenfeldsynthese. Zur Erweiterung des Forschungserfolges wird im Moment ein größeres Wellenfeldsynthese-System gebaut. Somit ist es nötig, die für das alte System entwickelten Schallfeld-Demonstrationen auf dem Neuen verfügbar zu machen. Diese Demonstrationen stehen problematischerweise größtenteils nur als 48-Kanal-.wav-Datei, also in kanal-basierter Form, zu Verfügung.

Audio-Dateien in kanal-basierter Form sind speziell auf ein Wiedergabe-System und -Verfahren ausgelegt. Informationen über z.B. räumliche Bewegung der Schallquellen, Lautstärkeveränderung und Echo-Effekte sind in den Signalen der einzelnen Kanäle überlagert und können nur schwer extrahiert werden. Um die Portierung von Audio-Szenen zwischen den verschiedenen Wiedergabe-Systemen und -Verfahrenen zu erleichtern, ist es sinnvoll eine Objektorientierung auf die Szenen anzuwenden. Dabei wird jede der genutzten Schallquellen einzeln aufgezeichnet und zusätzlich Informationen über deren Eigenschaften gespeichert. Mögliche Eigenschaften wären räumliche Bewegung, Lautstärkeveränderung und die akustischen Eigenschaften des Aufnahmerraums. Im Rahmen dieser Arbeit sollen Möglichkeiten zur Portierung von kanal-basierten Audio-Dateien zwischen verschiedenen Wiedergabe-Systemen und -Verfahrenen be-

schrieben werden. Zusätzlich werden Ansätze zur Beschreibung von objekt-basierten Audio-Szenen aufgezeigt.

## Kapitel 2

# Räumliche Wiedergabeverfahren

Zur Wiedergabe von Audiomaterial gibt es zahlreiche Techniken um ein akustisches Ereignis mehr oder weniger originalgetreu wiederzugeben. Im folgenden Kapitel sollen mehrere Möglichkeiten im näheren erläutert und auf Vor- und Nachteile eingegangen werden. Des Weiteren stehen die Aufnahmetechniken der jeweiligen Wiedergabemethode im Fokus dieses Kapitels.

### 2.1 Amplitudendifferenz-Stereofonie

Dieses weit verbreite zwei-kanalige Wiedergabe-Verfahren, auch Intensitätsstereophonie genannt, nutzt nach [1][2][3] Unterschiede zwischen den beiden Lautsprecheramplituden um Hörereignisse zu erzeugen. Die Lautsprecher senden somit simultan das gleiche Signal nur mit unterschiedlicher Amplitude. Im ein-dimensionalen Bereich können dabei für den Zuhörer oftmals ausreichende Ergebnisse erzielt werden. Durch eine vergleichsweise einfache und finanziell wenig aufwendige Umsetzung ist es das seit den 50er Jahren populärste Wiedergabeverfahren.

### 2.1.1 Wiedergabetechnik

Abbildung 2.1 zeigt den schematischen Aufbau eines 2-Kanal-Stereosystems. Im Koordinatenursprung befindet sich der Zuhörer, welcher von den Lautsprechern L und R beschallt wird. Dabei stellt die Anordnung bei einem Winkel von  $\theta_0 = 30^\circ$  ein gleichseitiges Dreieck dar. Durch die Pegeldifferenz der beiden Lautsprecher können Phantomschallquellen auf der Strecke zwischen L und R erzeugt werden, wobei die Quelle je nach Amplitudenunterschied im Winkel  $\theta$  erscheinen. Ein Tiefeneindruck kann jedoch eher weniger vermittelt werden. Unter der Annahme, dass sich der Zuhörer bei  $x = 0$  im Fernfeld befindet, kann die ankommende Schallwelle als eben angesehen werden. Aus dieser Vereinfachung heraus leitet sich das **Sinus-Gesetz der Stereophonie** ab, welches in Abbildung 2.2 als durchgehende Linie aufgetragen ist. Es bildet folgenden Zusammenhang

$$\sin \theta = \frac{1 - \frac{g_l}{g_r}}{1 + \frac{g_l}{g_r}} \sin \theta_0, \quad (2.1)$$

wobei  $g_l$  bzw.  $g_r$  jeweils für die Amplitude des linken und des rechten Lautsprechers stehen.

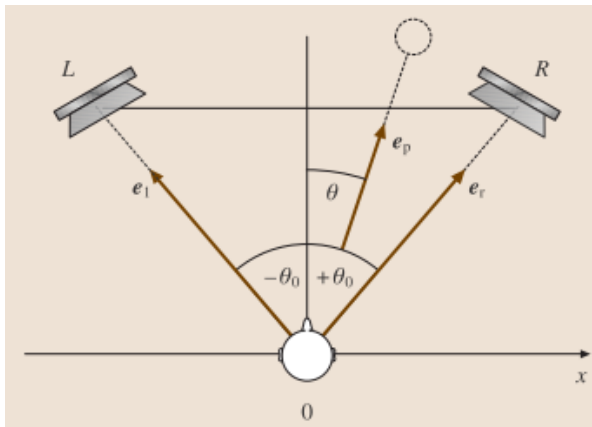


Abbildung 2.1: Geometrisches Schema 2-Kanal-Stereosystems mit den Einheitsvektoren  $e_l$  und  $e_r$  in Richtung der Lautsprecher bzw.  $e_p$  in Richtung der Phantomschallquelle [2].

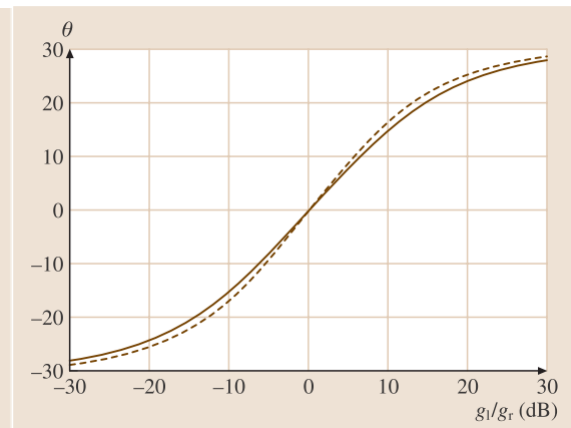


Abbildung 2.2: Beziehung zwischen Phantomschallquellenwinkel  $\theta$  und Amplitudenverhältnis  $g_l/g_r$  des **Tangens-Gesetz** (gestrichelte Linie) bzw. **Sinus-Gesetz** (durchgehende Linie) [2].



Die gestrichelte Linie in Abbildung 2.2 stellt das **Tangens-Gesetz der Stereophonie** dar. Es wird abgeleitet aus der Projektion der Phantomschallquelle bei Winkel  $\theta$  auf ein Koordinatensystem erzeugt durch die Einheitsvektoren  $\mathbf{e}_l$  und  $\mathbf{e}_r$  in Richtung der beiden Lautsprecher. Man erhält dadurch folgenden Zusammenhang

$$\tan \theta = \frac{1 - \frac{g_l}{g_r}}{1 + \frac{g_l}{g_r}} \tan \theta_0 \quad (2.2)$$

mit den Gewichtungsfaktoren der beiden Lautsprecher  $g_l$  und  $g_r$ . Offensichtlich bieten beide Berechnungsverfahren nur geringfügig unterschiedliche Ergebnisse.

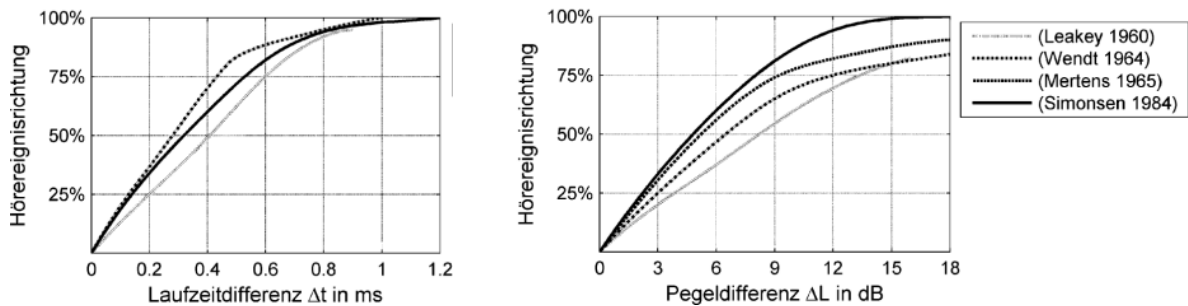


Abbildung 2.3: Direkter Vergleich zwischen Einfluss der Laufzeit- bzw. Pegeldifferenz auf die Wahrnehmung der Hörereignisrichtung (  $0\% \rightarrow \theta = 0^\circ$ ,  $100\% \rightarrow \theta = \theta_0$ ) [4].

Neben der Möglichkeit Phantomschallquellen durch Pegelunterschiede zu erzeugen, kann eine räumliche Schalldarstellung auch durch Laufzeitunterschiede zwischen den beiden Lautsprechern entstehen. Das menschliche Gehör nutzt hierbei interaurale Laufzeitdifferenzen zur Lokalisation der Schallquelle. In Abbildung 2.3 sind im direkten Vergleich die Ergebnisse mehrerer Hörversuche aufgezeigt. Auf der Ordinate ist jeweils die Wahrnehmung der Hörereignisrichtung aufgetragen. Hierbei entsprechen 0% dem Mittelpunkt zwischen den Lautsprechern und 100% der Lautsprecherposition. Die Hörtests wurden hinsichtlich verschiedener Bedingungen durchgeführt. Näheres dazu kann in [4] nachgelesen werden. Bis 75% der Hörereignisrichtung ist der Verlauf beider Graphen weitestgehend linear. Bei den Laufzeitdifferenzen weisen die empirisch ermittelten Kurven danach einen deutlicheren Anstieg auf. Vergleicht man nun den Graphen aus Abbildung 2.2 des **Sinus-** bzw. **Tangens-Gesetz** der Stereophonie mit dem empirisch ermittelten Graphen aus Abbildung 2.3 hinsichtlich der Pegeldifferenz, fällt ein deutlicher

Unterschied auf. Die subjektive Wahrnehmung insbesondere bei Simonsen, wie in Abbildung 2.3 zu sehen ist, lokalisiert die Schallquelle bereits bei einer Pegeldifferenz  $\Delta L$  von 15 dB in Lautsprecher-nähe. Die objektiven Sinus- bzw. Tangens-Gesetze befinden sich hier erst bei ca.  $\tan(20^\circ)/\tan(30^\circ) = 0.63$ , also 63% der Strecke zwischen Mittelpunkt und Lautsprecher. Die Gesetze, welche hinsichtlich des psychoakustischen Wissens aus Abbildung 2.3 entwickelt wurden, vermitteln daher nur einen Eindruck über die Zusammenhänge. Somit bietet die Einfachheit der Amplitudendifferenz-Stereofonie deutliche Nachteile im Hinblick auf die räumliche Wahrnehmung der Schallquelle.

### 2.1.2 Aufnahmetechnik

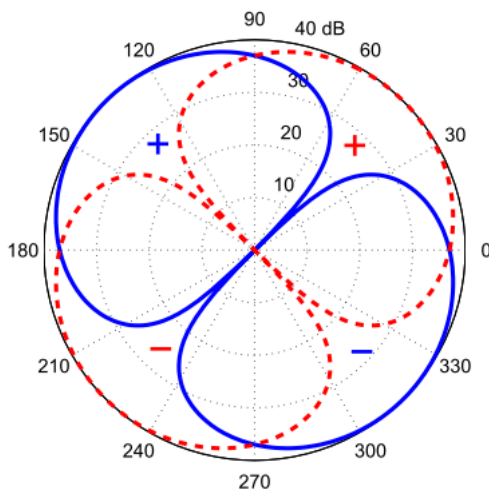


Abbildung 2.4: Blumlein Mikrofonanordnung: Die beiden Linien stellen das typische Aufnahmeverhalten von zwei um  $90^\circ$  versetzte Druckgradienten-Mikrofone abhängig vom Einfallswinkel  $\alpha$  dar [1].

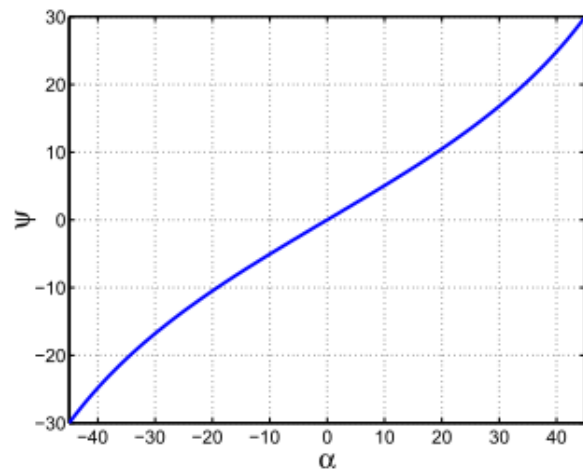


Abbildung 2.5: Beziehung zwischen Aufnahme-  $\alpha$  und Wiedergabewinkel  $\psi$  der Schallquelle [1].

Eine etablierte Stereoaufnahmetechnik ist das sogenannte Blumlein-Paar von 1931 [1]. Dabei werden zwei Druckgradienten-Mikrofone im  $90^\circ$  Winkel zueinander angeordnet. Das Schema ist in Abbildung 2.4 dargestellt. Dabei ergibt sich folgende Beziehung zwischen Empfangsamplituden  $g_l$  bzw.  $g_r$  und dem Winkel  $\alpha$  zwischen der Schallquelle

und der Mittelstellung des Stereo-Mikrofons

$$\tan \alpha = \frac{1 - \frac{g_l}{g_r}}{1 + \frac{g_l}{g_r}}. \quad (2.3)$$

Setzt man nun (2.3) in das Sinus-Gesetz (2.1) ein, erhält man einen Zusammenhang zwischen dem Winkel  $\alpha$  der aufgenommenen Schallquelle und dem Winkel  $\theta$  der bei der Wiedergabe erzeugten Phantomschallquelle

$$\sin \theta = \tan \alpha \cdot \sin \theta_0. \quad (2.4)$$

Dieser Zusammenhang zwischen den beiden Winkeln  $\alpha$  und  $\theta$  ist in Abbildung 2.5 aufgetragen. Eine lineare Beziehung wäre hier natürlich ideal um eine räumliche Verzerrung der Phantomschallquellenerzeugung zu vermeiden.

### 2.1.3 Beurteilung und Erweiterung des Verfahrens

Ungeachtet der Vorteile hinsichtlich der einfachen Umsetzung der Amplitudendifferenz-Stereofonie ergeben sich deutliche Nachteile im Bezug auf die räumliche Darstellung von Schallquellen. Das Sinus-Gesetz der Stereofonie gilt nur im schmalen Bereich des Koordinatenursprungs, da sonst die Wellen nicht mehr als eben angenommen werden können. Befindet man sich außerhalb der Symmetrieachse, ist die Lokalisation stark frequenzabhängig. Eine optimale räumliche Wahrnehmung ist somit nur in einem sogenannten 'sweet spot' gegeben. Die Methode der Laufzeitdifferenzen-Stereofonie spielt eine eher untergeordnete Rolle und wird im Allgemeinen nur zusätzlich zur Amplitudendifferenz-Stereofonie eingesetzt. Als vereinfachte Regel kann der Winkel, welcher durch die beiden Verfahren erzeugt wird, addiert werden wie in [1] beschrieben ist.

In diesem Abschnitt wurde die Methodik der Phantomschallquellenbildung mit Hilfe von Pegeldifferenzen im ein-dimensionalen Bereich, also auf der Strecke zwischen linkem und rechtem Lautsprecher, betrachtet. Es ist durchaus möglich, dieses Verfahren auf mehrere Dimensionen zu erweitern. Die Regeln zur Berechnung der Lautsprecherampli-

tuden werden üblicherweise *Panning-Rules* genannt. Ab einer gewissen Menge an Lautsprechern werden Vektoren zur Beschreibung der Panning-Rules genutzt. Auf Grund der vektoriellen Notation wird diese Methode *vector based amplitude panning*, kurz *VBAP*, genannt [5].

## 2.2 Rundumklang

Rundum-Klang ist nach [1] [6] ein spezielles Mehrkanal-Wiedergabeverfahren. Laut [3] kann Rund-um-Klang als „Variation des Amplitude-Panning-Algorithmus“ [3] angesehen werden. Im Allgemeinen werden jedoch keine Panning-Rules zur Lautsprecher-Signalberechnung verwendet. Es ist jedoch möglich die Panning-Rules auf die spezielle Lautsprecherkonstellation anzuwenden. Die Höreignisbildung, die durch Schallfeldüberlagerung erzeugt wird, beschränkt sich dabei auf eine Ebene und bewegt sich somit innerhalb von zwei Dimensionen.

### 2.2.1 Wiedergabetechnik

Der populärste Abkömmling ist das sogenannte *3/2-Surround*- oder *5.1-Surround*-Verfahren. Es findet insbesondere bei Wiedergabe von Kinofilmen Verwendung und erreicht dadurch einen hohen Popularitätsgrad. Durch den vergleichsweise geringen finanziellen Aufwand ist das Verfahren daher auch für heimische Kino-Systeme geeignet. Abbildung 2.6 zeigt die geometrische Anordnung der Lautsprecher. Im Frontbereich ist des Weiteren das Fernsehgerät zu erkennen, um den Bezug zur Kinofilmwiedergabe herzustellen. In der Blickrichtung des Zuhörers sind zwei Lautsprecher in gleicher Konstellation wie bei 2-Kanal-Stereo-Systemen angeordnet, also bei  $\pm 30^\circ$ . Zusätzlich befindet sich ein weiterer Lautsprecher bei  $0^\circ$ . Dieser Kanal wird vor allem genutzt um die bei Kinofilmen häufig vorkommenden Dialoge auch außerhalb des zentralen *sweet spots* gleichermaßen im Zentrum der Video-Wiedergabe lokalisierbar zu halten. Bei ungefähr  $\pm 120^\circ$  befinden sich die sogenannten *Surround*-Kanäle, welche den Rundumklang ermöglichen und eine Räumlichkeit von Kinofilmen vermitteln. Die Lokalisation seitlicher Höreignisse ist dabei relativ ungenau. Dieses System als solches bezeichnet man dann als *3/2-Surround*-Verfahren. Erweitert man es um einen *Low Frequency Effect*- oder auch *Low Frequency Enhancement*-Kanal, kurz *LFE*, nennt man es *5.1-Surround*-Verfahren. Für den LFE-Kanal wird aufgrund der begrenzten Bandbreite zwischen 20Hz und 100Hz ein Subwoofer zur Wiedergabe verwendet, also ein Lautsprecher, der

vor allem für niedrige Frequenzen konzipiert ist.

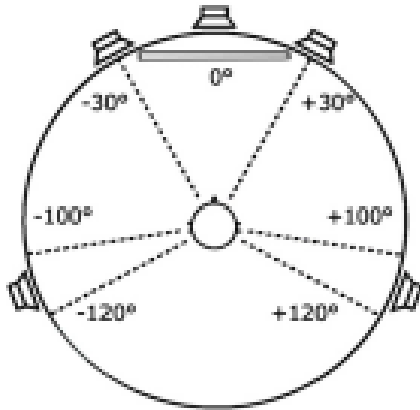


Abbildung 2.6:  
5.1-Surround Wiedergabeschema [3].

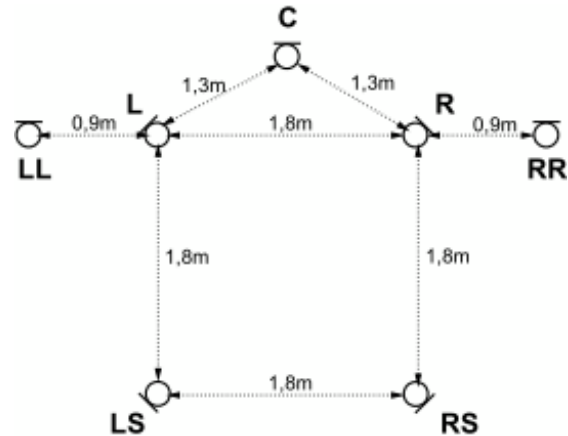


Abbildung 2.7:  
5.1-Surround Aufnahmeschema [1].

## 2.2.2 Aufnahmetechnik

Ein Beispiel für eine Aufnahmemethode für 5.1-Systeme ist der sogenannte *Fukada-Tree*, welcher in Abbildung 2.7 geometrisch dargestellt ist. Die Abstandswerte der hier abgebildeten Nierenmikrofone können dabei variieren. Es ist durchaus möglich weitere Mikrofone zur Umgebungsräuschaufnahme, wie z.B. Applaus, zu platzieren.

## 2.2.3 Beurteilung des Verfahrens

Insgesamt stellt Rundumklang eine recht einfach umsetzbare Möglichkeit zur räumlichen Wiedergabe innerhalb der horizontalen Ebene da. Die optimale Wahrnehmung ist jedoch wieder nur in einem *sweet spot* gegeben. Die Hörereignisse erscheinen außerdem nur in etwa in Lautsprecherentfernung. Des Weiteren gibt es Klangverfärbungen durch das überlagerte Schallfeld.

## 2.3 Ambisonics

Ambisonics ist ein weiteres Verfahren, welches auf dem Prinzip der Schallfeldüberlagerung basiert. In seiner ursprünglichen Form wurde es von Michael Gerzon in den 1970er Jahren entwickelt. Der Unterschied zu den bisher genannten Verfahren ist, dass sich die Wiedergabe nicht mehr auf eine bestimmte Anzahl Kanäle beschränkt. Die Lautsprechersignale können abhängig von der Lautsprecherkonstellation angepasst werden.

### 2.3.1 Beziehung zwischen Aufnahme- und Wiedergabetechnik

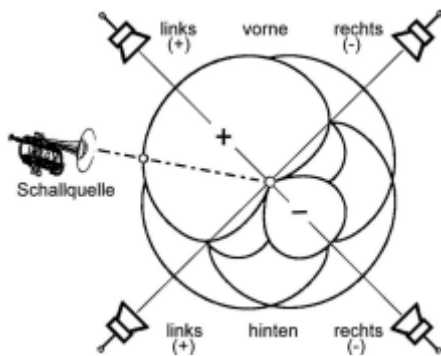


Abbildung 2.8: Zirkuläre Ambisonics Lautsprecheranordnung mit Deutung der Aufnahmeverhaltens der verwendeten Druckgradientenmikrofone, Vergleich Abbildung 2.4 [3].

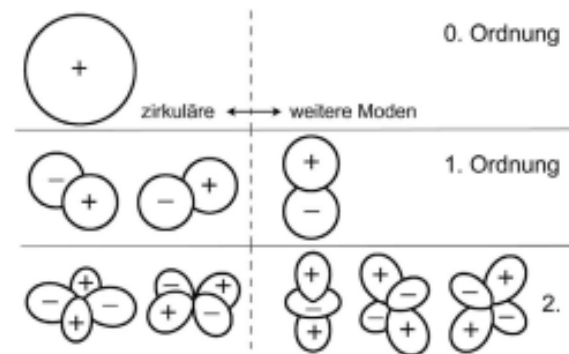


Abbildung 2.9: Vergleich zwischen zirkulären und sphärischen Moden [3].

Denkt man zurück an die Blumlein Mikrofonanordnung von Abbildung 2.4 und erweitert diese um ein zusätzliches Druckgradientenmikrofon in senkrechter Stellung und ein Kugelmikrofon so erhält man das sogenannte B-Format. Durch die Erweiterung der Druckgradientenmikrofone mit dem Kugelmikrofon erhält man eine Nieren oder Supernieren Mikrofon-Richtcharakteristik. Die Supernieren zeigen dann jeweils um  $90^\circ$  versetzt in die vier Hauptrichtungen der horizontalen Ebene. Diese zirkuläre Ambisonics Anordnung ist in Abbildung 2.8 dargestellt. Die Anregungsintensität der Superniere bestimmt bei der Wiedergabe den Pegel des jeweiligen Lautsprechers. Zur Berechnung der zirkulären Anordnung nehmen wir an, dass die Druckgradientenmikrofone jeweils par-

allel zur x- bzw. y-Achse angeordnet sind und sich im Koordinatenursprung befinden. Sie erzeugen dann Informationen über die Schallschnelle  $V_x$  bzw.  $V_y$  aus der jeweiligen Richtung. Das Kugelmikrofon, welches sich ebenfalls im Koordinatenursprung befindet, erzeugt Informationen über den Schalldruck  $P$ . Zur Berechnung eines Wiedergabesystems mit  $N$  Lautsprechern gilt

$$L_n = \frac{1}{N}(P + 2z_0(V_x \cos \varphi_n + V_y \sin \varphi_n)) \quad (2.5)$$

wobei  $L_n$  das Signal des  $n$ -ten Lautsprechers ist und  $\varphi_n$  dessen Abweichung von der x- bzw. y-Achse. Mit  $z_0 = c_s \cdot \rho$  ist hier die Schallkennimpedanz gemeint, welche sich aus Schallgeschwindigkeit  $c_s$  und Luftdichte  $\rho$  berechnen lässt und bei 20°C circa 413.6Ns/m<sup>3</sup> beträgt. Mit dieser Rechnung ist theoretisch eine exakte Rekonstruktion eines ebenen Schallfeldes im *sweet spot* möglich, in welchem es jedoch zu Abschattungseffekten durch den Kopf des Hörers kommt.

### 2.3.2 Mathematischer Hintergrund

Das Schallverhalten im Ursprung kann über sphärische harmonische Kugelflächenfunktionen (sphärische Harmonische) beschrieben werden. Im einfachsten Fall entspricht dies einer sphärischen Harmonischen nullter Ordnung. Man kann das Verhalten einer solchen mit einer atmenden Kugel vergleichen, also eine Kugel deren Radius sich je nach Schalldruck verändert. Dieses Verhalten kann mit Hilfe eines Kugelmikrofons ausreichend aufgezeichnet werden um später das Schallfeld in seiner ursprünglichen Form zu reproduzieren. Sphärisch Harmonische erster Ordnung kann man sich als starre, auf einem Streckenabschnitt oszillierende Kugel vorstellen. Zur korrekten Reproduktion des Schallfeldes sphärisch Harmonischer erster Ordnung können Druckgradientenmikrofone bei der Aufnahme genutzt werden. Das bereits erwähnte B-Format nutzt also sphärisch harmonische nullter und erster Ordnung zur Beschreibung des Schallfeldes im Ursprung bzw. der Position der Mikrofone. Die zweite Ordnung bildet bereits ein recht komplexes Schwingungsbild aus, bei dem sich die Kugel in vier um 90° versetzte Richtungen verzerren kann. In Abbildung 2.9 sind alle möglichen Moden von sphärischen bzw. zir-



kulären Harmonischen bis zur zweiten Ordnung aufgezeigt. Werden höhere Ordnungen zur Erzeugung des Schallfeldes genutzt, wird das Wiedergabeverfahren *Higher-order Ambisonics (HOA)* genannt. Die Entwicklung in sphärische Harmonische entspricht einer *Fourier*-Entwicklung einer periodischen Funktion.

Für eine zirkuläre Entwicklung kann folgendes für die periodische Funktion  $p_p$  angenommen werden

$$p_p(r, \delta, t) = e^{j\omega_0 t} \sum_{v=-\infty}^{\infty} a_v(\theta) \cdot b_v(r) \cdot e^{jv\delta} \quad (2.6)$$

um das Schallfeld in Entfernung  $r$  und Azimut-Winkel  $\delta$  zu berechnen. Die *Fourier*-Koeffizienten werden aus  $e^{j\omega_0 t} \cdot a_v(\theta) \cdot b_v(r)$  gebildet, indes der Winkel  $\theta$  die Richtung der ankommenden Welle angibt. Für  $a_v$  und  $b_v$  gilt folgender Zusammenhang,

$$a_v(\theta) = j^v e^{-jv\theta}, b_v(r) = J_v(\beta r), \quad (2.7)$$

wobei die Bessel-Funktion  $J_v$   $v$ -ter Ordnung vom Radius  $r$  und der räumlichen Frequenz  $\beta$  abhängt. Abbildung 2.10 zeigt die geometrische Deutung der hier genutzten Variablen  $r, \delta, \beta$  und  $\theta$ . Die Herleitung der Funktion kann in [7] nachgelesen werden.

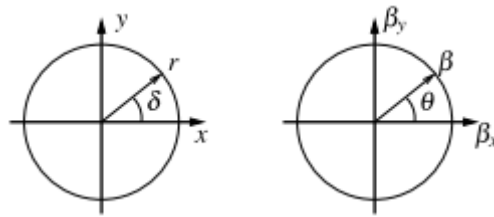


Abbildung 2.10: Geometrische Deutung in Polarkoordinaten für die Ortskoordinaten  $(x, y)$  und die räumliche Frequenz  $\beta$  [7].

In Abbildung 2.11 ist das Verhalten der *Fourier*-Koeffizienten  $a_v(\theta)$  und  $b_v(r)$  bis zur dritten Ordnung  $v$  aufgezeigt. Vom komplexen Wert  $a_v(\theta)$  ist in der ersten Zeile nur der Realteil  $\Re$  dargestellt, also  $\Re \{a_v(\theta)\} = \cos(-v\theta + v\frac{\pi}{2}) = \cos(v\theta - v\frac{\pi}{2})$ . Bei  $v = 0$  entspricht dieser Term einem konstanten Wert und besitzt somit keine Winkelabhängigkeit. Für höhere Ordnungen steigt die Variation der Gleichung innerhalb einer Kreisdrehung. Gerade in der Nähe des Ursprungs steigt dadurch die Auflösung deutlich.

Betrachtet man jedoch den Realteil des Produktes der Koeffizienten  $\Re\{a_v(\theta) \cdot b_v(r)\}$ , welches in der untersten Spalte der Grafik zu sehen ist, so wird die örtliche Auflösung durch das Verhalten der Bessel-Funktion für kleine Werte von  $r$  deutlich abgeschwächt. Bedenkt man, dass die Funktion  $b_v(r) = J_v(\beta r)$  jedoch zusätzlich von  $\beta$  abhängt, so steigt die örtliche Auflösung in Ursprungsnähe für hohe Frequenzen. Die Welle nullter Ordnung kann man als Punktquelle deuten, die erster Ordnung als Welle eines Dipols.

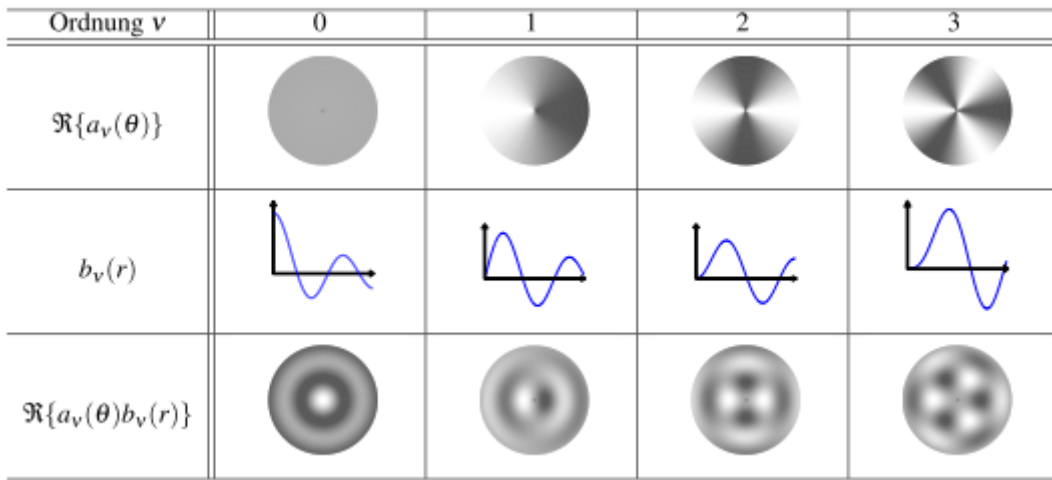


Abbildung 2.11: Berechnung der winkel- bzw. radiusabhängigen Anteile der *Fourier*-Koeffizienten bis zur dritten Ordnung. Oben: Realteil  $\Re\{\}$  der winkelabhängigen Funktion  $a_v(\theta)$ . Mitte: Radiusabhängige Bessel-Funktion  $b_v(r)$ . Unten: Realteil  $\Re\{\}$  des Produkts [7].

### 2.3.3 Beurteilung des Verfahrens

Das Ambisonic-Verfahren nach Gerzon hat sich als Wiedergabesystem nicht durchgesetzt, da das bereits erwähnte B-Format durch die spärliche Diskretisierung des Raumes eine zu geringe Aufzeichnung der sphärischen Harmonischen liefert. Dadurch beschränkt sich das volle Lokalisationspotential auf einen engen *sweet-spot* und ist in seitlicher Richtung zusätzlich relativ ungenau. Im Gegensatz dazu bietet HOA durch Erzeugung der sphärischen Harmonischen höherer Ordnung eine deutlich bessere Darstellung der Hörereignisse. Die Aufzeichnung des Schallfeldes ist dabei durch moderne

---

Mikrofone bis etwa zur vierten Ordnung möglich. Für die Wiedergabe synthetischer Quellen besteht jedoch keine Ordnungsbegrenzung. HOA wird daher oftmals bei der Wiedergabe elektronischer Musik genutzt.

## 2.4 Wellenfeldsynthese

Wellenfeldsynthese verfolgt genau *Higher Order Ambisonics* das Ziel ein Schallfeld innerhalb eines begrenzten Gebiets originalgetreu wiederzugeben. Während sich das Ambisonics Verfahren jedoch auf Pegelunterschiede zur Berechnung der Lautsprecher-signale beschränkt, spielen bei der Wellenfeldsynthese (WFS) Laufzeit- und Phasendifferenzen eine Rolle. Der mathematische Hintergrund der WFS unterliegt dem Kirchhoff-Helmholtz-Integral, welches später in diesem Abschnitt näher betrachtet werden soll. Eine bildliche Interpretation auf Wellenebene des Integrals bietet das Prinzip von Christian Huygens von 1690 welches Folgende besagt:

*Any point on a propagating wavefront can be taken as a point source for the production of spherical secondary waves.* Christiaan Huygens (1690) [1]

Also auf Deutsch übersetzt: Jeder Punkt einer sich ausbreitenden Wellenfront kann als Punktquelle angesehen werden, welche eine Kugelwelle erzeugt.

### 2.4.1 Beschreibung des Wellenfeldsynthese-Wiedergabeverfahrens



Abbildung 2.12: Deutung des Huygensschen Prinzips als Grundlage der Wellenfeldsynthese [1].

Mit Hilfe des Huygensschen Prinzips wird es ermöglicht jede beliebige Wellenfront durch sogenannte Elementarwellen nachzubilden. Ein Blick auf Abbildung 2.12 [1] hilft den Umfang der Aussage des Zitats zu begreifen. Das Huygenssche Prinzip bezieht sich jedoch auf infinitesimal kleine Abstände zwischen den erzeugten Kugelwellen. Un-

ter Beachtung dieser Restriktion, werden bereits Grenzen des Wiedergabeverfahrens aufgezeigt, da die Lautsprecher aufgrund ihrer Bauweise nur in gewissen Mindestabständen platziert werden können. Eine Welle kann mit der heutigen Lautsprechertechnik also niemals wirklich exakt rekonstruiert werden. Das menschliche Gehör benötigt allerdings keine exakte Rekonstruktion der Schallwelle, um ein Hörereignis richtig einordnen zu können. Als Daumenregel für den Abstand der Lautsprecher wird oftmals in etwa der Ohrabstand genommen werden. Näheres zu den Auswirkungen verschiedener Wellen auf das Schallfeld kann in [8] nachgelesen werden.

### 2.4.2 Mathematischer Hintergrund der Wellenfeldsynthese

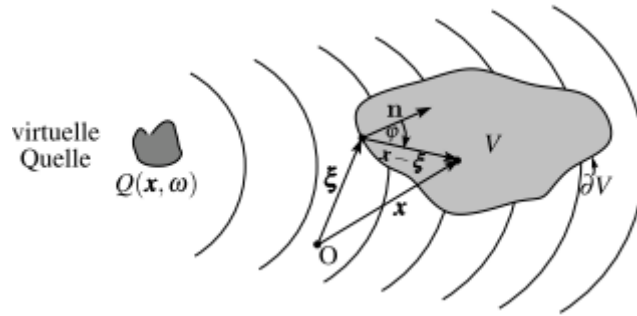


Abbildung 2.13: Geometrische Hilfe zur Berechnung des Schalldrucks  $P(\mathbf{x}, \omega)$  ausgelöst durch eine Quelle  $Q(\mathbf{x}, \omega)$  innerhalb eines Volumens  $V$ . Der Vektor  $\boldsymbol{\xi}$  zeigt auf den Rand  $\partial V$  des Volumens mit dem Normalenvektor  $\mathbf{n}$ , der Vektor  $\mathbf{x}$  auf den zu bestimmenden Schalldruck [7].

Nun soll die Herkunft des Kirchhoff-Helmholtz-Integrals näher betrachtet werden.

Ein Blick auf Abbildung 2.13 hilft die geometrischen Größen zu begreifen. Folgende Gleichung beschreibt den Schalldruck  $P$  innerhalb eines Volumens  $V$

$$P(\mathbf{x}) = \int_V G(\mathbf{x}|\boldsymbol{\xi}) P_0(\boldsymbol{\xi}) d\boldsymbol{\xi} + \int_{\partial V} \mathbf{n}^T (\nabla G(\mathbf{x}|\boldsymbol{\xi}) \cdot \mathbf{P}(\boldsymbol{\xi}) - G(\mathbf{x}|\boldsymbol{\xi}) \cdot \nabla P(\boldsymbol{\xi})) d\boldsymbol{\xi}, \quad (2.8)$$

wobei jede der Funktionen  $P$  und  $G$  zusätzlich zusätzlich von der Kreisfrequenz  $\omega$  abhängt. Mit  $G$  ist hier die *Greensche*-Funktion gemeint. Sie beschreibt den Einfluss einer Punktquelle am Rand eines Volumens an Position  $\boldsymbol{\xi}$  auf das Schallfeld innerhalb

des Volumens an Position  $\mathbf{x}$  und für welche in Freiraumausbreitung Folgendes gilt

$$G(\mathbf{x}|\boldsymbol{\xi}, \omega) = \frac{\exp(j\frac{\omega}{c}|\mathbf{x} - \boldsymbol{\xi}|)}{|\mathbf{x} - \boldsymbol{\xi}|}. \quad (2.9)$$

$G$  kann daher als Punktquelle bzw. als Monopol angesehen werden. Der Gradient der Greenschen Funktion  $\nabla G(\mathbf{x}, \omega)$  stellt einen Dipol dar. Somit werden, technisch gesehen, zwei Arten von Lautsprechern benötigt. Mehr zur Herleitung von Gleichung 2.9 und 2.8 kann in [7] nachgelesen werden. Betrachtet man Gleichung 2.8 genauer, fällt auf, dass sie zwei Schalldruck-Funktionen beinhaltet,  $P_0(\boldsymbol{\xi})$  und  $P(\boldsymbol{\xi})$ .  $P_0(\boldsymbol{\xi})$  bezeichnet hier den Schalldruck für Quellen, welche sich innerhalb des begrenzten Raumes befinden,  $P(\boldsymbol{\xi})$  den Schalldruck am Rand des Raumes, welcher durch Quellen außerhalb erzeugt wird. Somit lässt sich der Term in zwei Summanden aufteilen. Der erste Summand ist also verantwortlich für den Schalldruck, welcher durch Quellen innerhalb des Volumens verursacht wird und für den folgendes gilt

$$P(\mathbf{x}, \omega) = \int_V G(\mathbf{x}|\boldsymbol{\xi}, \omega) P_0(\boldsymbol{\xi}, \omega) d\boldsymbol{\xi}. \quad (2.10)$$

Nun wird der zweite Summand aus Gleichung 2.8 betrachtet, welcher durch Quellen ausserhalb des begrenzten Raumes entsteht. Multipliziert man den Normalenvektor  $\mathbf{n}^T$  des Volumenrandes  $\partial V$  in die Klammer und verknüpft ihn mit dem Gradienten  $\nabla$ , führt das Skalarprodukt der beiden zu einer Richtungsableitung  $\partial/\partial\mathbf{n}$ . Dieser Schritt führt uns zur Formulierung des bereits erwähnten Kirchhoff-Helmholtz-Integrals

$$P(\mathbf{x}, \omega) = \int_{\partial V} \left( \frac{\partial}{\partial\mathbf{n}} G(\mathbf{x}|\boldsymbol{\xi}, \omega) P(\boldsymbol{\xi}, \omega) - G(\mathbf{x}|\boldsymbol{\xi}, \omega) \frac{\partial}{\partial\mathbf{n}} P(\boldsymbol{\xi}, \omega) \right) d\boldsymbol{\xi}. \quad (2.11)$$

Somit kann theoretisch innerhalb eines begrenzten Bereichs der Schalldruck originalgetreu rekonstruiert werden. Diese Gleichung bildet die mathematische Grundlage des Huygensschen Prinzips und somit der Wellenfeldsynthese. Eine genauere Analyse hinsichtlich der Berechnung der Lautsprechersignale ist in [8] zu finden.

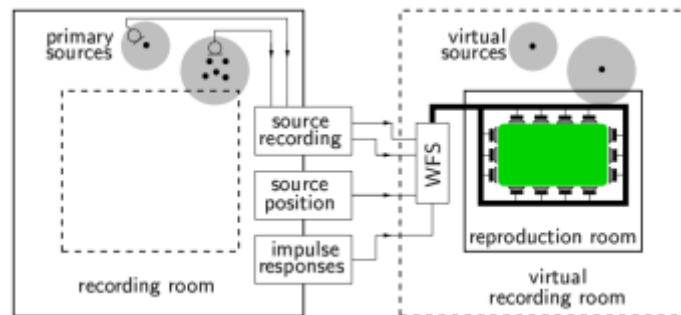


Abbildung 2.14: Beziehung zwischen Aufnahme und Wiedergabe im WFS-System [9].

### 2.4.3 Aufnahmetechnik

Im Gegensatz zu den bisher genannten räumlichen Wiedergabeverfahren, gibt es bei der Wellenfeldsynthese einen entscheidenden Unterschied im Bezug auf die Aufnahme der Audio-Szene. Die erwähnten Verfahren zeichnen die Hörereignisse mit Hilfe von speziellen Mikrofonen an einer zentralen Position auf. Die Mikrofone müssen dabei spezielle Richtcharakteristiken aufweisen und bei der Wiedergabe muss zusätzlich auf eine korrekte Dekodierung des Signals geachtet werden. Im Gegensatz dazu kann die Aufnahmetechnik bei der Wellenfeldsynthese als objektorientiert bezeichnet werden. In Abbildung 2.14 ist links der Aufnahmeraum mit mehreren Schallquellen zu sehen. Diese Quellen werden nun separat aufgezeichnet und ihre Positionen ermittelt. Zusätzlich können Raumimpulsantworten aufgezeichnet werden, um die akustischen Eigenschaften des Aufnahmeraums später rekonstruieren zu können. Diese Informationen werden nun an ein Wellenfeldsynthese-Wiedergabesystem übermittelt, wo letztendlich das ursprüngliche Schallfeld mit Hilfe von virtuellen Quellen wieder reproduziert wird.

### 2.4.4 Beurteilung des Verfahrens

Die Theorie bildet die Grundlage für die technische Umsetzung von WFS-Systemen. Es gibt jedoch Restriktionen hinsichtlich der Realisierbarkeit. Beispielsweise benötigt man wie bereits in diesem Abschnitt erwähnt zwei verschiedene Lautsprechertypen am Rand des Wiedergabebereichs. Monopol- und Dipolquellen haben unter Anbetracht des Normalenvektors in einem gewissen Bereich eine recht ähnliche Abstrahlcharak-

teristik. Oftmals begrenzt man sich bei der Wiedergabe auf eine Strecke, welche mit Lautsprechern bestückt, da Oberflächen einen deutlichen Mehraufwand erzeugen. Weiterhin ist eine kontinuierliche Lautsprecherverteilung offensichtlich unmöglich, weshalb es durch die Diskretisierung der Wellenfront zu räumlichen Aliasing-Effekten kommt, welche das menschliche Gehör mehr oder weniger wahrnimmt. Obwohl bei der technischen Umsetzung die Theorie nicht eins zu eins umgesetzt werden kann, liefern die Wellenfeldsynthese-Systeme eine relativ klare Wiedergabe von Hörereignissen.



## 2.5 Binaurale Wiedergabe

In diesem Kapitel wurden bisher einige Verfahren in Augenschein genommen, welche insbesondere eine mehr oder weniger exakte Reproduktion des Schallfeldes als Ziel verfolgen. Dies ist jedoch oftmals relativ aufwendig und liefert nicht immer die gewünschten Ergebnisse. Im Gegensatz dazu orientiert sich die binaurale Wiedergabe am menschlichen Gehör mit dessen charakteristischen psychoakustischen Eigenschaften. Das Signal wird dabei direkt am Kopf durch die Wiedergabe mit Kopfhörern erzeugt. Zunächst muss man verstehen wie das Gehör die am Kopf ankommenden Schallwellen verarbeitet.

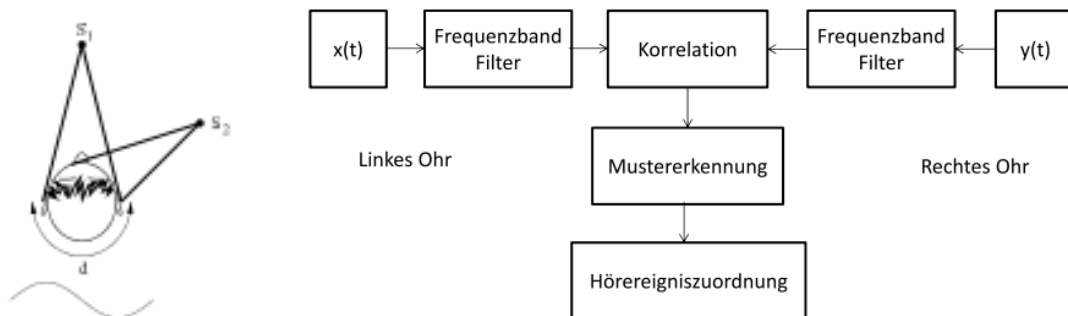


Abbildung 2.15: Links: Veranschaulichung zweier am Kopf ankommender Wellen [1]. Rechts: Verarbeitung der Signale durch das menschliche Gehör (adaptiert von [3]).

### 2.5.1 Beschreibung des menschlichen Gehörs

Es gibt nach [3] [1] zwei verschiedene Identifikationsmerkmale, welche das Gehör nutzt um ein Hörereignis einzuordnen. Diese Merkmale basieren auf Lord Raleighs Duplex-Theorie von 1907. Zum einen ist die Zeitverzögerung zu nennen, die zwischen den beiden Ohren entsteht, die sogenannte *Interaural Time Difference (ITD)*, zum anderen die Pegeldifferenz, welche durch Abschattung des Signals zwischen den beiden Ohren erzeugt wird, die sogenannte *Interaural Level Difference (ILD)*. Auf der linken Seite von Abbildung 2.15 ist ein Kopf zu erkennen, auf welchen zwei Schallwellen treffen. Angedeutet ist außerdem der halbe Kopfumfang  $d$ , welcher ungefähr 0.23 m entspricht. Nimmt man diesen Wert als Wellenlänge, kann daraus eine Frequenz  $f_0 = c_s/d \approx 1.5\text{kHz}$

berechnet werden, mit der Schallgeschwindigkeit  $c_s$ . Die Duplex-Theorie besagt nun, dass das menschliche Gehör für Frequenzen unterhalb dieser Schwelle  $f_0 = 1.5\text{kHz}$  vor allem die Laufzeitunterschiede zwischen den Signalen (ITD) nutzt, da nur geringe Pegelunterschiede zwischen den beiden Ohren entstehen. Frequenzen unter  $1.5\text{kHz}$  besitzen Wellenlängen von circa  $0.23\text{m}$  und befinden sich somit im Bereich des halben Kopfumfangs. Sie werden daher am Kopf gebrochen und somit nicht durch den Kopf abgeschattet. Daher ist der ILD für diese Frequenzen relativ gering. Für Frequenzen  $f > 1.5\text{kHz}$  nutzt das Gehör wiederum Pegelunterschiede (ILD), welche durch Abschattung des Kopfes entstehen.

Die wirkliche Lokalisation ist jedoch deutlich komplexer. Dabei führt das Gehirn eine frequenzabhängige Kreuzkorrelation der an den beiden Ohren ankommenden Signale durch. Das Signal wird somit in mehrere Frequenzbänder aufgeteilt und in diesen dann jeweils eine Korrelation durchgeführt. Danach wird das Ergebnis mit Mustern aus dem Gedächtnis verglichen und daraufhin eine Hörereigniszuordnung vorgenommen. Auf der rechten Seite von Abbildung 2.15 ist dieser Zusammenhang als Signalfussgraf dargestellt. Wie kann also genau dieses frequenzabhängige Signal originalgetreu an das Gehör übermittelt werden, damit dieses eine solche Lokalisation durchführen kann. In diesem Punkt kommen sogenannte *Head Related Transfer Functions (HRTFs)* ins Spiel. Dies sind durch Messungen ermittelte Übertragungsfunktionen. Die Übertragungsfunktionen stellen sozusagen den anatomischen Filter des Kopfes bzw. der Ohrmuschel dar.

### 2.5.2 Messung der HRTFs und Wiedergabetechnik

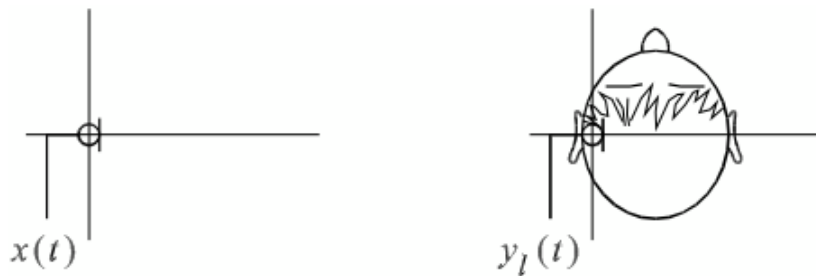


Abbildung 2.16: Messschema zur Bestimmung von HRTFs [1].

Abbildung 2.16 zeigt das Schema zur Aufnahme der Messungen wie in [1] beschrieben ist. Zum einen wird ein Zeitsignal in freier Position mit Hilfe eines Miniaturmikrofons gemessen. Danach wird an gleicher Position mit selbigem nur innerhalb eines Kopfes das Signal nochmals gemessen. Das im Freien gemessene Signal sieht man auf der Abbildung links und wird mit  $x(t)$  bezeichnet, das im Kopf im gemessene auf der rechten Seite der Abbildung mit  $y_l(t)$ . Nun bringt man die Fourier-Transformierten der beiden Signale in Beziehung

$$HRTF_{l/r}(j\omega) = \frac{Y_{l/r}(j\omega)}{X(j\omega)} \quad (2.12)$$

und erhalten somit die Head Related Transfer Function des linken bzw. rechten Ohres.



Abbildung 2.17: Abbildung eines Kunstkopfes zur Messung von HRTFs [10].

Ein Problem liegt darin, dass nicht jeder Kopf bzw. jedes Ohr gleich geformt ist und somit eigentlich HRTFs für jede Person speziell gemessen werden müssten. Um dies zu umgehen mittelt man die Werte über mehrere Messungen mit verschiedenen Per-

sonen. Eine weitere Möglichkeit stellt der Einsatz eines genormten Kunstkopfes dar, welcher eine genormte Kopfform besitzt. Ein solcher ist in Abbildung 2.17 dargestellt. Für die Messung der HRTFs sollte immer ein hallarmer Raum genutzt, um mögliche Echo-Effekte zu reduzieren. Die Messung wird in verschiedenen Winkeln durchgeführt, wobei die räumliche Auflösung der Wiedergabe von der Anzahl der Messpunkte abhängt. Die Entfernung hat lediglich Einfluss auf die Lautstärke des Signals. Oftmals ist es aber auch gewollt die Akustik des Raumes in die Messung zu integrieren. In diesem Fall wird die Impulsantwort der gemessenen Übertragungsfunktion *Binaural Room Impulse Response (BRIR)* genannt. Die Wiedergabe erfolgt ähnlich wie bei der Wellenfeldsynthese objektorientiert. Die hallarme Aufnahme einer Schallquelle muss je nach Position lediglich mit der jeweiligen Impulsantwort gefaltet werden um das Signal am rechten bzw. linken Kopfhörer zu erhalten. Das Schema für die Berechnung der Kopfhörersignale ist in Abbildung 4.1(c) dargestellt.

### 2.5.3 Beurteilung des Verfahrens

Die binaurale Wiedergabe liefert allgemein recht gute Ergebnisse zur Darstellung von Hörereignissen. Auch hinsichtlich des Aufwandes bietet sie einige Vorteile. Darüber hinaus ist die Erforschung der binauralen Technik auch für andere Themengebiete interessant, wie zum Beispiel Audio-Codierung oder Hörgeräteentwicklung. In Bezug auf die räumliche Wiedergabe ist es allerdings nicht immer gewollt Kopfhörer zu tragen.

## Kapitel 3

# Spiegelquellen Theorie

Zur Beschreibung von Raumakustik wird oftmals die Spiegelquellen Theorie genutzt. Sie wird genutzt um Echo-Effekte zu beschreiben. Das Grundprinzip folgt den gleichen Regeln, welche auch bei der Elektrodynamik oder Optik angewendet werden. Bei der Akustik wird von glatten Wänden ausgegangen und diffuse Reflektionen damit vernachlässigt. Am besten kann der Theorie-Ansatz anhand einer Zeichnung erklärt werden.

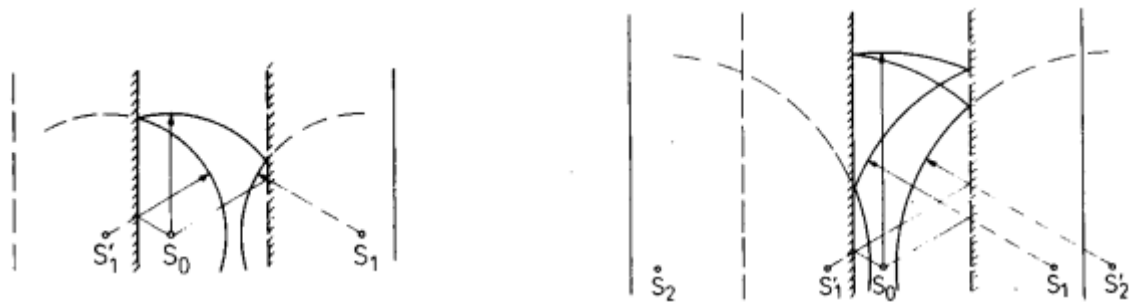


Abbildung 3.1: Darstellung des Spiegelquellenprinzips[11]

Auf der linken Seite von Abbildung 3.1 ist eine Punktquelle  $S_0$  zu sehen, welche sich zwischen zwei Wänden befindet. Die erzeugte Schallwelle breitet sich nun radial in alle Richtungen aus und trifft dabei auf die beiden Wände, an denen sie reflektiert. Nun kann man die erzeugte Reflektion mit Hilfe einer virtuellen Spiegelquelle erster Ordnung beschreiben, welche sich im gleichen Abstand zur Wand hinter dieser befindet.

Dieses Prinzip wendet man für beide Wände an. Die Welle der Spiegelquelle reflektiert nun zusätzlich an der gegenüber liegenden Wand und erzeugt somit wiederum eine Spiegelquelle zweiter Ordnung. Abbildung 3.1 zeigt auf der rechten Seite die Quellen erster und zweiter Ordnung.

Dieses Prinzip kann nun auch für vier Wände angewendet werden, wobei nun an allen vier Wänden gespiegelt werden muss. In Abbildung 3.2 sind die Spiegelquellen eines rechteckigen Raumes mit vier Wänden bis zur zwölften Ordnung aufgetragen. An dieser Zeichnung lässt sich das Prinzip deutlich erkennen.

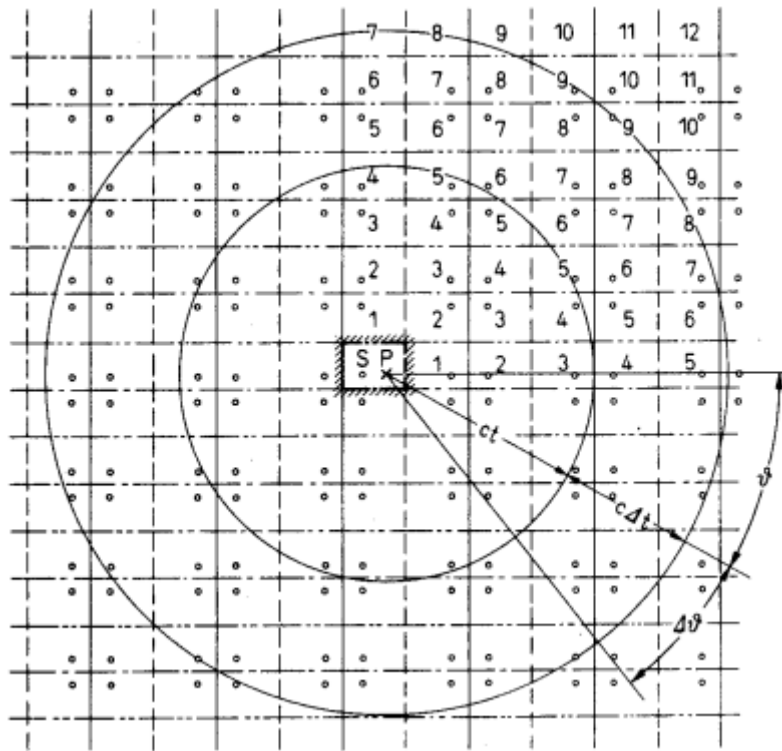


Abbildung 3.2: Konstruktion der Spiegelquellen für einen rechteckigen Raum[11]

Rein theoretisch kann man die Spiegelquellen-Theorie auf komplexere Wandformationen anwenden. Allerdings wird der Algorithmus zum Setzen der Spiegelquellen so kompliziert, dass es sinnvoller ist den Raumhall mit Hilfe von Raumimpulsantworten zu beschreiben. Die Beschreibung eines Algorithmus zur Berechnung von Spiegelquellen komplexere Wandformationen kann in [12] nachgelesen werden.

## Kapitel 4

# Objektbasierte Wiedergabe von Audio Szenen

Wie bereits im Abschnitt für Wellenfeldsynthese 2.4 und binaurale Wiedergabe 2.5 erwähnt, beziehen sich diese Verfahren auf eine objektorientierte Wiedergabe. Allerdings kann die Objektorientierung auch auf die anderen Wiedergabeverfahren angewendet werden. Bei objektorientierter Wiedergabe beschränkt man sich somit nicht mehr auf ein spezielles Wiedergabeverfahren. Die Lautsprecher-signale müssen abhängig von der Schallquelle und deren Position am Wiedergabeort selber berechnet werden. In diesem Kapitel sollen Möglichkeiten aufgezeigt werden, mit welchen eine solche Objektorientierung angewendet werden kann.

### 4.1 Soundscape Renderer

Der Soundscape Renderer (SSR) ist ein frei verfügbares Programm zur Wiedergabe objektorientierter Audio-Szenen im zweidimensionalen Bereich. Es wurde von den Quality and Usability Labs des Instituts für Softwaretechnik und theoretische Informatik der Technischen Universität Berlin entwickelt, um Audio-Szenen unabhängig vom Wiedergabesystem möglichst originalgetreu darzustellen. Die Audio-Szene wird dabei in einer sogenannten Szenen Beschreibung gespeichert, welche der SSR interpretiert und die

Szene modelliert. Näheres zur Szenenbeschreibung wird im nächsten Abschnitt 4.2 erläutert.

Der Soundscape Renderer ist in der Programmiersprache C++ geschrieben. Diese bietet gute Möglichkeiten zur Objektorientierung und diverse Funktionsbibliotheken. Dabei ist es trotzdem möglich einen maschinennahen und somit schnellen Programmcode zu erzeugen. Durch die Leistungsfähigkeit des Programms, ist es ohne Probleme möglich die Signale für jeden Lautsprecherkanal in Echtzeit zu berechnen. Zur Zeit stehen Versionen des SSR für Mac OS X und Linux zur Verfügung, wobei es möglich ist die Linux Version mit einem Smart-Phone über die integrierte Netzwerkschnittstelle des SSR zu steuern. Diese Netzwerkschnittstelle ermöglicht außerdem diverse Befehle an den SSR zu senden, die dieser in Echtzeit verarbeitet. Als Beispiel für die möglichen Befehle ist hier Quellen-Erzeugung, Positionsveränderung einer existierenden Quelle oder Lautstärkeveränderung zu nennen. Weiter Befehle können in [13] nachgelesen werden. Der SSR selber erzeugt keine Audio-Signale, er verarbeitet sie nur. Standardmäßig nutzt der SSR das Kommandozeilen-Programm Ecasound zur Audio-Datenstrom Erzeugung. Es ist aber auch möglich ein durch ein anderes Wiedergabeprogramm erzeugtes Signal in den SSR zur Verarbeitung umzuleiten. Die Steuerung Datenströme außerhalb des SSR übernimmt ein sogenannter JACK-Audio-Server. Eine genauere Beschreibung zur Verarbeitung der Signale außerhalb des SSR ist in [9] zu finden. Der Umgang mit den für den SSR benötigten Verarbeitungs-Programmen wird kurz im Anhang A.1 beschrieben oder kann in der Anleitung des SSR [13] ausführlich nachgelesen werden. Es ist bereits möglich den SSR auf alle im Kapitel 2 erwähnten Wiedergabe-Methoden anzuwenden. Abbildung 4.1 zeigt jeweils die Verarbeitungsschritte der Signale für die Wiedergabesysteme Wellenfeldsynthese, Ambisonic Amplitude Panning bzw. HOA, binaurale Wiedergabe und Vector Base Amplitude Panning. Die englischen Fachbegriffe stehen in der Bildunterschrift in deutscher Übersetzung.

Bei der Wellenfeldsynthese aus Abbildung 4.1(a) wird das Quellensignal zunächst durch einen Vorfilter spektral korrigiert. Dieser Vorfilter entsteht durch Übertragung des Kirchhoff-Helmholtz-Integrals auf ein ein-dimensionales Lautsprecher-Array. Mehr zu



diesem Vorfilter kann in [8] gefunden werden. Abhängig vom Einfallswinkel und Distanz zum Lautsprecher wird das Signal gewichtet und mit einer Verzögerung versehen, um dann unter Beachtung weiterer Quellen an den Lautsprecher gesendet zu werden. Die Berechnung des Ambisonic Lautsprechersignals aus Abbildung 4.1(b) erfolgt durch Gewichtung abhängig vom Einfallswinkel und der Distanz der Quelle zum Lautsprecher. Dabei fließen auch jeweils die Signale der anderen Quellen mit ein.

Bei der binauralen Lautsprechersignaleberechnung aus Abbildung 4.1(c) wird das Signal zunächst abhängig von der Distanz zum Zuhörer gewichtet. Danach erfolgt eine Filterung mit der jeweiligen HRTF aus einer Datenbank abhängig vom Einfallswinkel für das linke bzw. rechte Ohr. Hier können zusätzlich die raumakustischen Informationen aus den BRIRs einfließen.

Die VBAP-Signal-Berechnung aus Abbildung 4.1(d) wird zunächst abhängig von der Distanz zwischen Quelle und Lautsprecher gewichtet. Eine Quelle erzeugt abhängig vom Einfallswinkel immer nur Signale für zwei Lautsprecher.

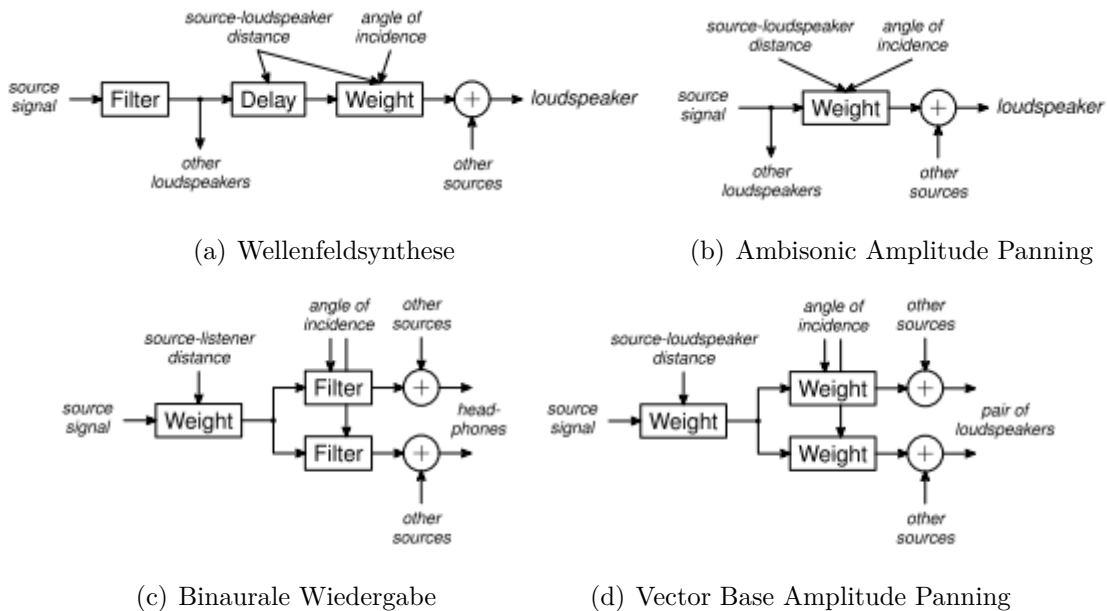


Abbildung 4.1: Flussgraphen zur Bestimmung der Lautsprecher bzw. Kopfhörersignale. Source signal - Quellensignal; angle of incidence - Einfallswinkel; delay - Verzögerung; weight - Gewichtung; [14].

Die grafische Oberfläche des SSR bietet einige Möglichkeiten eine 2-dimensionale Audio-Szene zu gestalten. Dabei ist die Bedienung sehr intuitiv über die Maus zu steuern. Eine weitreichende Beschreibung zur Bedienung des SSR ist im Abschnitt A.1 zu finden. Abbildung 4.2 zeigt die grafische Oberfläche des SSR, wobei in diesem Fall ein Wellenfeldsynthese-System als Wiedergabemethode genutzt wird. Gut zu erkennen ist der Einfluss der Punktquelle auf die Lautsprecher. Vergleicht man diesen Einfluss nun mit dem Einfluss einer planaren Welle aus Abbildung 4.3, so fällt auf, dass bei der Punktquelle weniger Lautsprecher für das Schallfeld genutzt werden. Denkt man zurück an Abbildung 2.13 und Gleichung 2.8, so wird klar, dass bei einer Punktquelle die Orthogonalität zwischen dem Normalenvektor der radialen Welle und dem Normalenvektor des Volumenrandes bereits bei geringen Abständen zum Lautsprecher erreicht wird. Entfernt man die Punktquelle vom Lautsprecherring, so nähert sich diese in einer gewissen Entfernung einer ebenen Welle an und es werden mehr Lautsprecher zur Wiedergabe genutzt. Dahingegen nutzt die ebene Welle immer in etwa einen Lautsprecherhalbkreis zur Wiedergabe. Die Orthogonalität hat auf ihre Lautsprechersignalberechnung somit einen konstanten Einfluss unabhängig von der Distanz.

Abschließend kann man sagen, dass der SSR eine gute Möglichkeit zur Darstellung zwei-dimensionaler Audio-Szenen ist. Die in Echtzeit berechneten Signale führen zur gewünschten Hörereignis-Wahrnehmung im Rahmen des genutzten Wiedergabesystems.

## 4.2 Audio-Szenen Beschreibungsformat

Der Soundscape Renderer bietet die Möglichkeit zweidimensionale Audio-Szenen darzustellen. Nun stellt sich noch die Frage wie eine solche objektorientiert Szene beschrieben werden kann. Der SSR nutzt zur Beschreibung von Audio-Szenen eine auf XML (eXtensible Markup Language)-Code basierende Syntax, das sogenannte *Audio Scene Description Format (ASDF)*. Die Syntax des XML-Formats bietet die Möglichkeit in gut begreifbarer Weise sämtliche Parameter einer Audioszene fest zu legen, ohne dass besondere Kenntnisse in der Verarbeitung der Lautsprechersignale von Nöten

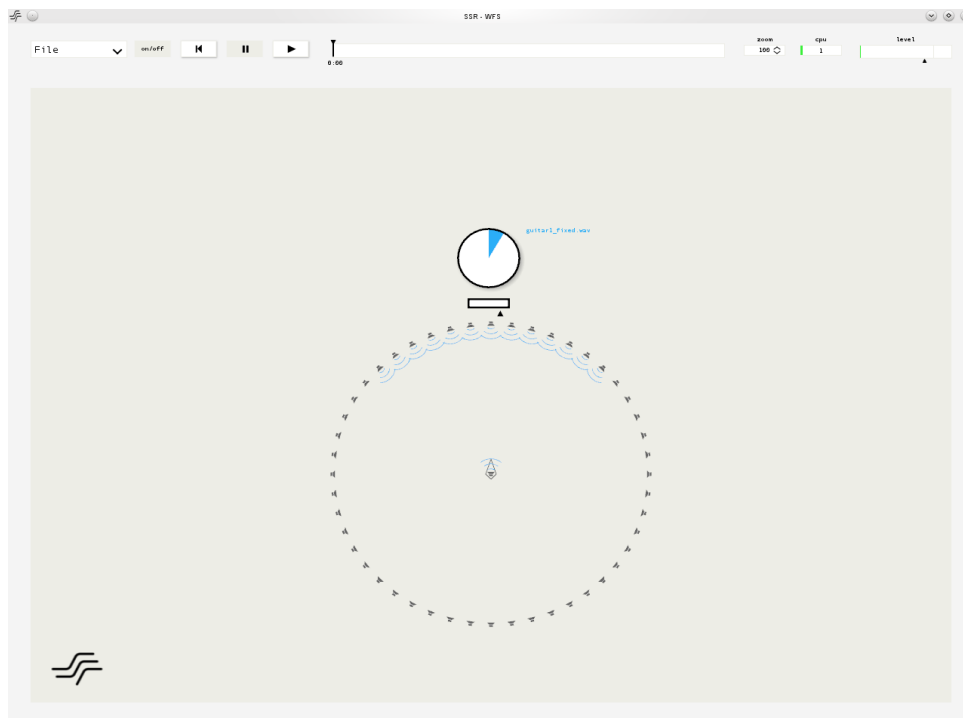


Abbildung 4.2: Einfluss einer Punktquelle auf die Lautsprechersignale eines Wellenfeldsynthese-Systems im Soundscape Renderer.

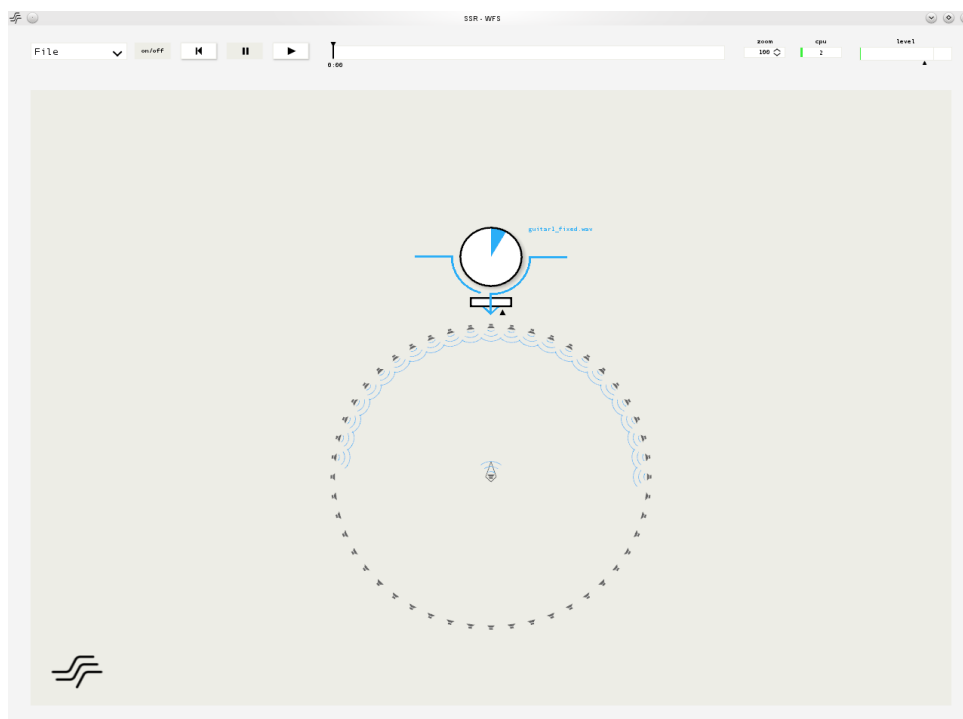


Abbildung 4.3: Einfluss einer planaren Welle auf die Lautsprechersignale eines Wellenfeldsynthese-Systems im Soundscape Renderer.

sind. Somit wird es möglich Audio-Szenen zwischen verschiedenen Szenen auszutauschen und dabei Korrekturen in den Parametern ohne fachspezifisches Programmier- und Signalverarbeitungswissen vorzunehmen. Nachfolgend ist ein Listing 4.1 des ASDF dargestellt.

Listing 4.1: Auf XML basierender Beispielcode der ASDF für eine Audio-Szene.

```
1 <?xml version="1.0" encoding="utf-8"?>
2 <asdf
3     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
4     xsi:noNamespaceSchemaLocation="asdf.xsd"
5     version="0.1">
6
7     <header>
8         <name>Beispiel-Szene</name>
9         <description>
10             Beispiel-Szene
11         </description>
12     </header>
13
14     <scene_setup>
15
16         <source name="Punktquelle">
17             <file>audio/music.wav</file>
18             <position x="-2" y="1.5"/>
19         </source>
20
21         <source name="Planare_Welle" model="plane">
22             <file channel="2">audio/sound.wav</file>
23             <position x="2" y="2"/>
24             <orientation azimuth="-135"/>
25         </source>
26
27     </scene_setup>
28 </asdf>
```

Die Szene beschreibt in diesem Fall eine Punktquelle mit kartesischen Koordinaten. Bei der Beschreibung der planaren Welle ist zusätzlich ein Orientierungswinkel angegeben, in welchen die Welle abstrahlt.

Neben der Beschreibung der Szene nutzt der SSR das ASDF zusätzlich zur Interpretation der Lautsprecher-Konstellation des Wiedergabe-Systems. Nachfolgend ist ein Beispielcode 4.2 für ein Reproduktions-Setup mit einem Lautsprecher-Ring aus 48 Lautsprechern dargestellt.

Listing 4.2: Reproduktions-Setup im ASD-Format für einen Lautsprecherring.

```
1 <?xml version="1.0" encoding="utf-8"?>
2 <?xml-stylesheet type="text/xsl" href="asdf2html.xsl"?>
3 <asdf xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
4     xsi:noNamespaceSchemaLocation="asdf.xsd">
5   <header>
6     <name>Lautsprecher-Ring</name>
7   </header>
8
9   <reproduction_setup>
10     <circular_array number="48">
11       <first>
12         <position x="1.5" y="0"/>
13         <orientation azimuth="-180"/>
14       </first>
15     </circular_array>
16   </reproduction_setup>
17 </asdf>
```

Die Beschreibung von Audio-Szenen im ASD-Format bietet noch weitere Möglichkeiten, welche genauer in der Anleitung des SSR [13] beschrieben sind.

# Kapitel 5

## Vohandene

## Wellenfeldsynthese-Demos am

## Lehrstuhl für

## Multimediakommunikation

Der Grundgedanke der praktischen Ausführung dieser Arbeit ist die vorhandenen Wellenfeldsynthese-Hörbeispiele des Lehrstuhls für Multimediakommunikation und Signalverarbeitung( LMS) der Friedrich-Alexander Universität Erlangen-Nürnberg für ein sich im Bau befindliches Wellenfeldsynthese-System verfügbar zu machen. Im Gegensatz zum alten WFS-System, welches mit einem kreisrunden 48-Lautsprecher-Array ausgestattet ist, besitzt das neue System deutlich mehr Lautsprecher und ist rechteckig geformt. Eine Übertragbarkeit der einzelnen Lautsprecher-Signale ist somit nicht direkt möglich. Die Signale müssten somit für jede der vorhandenen Demos neu berechnet werden.

Ein Blick auf Tabelle 5.1 ermöglicht eine Übersicht über die vorhandenen Demos des LMS. Für alle Projekte sind für das aktuelle Wellenfeldsynthese-System berechnete 49-Kanal wav-Dateien vorhanden, wobei der 49-zigste Kanal an den Subwoofer wei-

tergeleitet wird. Besonderes Augenmerk soll auf die vierte Spalte der Tabelle gelegt werden. Diese kennzeichnet die Projekte, bei welchen die einzelnen Quellen vorhanden sind, die abhängig von deren Position zu Berechnung der Lautsprecher-Signale genutzt wurden. Es fällt auf, dass ein Großteil der Projekte somit nicht objektorientiert gespeichert wurde und somit keine Quellensignale vorhanden sind.

Projekt	Name	Kategorie	Quellen vorhanden
Eagles	Eagles	Super-Audio 5.1	✓
Eichenau	Eichenau	Konzert	✓
Metropolis	Metropolis	vertonter Film	×
Project	Project	Musik	✓
Simple Sources	guitar1	Musik	✓
	chor	Musik	×
	guitarvoiceold	Musik	×
	pop	Musik	×
	rock	Musik	×
	sine	Ton	×
	siren	Ton	×
	voice	Musik	✓
Sting	Sting	Super-Audio 5.1	×
Stufen	Stufen	Video	×
Wachphase	Wachphase	Hörspiel	×
Demo M. Fischer	Demo M. Fischer	Band Simulation	×

Tabelle 5.1: Übersicht über die vorhandenen Hörbeispiele des aktuellen WFS-Systems des LMS. Eine Beschreibung der einzelnen Projekte ist im Text zu finden.

**Eagles** Das 'Eagles' Projekt der gleichnamigen Band beinhaltet neben der gerenderten 49-Kanal wav-Datei zusätzlich alle Kanäle( 6) des 5.1-Systems. Aufgenommen



wurde hier das berühmte Stück 'Hotel California' bei einem Live-Konzert im 5.1 Format. . . .

**Eichenau** Im 'Eichenau' Projekt wurde ein Konzert in einer Kirche in Eichenau aufgenommen. Dazu wurden die drei Quellen von Chor und Orchester als Monoquelle aufgezeichnet. Um den räumlichen Eindruck der Kirche reproduzieren zu können, wurden in regelmäßigen Abständen verteilt auf einem Kreis zusätzlich die Impulsantworten der Kirche gemessen. Danach wurden aus den vorhandenen Messungen die Impulsantworten für acht Positionen berechnet. Durch Faltung der Quellen mit den Impulsantworten kann dadurch die Akustik des Raumes reproduziert werden.

**Metropolis** ist ein deutscher Stummfilm aus den 30er Jahren, aus welchem eine Labor-Szene mit im Raum verteilten Geräuschen untermalt wurde. In diesem Fall sind keine Quellen der einzelnen Geräusche vorhanden.

**Project** Beim Projekt 'project' handelt es sich um ein vertontes Musikstück. Quellen der Instrumente oder des Gesangs sind hier nicht vorhanden.

**Simple Sources** Das Projekt 'simple sources' beinhaltet diverse Demos zur Verdeutlichung von zwei-dimensionalen Hörereignissen. Es sind nur einzelne Quellen einer Gitarre und eines Gesangs vorhanden. Die Demos sind jeweils als statische und dynamische Audio-Szene vorhanden. Neben der bereits erwähnten Gitarre und des Gesangs existieren eine Kombination aus beiden und ein Chor, Popsong, Rocksong, ein sich bewegender Sinus-Ton und eine sich bewegende Sirene. Die Bewegung stellt jeweils einen Kreis da.

**Sting** Eine weiteres Projekt im 5.1 Format das Lieder 'Desert Rose' des Künstlers 'Sting'. In diesem Fall existiert jedoch nur die gerenderte wav-Datei. Die Vertonung dieses Stückes ist in meinen Augen sehr imposant.

**Stufen** Das 'Stufen' Projekt ist ein künstlerisch vertontes und bildlich bearbeitetes

Video einer Rolltreppe. Es wurde 2005 auf dem Hörkunstfestival in Erlangen aufgeführt. Quellen der genutzten Geräusche sind auch hier nicht vorhanden.

**Demo M. Fischer** Ein besonderes Projekt ist die Band-Simulation von Michael Fischer. Die in einer Ebene dargestellten Instrumente können über Midi-Befehle gesteuert werden. Zusätzlich wird die Befehlseingabe über eine Visualisierung auf einer grafischen Oberfläche kenntlich gemacht. Eine Demonstration der Möglichkeiten ist in einem Musik-Video zusammengefasst.

Das nachfolgende Kapitel erläutert die durchgeführten Lösungsansätze zur Portierung der vorhandenen Hörbeispiele auf das neu gebaute WFS-System, wobei durch Einsatz des SSRs die Portierung auch auf andere Wiedergabeverfahren möglich ist.

## Kapitel 6

# Audio-Szenen-Portierung durch Nutzung des SSR

Der SSR bietet bereits viele Möglichkeiten um Audio-Szenen nach objektorientiert erfolgter Aufnahme reproduzieren zu können. Zur Beschreibung der Szene wird wie bereits erwähnt das ASD-Format genutzt. Im Moment können in diesem bereits Merkmale wie Position, Lautstärke und Abstrahlverhalten einer Quelle festgehalten werden. Jedoch ist es oftmals erwünscht weitere Merkmale einer Audio-Szene festzuhalten, wie zum Beispiel Quellenbewegung oder raumakustisches Verhalten. Ansätze zur Umsetzung dieser Szenen-Inhalte, dessen Beschreibung im Stil des XML-Formats erfolgt, werden im Verlauf dieses Kapitel beschrieben. Ungeachtet dessen ermöglicht das ASD-Format bereits einige der vorhandenen Demos des LMS in den SSR zu integrieren und somit portierbar auf andere Wiedergabesysteme zu machen, insbesondere dem im Bau befindlichen neuen WFS-System. Um die Portierung aller Demos zu gewährleisten und die neuen Szenen-Inhalte realisieren, mussten einige Zusatzmodule generiert werden. Zur Erzeugung der Zusatzmodule ist die Skript-Sprache Python gut geeignet. Python besitzt neben der Möglichkeit zur Objektorientierung, eine Funktion zum erkennen regulären Ausdrücken (Zeichen-Mustererkennung) zum Auslesen des ASD-Formats, eine ergiebige Auswahl an mathematischen Funktionen zur Berechnung der Quellenpositio-

nen und insbesondere die Möglichkeit Befehle an die Netzwerkschnittstelle des SSR zu senden.

In Abbildung 6.1 ist ein Beispiel für eine Portierung mit den gegebenen Möglichkeiten des ASD-Formats zu sehen. Es handelt sich hierbei um das bereits erwähnte Konzert in der katholischen Kirche von Eichenau. Die drei Quellen des Orchesters bzw. Chors sind jeweils hervorgehoben. Um diese befinden sich in einem Kreis angeordnet die mit der Impulsantwort gefalteten Hallquellen als planare Welle. Ein kleines Python Skript bietet die Möglichkeit den Hall jeweils ein oder auszuschalten und ermöglicht somit die Auswirkungen des Raumhalls im direkten Vergleich. Der Umgang mit dem Skript, wie auch der Umgang mit allen anderen Skripten ist in Anhang A.1 beschrieben.

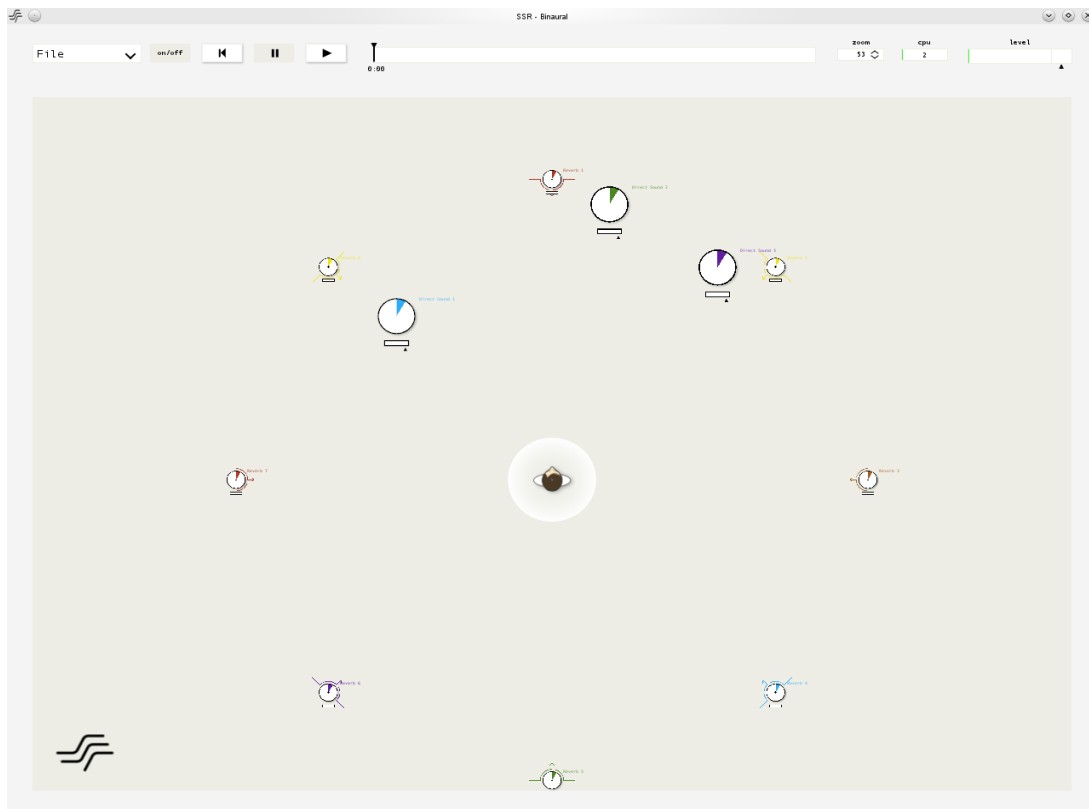


Abbildung 6.1: Audio-Szene eines Konzerts in der katholischen Kirche in Eichenau: Die Quellen des Orchesters bzw. Chors sind hervorgehoben, die planaren Hallquellen in einem Kreis angeordnet.

Im nachfolgenden Abschnitt soll die Portierung von kanal-basierenden Wiedergabe-

verfahren betrachtet werden. Darauf folgen Ansätze zur Portierung der akustischen Gegebenheiten des Aufnahmeraums und zur Portierung von Schallquellenbewegungen.

## 6.1 Portierung Kanalbasierter Audio-Dateien

Im Kapitel 2 wurden diverse Verfahren zur räumlichen Wiedergabe aufgezählt. Insbesondere Amplituden-Differenz Stereophonie und 5.1 Surround stellen an sich kanalbasierte Verfahren dar. Die Lautsprechersignale werden also nach der Aufnahme speziell für das jeweilige Wiedergabesystem einmal berechnet, um sie danach gleichermaßen wiederzugeben. Wie Tabelle 5.1 zeigt, fehlt auch einigen Wellenfeldsynthese-Demos des LMS die Objektorientierung der Audio-Szene und sie sind somit teilweise kanalbasierend. Da der Soundscape Renderer jedoch objektorientiert arbeitet, müssen auch diese portierbar gemacht werden. Dieser Abschnitt zeigt Lösungsmöglichkeiten auf, um diese kanalbasierenden Verfahren mittels des SSR auf beliebigen Wiedergabesystemen verfügbar zu machen.

Die Portierung einer Stereo-Datei ist im SSR nicht sehr aufwendig. Wird eine zwei Kanal wav-Datei im SSR geöffnet, so positioniert dieser die beiden Quellen (Kanäle) in der für Amplituden-Differenz Stereophonie üblichen Position als planare Welle.

Für 5.1 Surround Dateien bedarf es jedoch einer Szenen-Beschreibung, um die Quellen am richtigen Ort und mit der richtigen Wellenform zu positionieren. In Abbildung 6.2 ist die Audio-Szenen-Umsetzung einer im 5.1-Format vorhandenen wav-Datei dargestellt. Die winkelabhängige Positionierung der Quellen orientiert sich an der Lautsprecheranordnung von 5.1-Surround Systemen, wobei der Radius willkürlich gewählt wurde, da er nur Einfluss auf die Lautstärke hat. Zu erwähnen bleibt, dass die vier für den Raumklang verantwortlichen Kanäle als planare Quelle gewählt wurden, um eventuelle Panning-Effekte nicht zu verzehren. Der zentrierte für Dialoge zuständige Lautsprecher breitet sich radial aus, damit dieser im Wiedergabebereich für alle gleichmäßige Verständlichkeit und Orientierung ermöglicht. Die Position des LFE-Kanal hat auch keinen Einfluss auf die räumliche Wahrnehmung und kann somit beliebig gewählt

werden, in diesem Fall bei  $\pi/4$ .

Für die Umsetzung der für das WFS-System des LMS konzipierten 48-Kanal-wav-Dateien wurde ebenfalls die ursprüngliche Lautsprecherkonstellation zur Positionierung genutzt. Nach dem Huygenschen Prinzip entstehen hierdurch keine Nachteile hinsichtlich der Wellennachbildung. Die einzelnen Kanäle bzw. Punktquellen werden gleichmäßig auf einem Kreis unter Beachtung der ursprünglichen Reihenfolge verteilt wie in Abbildung 6.3 zu sehen ist. Zu erwähnen ist, dass der 49-zigste Kanal, welcher den Subwoofer bestimmt ist, weggelassen wurde, da dieser keine besonderen räumlichen Darstellungsmerkmale besitzt und der Soundscape Renderer die Signale aller vorhandenen Quellen zur Subwoofer-Signal-Berechnung nutzt.

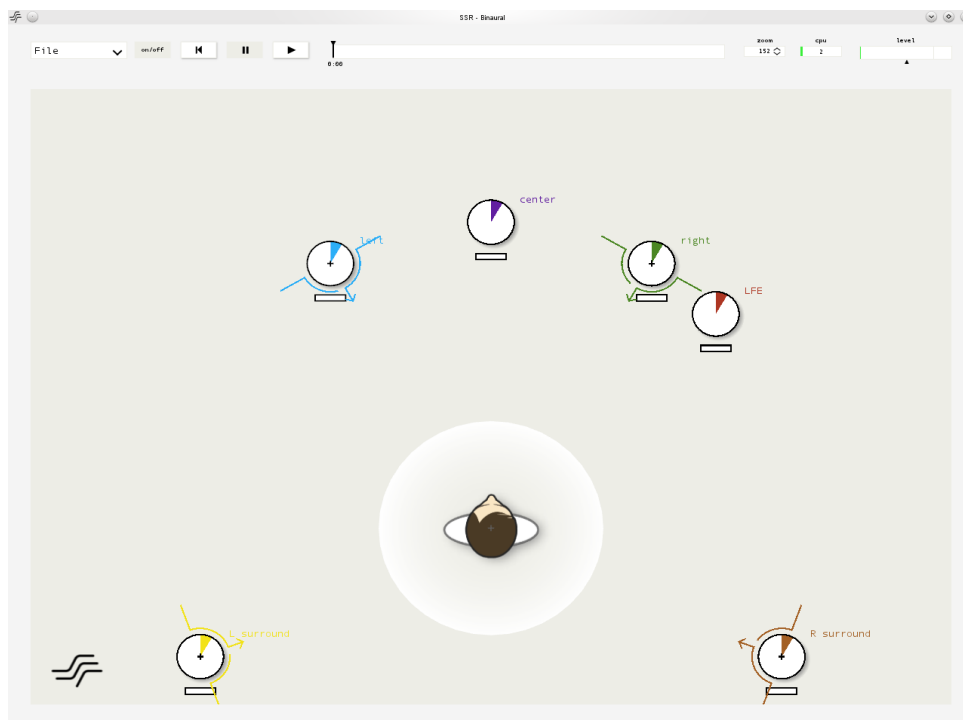


Abbildung 6.2: Umsetzung einer 5.1 Surround Audio-Szene im SSR.

Eine Verknüpfung aus einem Shell- bzw. Python-Skript ermöglicht es eine Audio-Szene im ASD-Format für kanalbasierte Audiodateien zu erstellen. Dem Skript muss neben der mehrkanaligen wav-Datei, die Anzahl der Kanäle und der Radius übergeben werden. Werden sechs Kanäle angegeben, so enthält die ASD-Datei die Beschreibung einer

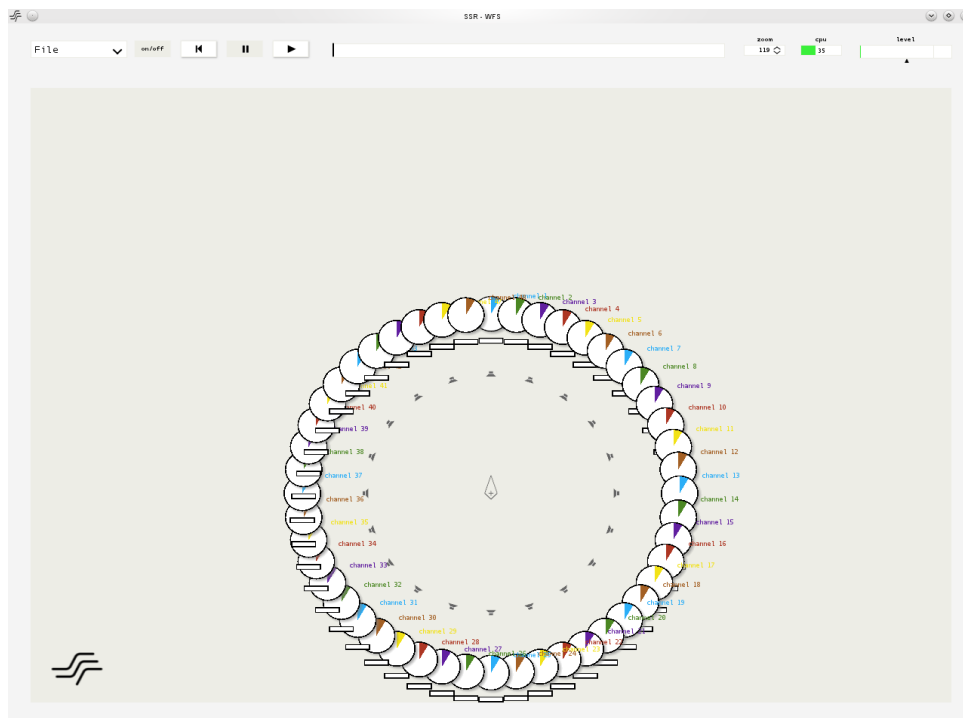


Abbildung 6.3: Umsetzung einer für einen Lautsprecherring konzipierten 48-Kanal-wav-Datei.

5.1-Audio-Szene. Werden mehr als sechs Kanäle übergeben, so beschreibt das Skript eine Szene mit einem gleichmäßig verteilten Quellenring. Hierbei sollte der Radius des Lautsprecherrings, für welchen die Audio-Quelldatei erzeugt wurde, übergeben werden um das Schallfeld möglichst nicht zu verfälschen. Das Skript ermöglicht es somit beliebige kanalbasierende Audio-Dateien im 5.1-Format und im 'Lautsprecher-Ring-Format' auf beliebigen Wiedergabe-Systemen darzustellen. Ob die Lautsprechersignale des 'Lautsprecher-Ring-Formats' nun ursprünglich für Wellenfeldsynthese, HOA oder VBAP berechnet wurden, spielt für die Wiedergabe im SSR keine Rolle.

## 6.2 Portierung der akustischen Gegebenheiten des Aufnahmerraums

Das ASD-Format bietet im Moment nicht die Möglichkeit die raumakustischen Merkmale des Aufnahmerraumes zu beschreiben. Wie in Abbildung 6.1 zu sehen ist, können diese Merkmale jedoch mit Hilfe von Impulsantwortmessungen reproduziert werden. Eine solche Messung erfordert jedoch einen gewissen Aufwand hinsichtlich der Zeit und des Aufnahmемaterials. Ein weiterer Ansatz ist die räumlichen Gegebenheiten geometrisch zu beschreiben und dann durch Einsatz des Spiegelquellenverfahrens aus Kapitel 3 die Schallreflexionen durch weitere Quellenpositionierungen im SSR zu erzeugen.

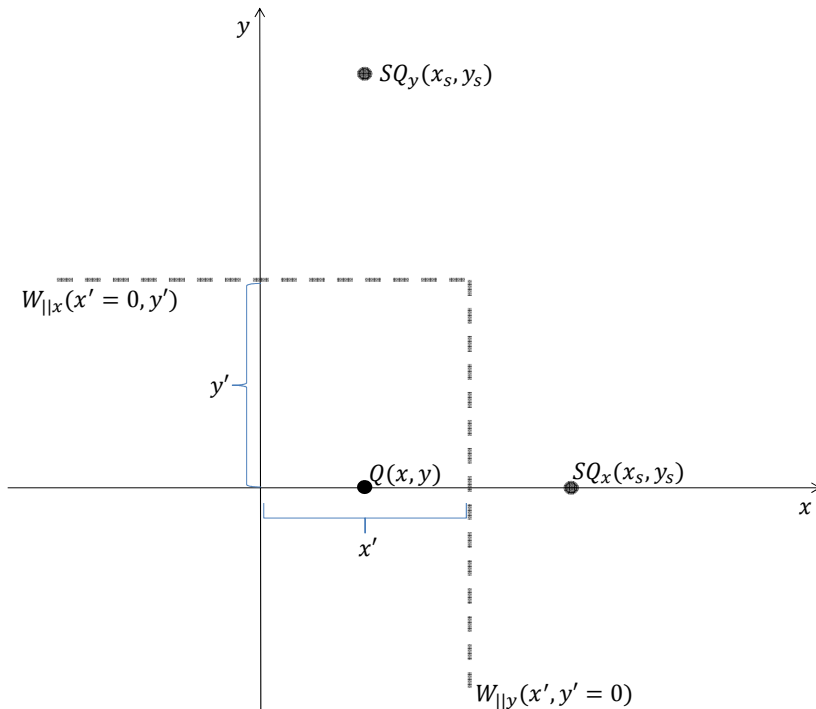


Abbildung 6.4: Geometrische Hilfe zur Berechnung der Spiegelquellen  $SQ_{x/y}(x_s, y_s)$  durch Spiegelung der Quelle  $Q(x, y)$  an der Wänden  $W_{||x}(x' = 0, y')$ / $W_{||y}(x', y' = 0)$ .



Existiert wie in Abbildung 6.4 zu sehen eine Quelle  $Q(x, y)$  in einem Raum mit den jeweiligen Wänden  $W(x', y')$ , wobei eine Wand parallel zur x-Achse bzw. y-Achse die Spiegeloperationseigenschaft  $W_{||x}(x' = 0, y')$  bzw.  $W_{||y}(x', y' = 0)$  besitzt, so kann die Spiegelquelle  $SQ(x_s, y_s)$  durch

$$SQ(x_s, y_s) = Q(x' - x + x', y' - y + y) = Q(2x' - x, 2y' - y) \quad (6.1)$$

ausgedrückt werden. Im Rahmen dieser Arbeit wurde ein Python-Skript entwickelt, welches abhängig von der im ASD-Format beschriebenen Raumgeometrie die Spiegelquellen bis zur zweiten Ordnung berechnen kann. Die Audio-Szene kann dann im ASD-Format folgendermaßen beschrieben werden:

Listing 6.1: Beschreibung einer Audio-Szene unter Berücksichtigung der Raumgeometrie.

```

1 <?xml version="1.0" encoding="utf-8"?>
2 <asdf version="0.2" coordinates="RFT">
3   <scene volume="3" name="Room_Acoustics">
4     <source name="Gitarre" model="point" volume="5">
5       <file channel="1">guitar.wav</file>
6       <position x="0.5" y="0"/>
7     </source>
8     <echo order="2">
9       <wall id="1" pos_x="1">
10      <wall id="2" pos_y="1">
11      <wall id="3" pos_x="-1">
12      <wall id="4" pos_y="-1">
13    </echo>
14  </scene>
15 </asdf>

```

Die Beschreibung beschränkt sich derzeit auf rechteckige Räume. Wie im Quellencode 6.1 zu sehen ist, können Wände in *pos\_x* oder *pos\_y* positioniert werden, wobei diese

jeweils parallel zu y- bzw. x-Achse verlaufen. Die daraus resultierende Szene im SSR ist in Abbildung 6.9 dargestellt. Zur besseren Orientierung sind die Koordinatenachsen und die Umrisse der virtuellen Räume angedeutet.

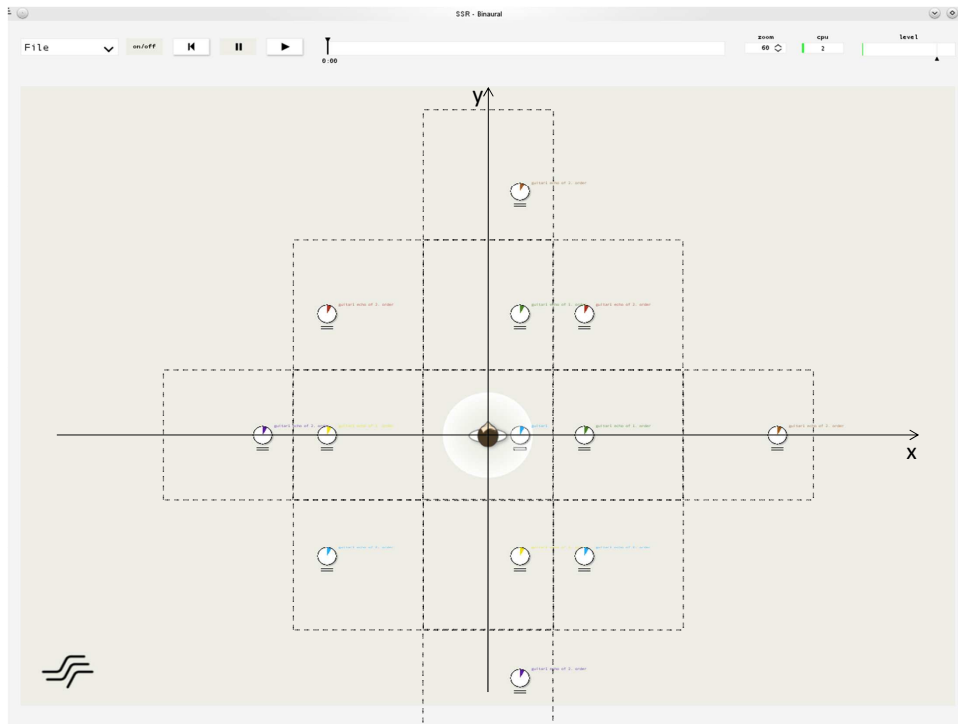


Abbildung 6.5: Darstellung der durch den virtuellen Raum erzeugten Spiegelquellen.

Das erwähnte Python-Skript ermöglicht es somit die Akustik eines rechteckigen Raumes zu reproduzieren. Dadurch ist es nicht mehr zwangsläufig notwendig Impulsantwortmessungen für Räumlichkeiten dieser Art durchzuführen. Eine Beschreibung der Raumgeometrie reicht somit aus, um deren Akustik reproduzieren zu können. Subjektiv ermöglichen Spiegelquellen bis zur zweiten Ordnung bereits einen gewissen Halleffekt, sofern die Raumbemaßungen groß genug gewählt werden. Das Skript ist jedoch deutlich ausbaufähig, indem Spiegelquellen höherer Ordnung genutzt werden. Ein effizienter Algorithmus zur Berechnung von Spiegelquellen höherer Ordnung in rechteckigen Räumen ist in [15] beschrieben. Trotz alledem ermöglichen Raumimpulsantwortmessung jedoch eine deutlich bessere Reproduktion der Raumakustik.

## 6.3 Portierung von Schallquellenbewegungen

Dieser Abschnitt beschreibt eine mögliche Umsetzung zur Beschreibung von Schallquellenbewegungen mit Hilfe des ASD-Formats. Im Quellcode 6.2 ist eine solche Bewegung dargestellt.

Listing 6.2: Beschreibung einer Audio-Szene mit sich bewegenden Schallquellen.

```
1 <?xml version="1.0" encoding="utf-8"?>
2 <asdf version="0.2" coordinates="RFT">
3   <scene volume="3" name="Moving_Sources">
4     <source name="Gitarre" model="point" volume="5">
5       <file channel="1">guitar.wav</file>
6       <position x="1" y="0"/>
7     </source>
8   </scene>
9   <score>
10    <par>
11      <animate id="1">
12        <spiral start="1_0" end="0_2" time="0s_10s">
13          <linear start="0_2" end="1_0" time="10s_20s">
14            <circular start="1_0" rotations="2" time="20s_30s">
15              <circular start="1_0" angle="180" time="30s_40s">
16            </animate>
17          </par>
18        </score>
19 </asdf>
```

Jede gesetzte Schallquelle erhält im SSR eine feste 'id'. Die 'id' wird in der Reihenfolge der Quellenpositionierung vergeben. In diesem Fall erhält die Quelle Gitarre die id="1", da sie als erste Quelle in Szenenbeschreibung aus Listing 6.2, in diesem Fall als Einzige genannt wurde. Der Befehl <par> leitet eine Reihe von im Zeitablauf par-

allelen Prozessen ein. Werden einer Quelle mehrere Bewegungen zugewiesen, so sollte jede Bewegung in sich nicht überschneidenden Zeitabschnitten stattfinden. Der Befehl `<animate id="1">` kündigt Bewegungen für die Schallquelle mit `id="1"` an. Im Moment ist es möglich eine spiralförmige(`<spiral>`), eine linienförmige(`<linear>`) und eine kreisförmige(`<circular>`) Bewegung innerhalb eines Zeitbereichs(`time="0s_10s"`) zu beschreiben. Die Bewegungen werden in diesem Abschnitt noch genauer beschrieben. Zu erwähnen bleibt, dass Bewegungen stetig sein sollten, da es sonst zu undefiniertem Verhalten bei der Spiegelquellen-Bewegung kommt.

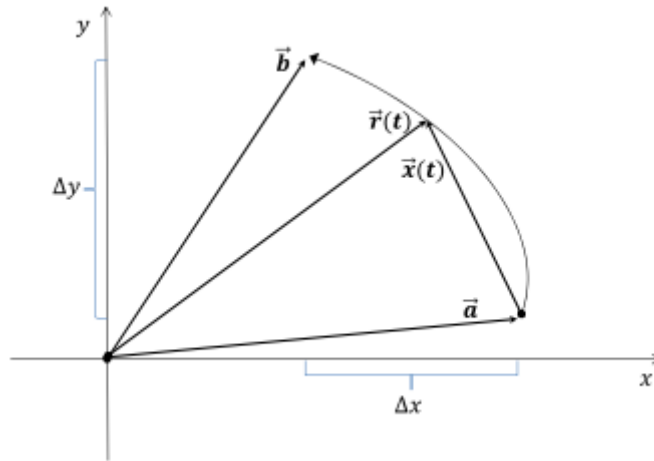


Abbildung 6.6: Geometrische Hilfe zur Beschreibung einer Bewegung entlang einer Kurve zwischen Anfangsvektor  $\vec{a}$  und Endvektor  $\vec{b}$ . Zum Zeitpunkt  $t$  befindet sich die Quelle um  $\vec{x}(t)$  von  $\vec{a}$  entfernt bei  $\vec{r}(t)$ .

Abbildung 6.6 zeigt geometrisch die Bewegung auf einer zeitabhängigen Bahnkurve  $\vec{r}(t)$  im Zeitabschnitt zwischen  $\tau_{\text{start}}$  und  $\tau_{\text{ende}}$ . Für die Gleichung der Bahnkurve gilt dann folgender Zusammenhang

$$\vec{r}(t) = (\varepsilon(t) - \varepsilon(t - \tau_{\text{ende}})) \cdot \vec{a} + (\varepsilon(t - \tau_{\text{start}}) - \varepsilon(t - \tau_{\text{ende}})) \cdot \vec{x}(t) + \varepsilon(t - \tau_{\text{ende}}) \cdot \vec{b}. \quad (6.2)$$

$\varepsilon(t - \tau)$  ist in diesem Fall die Sprungfunktion zu einem Zeitpunkt  $\tau$  und hilft die Zeitabschnitte zu definieren. Vor  $\tau_{\text{start}}$  befindet man sich somit am Startpunkt der Kurve bei  $\vec{a}$  und entfernt sich dann mit der Zeit um  $\vec{x}(t)$  und befindet sich zum Zeitpunkt

$\tau_{\text{ende}}$  bei  $\vec{b}$ .

Wie bereits im Quellcode 6.2 zu sehen war, ist es möglich diverse Bewegungen zu beschreiben. Abbildung 6.7 zeigt jeweils die möglichen Bewegungen, welche mit Hilfe von Vektoren dargestellt sind.

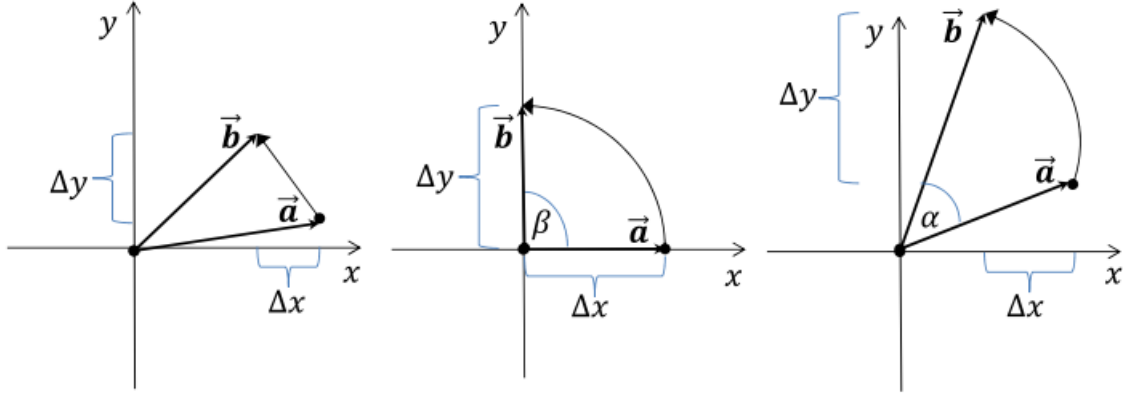


Abbildung 6.7: Darstellung der möglichen Bewegungen der Schallquellen. Links: Linear. Mitte: Zirkular. Rechts: Spiralförmig.

### 6.3.1 Lineare Bewegung

Für den Abstandsvektor  $\vec{x}$  der linearen Bewegung gilt dann folgender Zusammenhang

$$\vec{x}(t) = \frac{t - \tau_{\text{start}}}{\tau_{\text{ende}} - \tau_{\text{start}}} (\vec{b} - \vec{a}). \quad (6.3)$$

$\vec{x}(t)$  wird somit innerhalb der Zeit zwischen  $\tau_{\text{start}}$  und  $\tau_{\text{ende}}$  gleichmäßig gestreckt um zum Zeitpunkt  $\tau_{\text{ende}}$   $\vec{x}(t)$  der Differenz  $(\vec{b} - \vec{a})$  zu gleichen.

### 6.3.2 Zirkulare Bewegung

Um die Kreisbewegung zu beschreiben, benötigt man zunächst eine Kreisfrequenz  $\omega$  die folgendermaßen beschrieben werden kann

$$\omega = \frac{2\pi \cdot n}{\tau_{\text{ende}} - \tau_{\text{start}}} = \frac{\beta}{T}. \quad (6.4)$$

Der Winkel  $\beta$  kann im ASD-Format entweder direkt angegeben werden oder über einen Rotationsfaktor  $n \in \mathbb{R}$  beschrieben werden, wie in Listing 6.2 zu sehen ist. Für  $n = 1$

würde eine komplette Kreisbewegung ausgeführt werden. Für den Bewegungsvektor  $\vec{x}(t)$  gilt dann der folgende Zusammenhang

$$\vec{x}(t) = |\vec{a}| \cdot \begin{pmatrix} \cos(\omega \cdot (t - \tau_{\text{start}}) + \varphi) \\ \sin(\omega \cdot (t - \tau_{\text{start}}) + \varphi) \end{pmatrix} - \vec{a} \quad (6.5)$$

Im Endeffekt stellt er nichts anderes dar wie die Differenz zwischen dem Bahnkurvenvektor  $\vec{r}(t)$  und dem Anfangsvektor  $\vec{a}$ . Warum man den Bahnkurvenvektor nun nicht direkt beschreibt, soll im nächsten Abschnitt erklärt werden.

### 6.3.3 Spirale Bewegung

Für die spiralförmige Bewegung benötigt man zunächst den Unterschied des Radius zwischen ( $\vec{b}$  und  $\vec{a}$ ) und kann folgendermaßen ausgedrückt werden

$$\Delta r = |\vec{b}| - |\vec{a}|. \quad (6.6)$$

Um den Winkel zwischen den beiden Vektoren zu berechnen nutzt man folgenden Zusammenhang

$$\cos(\alpha) = \frac{\vec{b} \cdot \vec{a}}{|\vec{b}| \cdot |\vec{a}|}, \quad (6.7)$$

wobei diese Formel in jeder besseren mathematischen Formelsammlung zu finden ist. Für die Kreisfrequenz muss nun lediglich der errechnete Winkel durch die Zeitabschnittsdifferenz  $T$  geteilt werden

$$\omega = \frac{\alpha}{\tau_{\text{ende}} - \tau_{\text{start}}} = \frac{\alpha}{T} \quad (6.8)$$

Nun hat man sämtliche Größen die zur Berechnung des Bewegungsvektors benötigt werden und für welchen folgendes gilt

$$\vec{x}(t) = \left( |\vec{a}| + \frac{(t - \tau_{\text{start}}) \cdot \Delta r}{T} \right) \cdot \begin{pmatrix} \cos(\omega \cdot (t - \tau_{\text{start}}) + \varphi) \\ \sin(\omega \cdot (t - \tau_{\text{start}}) + \varphi) \end{pmatrix} - \vec{a}. \quad (6.9)$$

Zu erwähnen bleibt, dass sich Radius der Bahnkurve  $\vec{r}(t)$  für die spiralförmige Bewegung gleichmäßig über dem Zeitabschnitt  $T$  erhöht, also  $|\vec{r}(\tau_{\text{start}})| = |\vec{a}|$  und  $|\vec{r}(\tau_{\text{ende}})| = |\vec{b}|$ .

In diesem Abschnitt wurden Möglichkeiten aufgezeigt, um Bewegungen in einer Ebene mit Hilfe des ASD-Formats in einfacher Weise zu beschreiben ohne dafür besonders weitreichende mathematische Kenntnisse zu besitzen.

## 6.4 Kombination von Schallquellenbewegungen und Raumakustik

Nun werden die Beschreibung der Raumakustik und die Beschreibung der Quellenbewegung im ASD-Format verknüpft. Diese kombinierte Beschreibung ist in folgendem Quellcode zu sehen.

Listing 6.3: Beschreibung einer Audio-Szene unter Berücksichtigung der Raumgeometrie und der Quellenbewegung.

```
1 <?xml version="1.0" encoding="utf-8"?>
2 <asdf version="0.2" coordinates="RFT">
3   <scene volume="3" name="Moving_Sources">
4     <source name="Direct_Sound_1" model="point" volume="5">
5       <file channel="1">guitar.wav</file>
6       <position x="0.5" y="0"/>
7     </source>
8     <echo order="2">
9       <wall id="1" pos_x="1">
10        <wall id="2" pos_y="1">
11        <wall id="3" pos_x="-1">
12        <wall id="4" pos_y="-1">
13      </echo>
14    </scene>
15    <score>
16      <par>
17        <animate id="1">
18          <linear start="0.5_0" end="0.7_0.7" time="10s_20s">
19        </animate>
20      </par>
21    </score>
```



22 </ asdf>

Nun stellt sich die Frage wie eine Bewegung der Basisquelleschallquelle sich auf die Bewegung der zugehörigen Spiegelquelle auswirkt. Eine bildliche Illustration der Bewegungen aus dem Beispiel des Quellcodes ist in Abbildung 6.8 dargestellt.

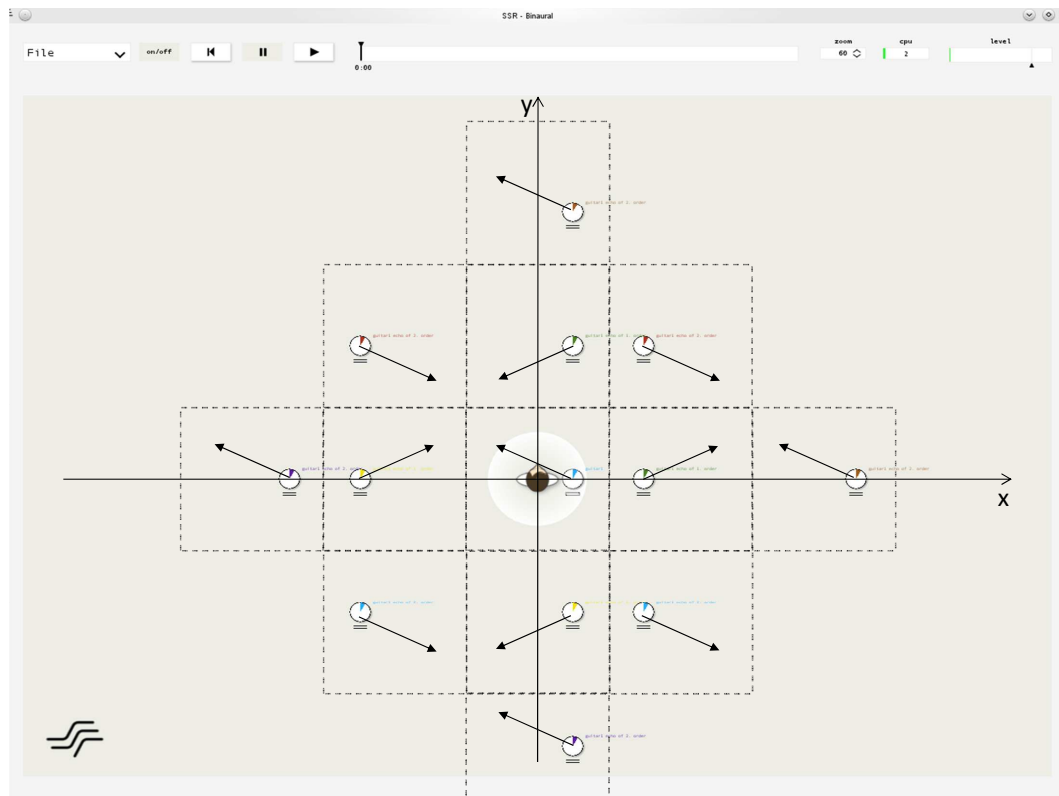


Abbildung 6.8: Zusammenhang zwischen den Bewegungen der Basisquelle und der Spiegelquellen.

Nun wird auch deutlich warum man vorhin den Umweg über den Bewegungsvektor  $\vec{x}(t)$  gegangen ist. Die Bewegungsdistanz ist für alle Quellen gleich. Sie unterscheidet sich lediglich in der Richtung und somit nur um einen eventuellen Vorfaktor  $(-1)$ .

Der Algorithmus zur Berechnung der Spiegelquellen und der zugehörigen Bewegungsvektoren ist in Abbildung 6.9 als Flussgraf dargestellt. Die Deutung der Größen ist in Abbildung 6.4 zu sehen. Eine Schallquelle  $Q(x, y)$  wird jeweils an den beiden parallel zur y-Achse bzw. x-Achse parallelen Wänden gespiegelt. Dadurch erhält der Bewe-

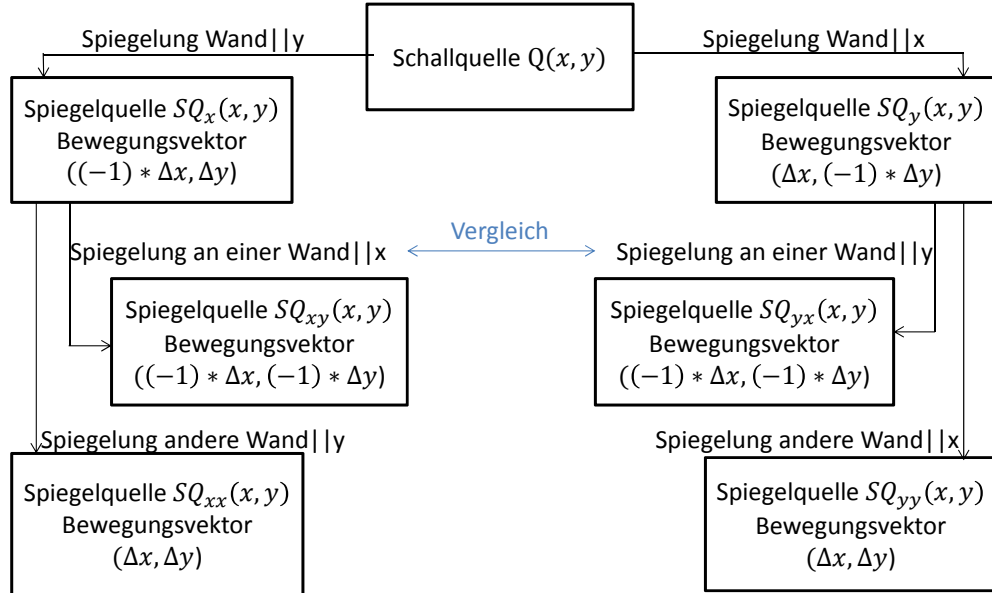


Abbildung 6.9: Algorithmus zum setzen der Spiegelquellen. Wand||x→Wand parallel zur x-Achse.

gungsvektor den Vorfaktor  $(-1)$  für seinen x/y-Anteil für eine die Spiegelung an einer zur y/x-Achse parallelen Wand. Danach wird die erzeugte Spiegelquelle  $SQ_x(x, y)$  bzw.  $SQ_y(x, y)$  an der gegenüber liegenden zur x/y-Achse parallelen Wand gespiegelt. Der Vorfaktor des Bewegungsvektor wird durch nochmalige Multiplikation mit  $(-1)$  negiert. Somit entspricht die Bewegung  $SQ_{xx}(x, y)$  bzw.  $SQ_{yy}(x, y)$  der Bewegung der Basisquelle  $Q(x, y)$ .  $SQ_x(x, y)/SQ_y(x, y)$  wird des Weiteren an einer zur x/y-Achse parallelen Wand gespiegelt. Dabei muss ein Vergleich erfolgen, auf welche Wand die Spiegelung angewendet wurde, um eine doppelte Platzierung einer Spiegelquelle an gleicher Position zu vermeiden. Der Vorfaktor für den Bewegungsvektor der Spiegelquellen  $SQ_{xy}(x, y)$  bzw.  $SQ_{yx}(x, y)$  ist  $(-1)$  für beide Anteile und ist somit genau entgegengesetzt der Bewegung der Basisquelle  $Q(x, y)$ . Der gesamte Spiegelprozess, welcher in

diesem Absatz beschrieben wurde, muss für beide zur y-Achse bzw. x-Achse parallelen Wände angewendet werden.

Wie bereits erwähnt sollten die Bewegungen stetig erfolgen, da sonst undefiniertes Verhalten der Spiegelquellenbewegungen auftritt. Insgesamt bietet die Kombination aus Spiegelquellen und Bewegung bereits einige Möglichkeiten um eine Audio-Szene zu gestalten.

# Kapitel 7

## Ausblick

Zur Portierung kanalbasierter Audio-Dateien steht bereits das beschriebene Skript aus Abschnitt 6.1 zu Verfügung. Dieses ermöglicht es jedoch im Moment nur die Portierung des 5.1-Formats und von für Lautsprecherringe konzipierten Dateien. Interessant wäre es, das Skript auf Lautsprecherrechtecke, Linien oder andere Konstellationen zu erweitern.

Die akustischen Gegebenheiten eines Aufnahmeraums können zwar bereits für rechteckige Räume im ASDF beschrieben werden. Es wäre jedoch interessant auch andere Raumformen beschreiben zu können. Des Weiteren können im Moment nur Spiegelquellen bis zur zweiten Ordnung berechnet werden. Dadurch ist der Hall nicht ganz so satt, wie man ihn aus der Realität gewöhnt ist. In der Realität werden außerdem hohe Frequenzen durch die Reflektion an der Wand gedämpft, was jedoch bei der Platzierung der Quellen im SSR nicht berücksichtigt wurde. So müsste an sich jede Spiegelquelle vor der Platzierung gefiltert werden. Deutlich kommt dieser Effekt bei Erzeugung von Spiegelquellen einer Kick-Drum zum Vorschein. Diese klingt dann recht verwaschen und blechern. Es wirkt so, als wäre der Lautsprecher beschädigt.

Das Bewegungsmodul selber funktioniert von den Berechnungen her bereits sehr gut. Allerdings gibt es noch Probleme beim Auslesen des ASD-Formats. Die regulären Ausdrücke, welche zum durchsuchen der asd-Datei genutzt werden, müssen noch an sämtliche Syntax-Konstellationen der Beschreibungssprache angepasst werden. Die dynami-

sche Beschreibung einer Szene könnte noch weiter ausgeführt werden. Im Moment sind bereits vier Bewegungen möglich. Die Möglichkeit, weitere Bewegungen zu beschreiben, wäre durchaus interessant. Des Weiteren wäre es interessant Veränderung in der Lautstärke beschreiben zu können.

Bei der Kombination der beiden Skripte können bisher nur stetige Bewegungen beschrieben werden. Kommt es zu örtlichen Sprüngen der Schallquelle, wird dieser Sprung nicht auf die Spiegelquellen angewendet. Insgesamt können natürlich sämtliche Skripte hinsichtlich ihrer Schnelligkeit optimiert werden. Dazu könnte zum einen der Programmcode selber optimiert werden, zum anderen eine systemnähere Programmiersprache wie C++ genutzt werden.

Es stellt sich weiterhin die Frage, warum die in Kapitel 2 erläuterten Wiedergabeverfahren sich bis heute größtenteils nicht massentauglich durchgesetzt haben, obwohl diese deutliche Verbesserung hinsichtlich der räumlichen Wahrnehmung von Hörereignissen bieten. In meinen Augen liegt das wohl an der aufwendigen Umsetzung des Wiedergabeverfahrens. Ein Versuch zur Verbreitung der binauralen Wiedergabetechnik ist im Moment durch die Planung einer sogenannten Kopfhörerparty in Planung. Diese Party ermöglicht es, fortschrittlichen Raumklang einer breiten Masse zu präsentieren und somit für eine Mundpropaganda dieses Verfahrens zu sorgen. Ein Test zur Durchführung der Kopfhörerparty findet bereits in diesem Jahr im LMS statt.

# Kapitel 8

## Zusammenfassung

Wie Kapitel 2 zeigt, gibt es eine Vielzahl an räumlichen Wiedergabeverfahren. Diese Arbeit gibt zunächst einen Überblick der existierenden Verfahren und erklärt deren Prinzipien. Zur Portierung von Audio-Szenen zwischen den verschiedenen Wiedergabesystemen wurde im Rahmen dieser Arbeit der Soundscape Renderer genutzt. Dieses Programm bietet bereits diverse Möglichkeiten zur Beschreibung objektorientierter Audio-Szenen in Form des ASD-Formats wie in Kapitel 4 beschrieben. Um die Demos des LMS, welche größtenteils nur in kanal-basierender Form zur Verfügung stehen, für das im Bau befindliche Wellenfeldsynthesesystem nutzbar zu machen, wurde ein Szenen-Beschreibungs-Skript entwickelt. Dieses Skript nutzt die Lautsprecherpositionen des Wiedergabesystems, für welche die kanal-basierende wav-Datei konzipiert ist, um die Szene im ASD-Format zu beschreiben. Im Rahmen dieser Arbeit wurde außerdem ein Konzept entwickelt, um das ASD-Format flexibler zu gestalten und mehr Eigenschaften einer Audio-Szene aufzuzeichnen. Eine Möglichkeit besteht darin, rechteckige Aufnahmeräume zu beschreiben und somit deren Akustik nachträglich zu rekonstruieren. Dies ist durch den Einsatz von Spiegelquellen möglich, wobei der Rechenaufwand dadurch deutlich ansteigt, da das entwickelte Skript für jede Schallquelle zusätzliche zwölf Spiegelquellen im SSR erzeugt. Eine weitere Möglichkeit ist es, Bewegungen einzelner Schallquellen zu beschreiben und somit dynamische Audio-Szenen zu erschaffen. Es wurden vier verschiedene Bewegungen entwickelt, welche auf einfache Weise im

ASD-Format beschrieben werden können. Kombiniert man die Aufnahmeraum- und Bewegungs-Beschreibung, so bewegen sich die Spiegelquellen abhängig von der Bewegung der ihnen zugeordneten Schallquelle.

# Anhang A

## Anhang Kapitel

### A.1 How To SSR

#### A.1.1 Installation

In diesem Abschnitt soll kurz erklärt werden, welche Schritte zur Installation des Soundscape Renderers nötig sind. Das Programm steht als Open Source Projekt unter [13] zum freien Download zur Verfügung. Um die heruntergeladene Datei zu entpacken und in das Verzeichnis des SSR zu gelangen, müssen folgende Befehle in der Kommandozeile einer Shell eingegeben werden.

```
1 tar xvzf ssr-x.x.x.tar.gz
2 cd ssr-x.x.x
```

Oftmals ist es nötig vor der Kompilierungs-Konfiguration einige Zusatzpakete für die jeweilige Linux-Distribution zu installieren. In unserem speziellen Fall am Rechner des LMS wurde die Ecasound Library benötigt. Des Weiteren traten Probleme mit der boost-system und boost-thread Library auf. Dies scheint aber an der Linux-Distribution openSuse zu liegen, da diese keine Verlinkung zwischen den Verzeichnissen der 32- und 64-bit Version der Library zu Verfügung stellt. Bei anderen Linux-Distributionen scheint dieses Problem nicht aufzutreten. Um dieses Problem zu umgehen muss in Zeile 16160 und 16192 im configure-Shell-Skript



```
1 for dir in /{usr,opt}/{,local/}lib; do
```

ersetzt werden durch

```
1 for dir in /{usr,opt}/{,local/}lib{,64}; do
```

wobei die Zeilenanzahl für aktuelle Version des SSR gilt(0.3.2). Falls sich die Zeile ändert kann auch einfach nach `boost_system` bzw. `boost_thread` in der `configure`-Datei gesucht werden. Werden noch andere Librarys zur Installation benötigt, so wird man darauf hingewiesen. Nicht alle Librarys sind zwingend erforderlich. Beispielsweise sind `boost_system` und `boost_thread` für die Netzwerkschnittstelle des SSR notwendig. Ohne diese kann also kein Netzwerkzugriff erfolgen, die andern Funktionen des SSR stehen jedoch zu Verfügung. Welche Funktionen verfügbar sind, wird am Ende der Ausführung des `configure`-Skripts aufgelistet. Folgende fünf Befehle führen bereits zur vollenden die Installation des SSR.

```
1 ./configure
```

```
2 make
```

```
3 make install
```

Manchmal geschieht es, dass nachträglich library-Pakete installiert werden, da man z.B. eventuell doch alle Funktionen des SSR nutzen will. In diesem Fall müssen die letzten fünf Befehle wiederholt werden, wobei im vorhinein ein

```
1 make clean
```

durchgeführt werden sollte.

### A.1.2 Jack Audio Server

Bevor der Soundscape Renderer gestartet werden, muss zunächst der Jack Audio Server gestartet werden. Entweder geschieht dies über den Befehl

```
1 qjackctl &
```

, wobei in diesem Fall eine grafische Oberfläche zur Audio-Stream-Koordination geöffnet wird. Andererseits ist es möglich den Server als reines Kommandozeilenprogramm über folgenden Befehl

```
1 jackd -d alsa -r 44100 -p 1024
```

zu starten. Mit `-d` wird das backend gewählt, wobei die Kommunikation mit der Soundkarte in diesem Fall über die Advanced Linux Sound Architecture (ALSA) erfolgt. Näheres dazu kann in [9] gefunden werden. Mit `-r` kann die Abtastrate festgelegt, welche der Abtastrate der genutzten wav-Dateien entsprechen muss. Der Wert hinter `-p` muss einer Zweierpotenz entsprechen und legt Puffer-Größe fest. Um den SSR für das Wellenfeldsynthesesystem zu starten ist das Kommandozeilenprogramm zu empfehlen. Für die binaurale Wiedergabe im SSR kann auch die grafische Oberfläche genutzt werden. Wird der Jack Server nicht mehr benötigt so kann er mit

```
1 killall jack
```

beendet werden.

### A.1.3 Starten des Soundscape Renderers

Um den SSR zu starten, muss zunächst im Verzeichnis des SSR in den Ordner `src/` gewechselt werden. Dann kann mit

```
1 ./ssr
```

der Soundscape Renderer standardmäßig im binauralen Modus gestartet werden, wobei der Parameter `--help` eine Übersicht der möglichen Optionen gibt. Als letztes Argument erhält der SSR eine Szenenbeschreibungs-Datei im ASD-Format (`szene.asd`) oder eine wav-Datei, wobei bei letzterem Fall die Quellenpositionierung je nach Anzahl der Kanäle der wav-Datei unterschiedlich ausfällt. Wird keine Szenenbeschreibung übergeben, kann nachträglich mit Hilfe der grafischen Oberfläche eine asd- bzw. wav-Datei ausgewählt werden. Der SSR kann Lautsprecher-signale für verschiedene Wiedergabeverfahren erzeugen wie in Abschnitt 4.1 zu sehen ist. Das Wiedergabeverfahren kann

über Parameter bestimmt werden, welche beim Aufruf des SSR mit `--help` aufgelistet werden. Zusätzlich muss die geometrische Lautsprecheranordnung des Wiedergabesystems über `--setup=lautsprecheranordnung.asd` angegeben werden. Ein solche Lautsprecherkonstitutionsbeschreibung ist in Listing 4.2 zu sehen. Weitere Beispiele sind im Unterordner `data/reproduction_setups` des SSR-Verzeichnisses zu finden.

#### A.1.4 Ausführen der Python- bzw. Shell-Skripte

Zum Ausführen eines Python-Skripts muss letztendlich nur der du Vorsatz `python` vor das auszuführende Skript gesetzt werden. Bei einem Shell-Skript stellt man den Vorsatz `sh` vor das auszuführende Skript. Folgende Skripte wurden im Rahmen dieser Arbeit entwickelt:

```
1 create_scene_in_ssr.sh
2 eichenau_add_reverb.py
3 eichenau_mute_reverb.py
4 moveSources_createEcho_by_ASDF.py
5 sceneParser.py
```

Das *sceneParser*-Skript ist nicht zur eigenständigen Ausführung gedacht. Es wird innerhalb *create\_scene\_in\_ssr*-Skripts aufgerufen. Der Aufruf des *create\_scene\_in\_ssr*-Skripts erfolgt folgendermaßen:

```
1 sh create_scene_in_ssr.sh audio.wav [Anzahl Kanäle] [Radius]
```

Werden dem Skript eine wav-Datei mit sechs Kanälen übergeben, so erstellt dieses eine Szenen Beschreibung im ASD-Format für eine 5.1-Lautsprecheranordnung im Home-Verzeichnis des Nutzers. Werden mehr als sechs Kanäle übergeben, so erstellt das Skript eine asd-Datei für eine Lautsprecherring-Anordnung.

Das *eichenau\_add\_reverb*- bzw. *eichenau\_mute\_reverb*-Skript schaltet die für den Hall verantwortlichen Schallquellen des Eichenau-Konzerts ein oder aus. Bevor diese ausgeführt werden können, muss die Netzwerkschnittstelle des SSR aktiviert sein und

die ASD-Datei eichenau\\_multichannel.asd im SSR geladen sein. Um die Netzwerkschnittstelle zu aktivieren, übergibt man beim Starten des SSR die Parameter `-i -t`. Auch beim *moveSources\_createEcho\_by\_ASDF*-Skript muss die Netzwerkschnittstelle aktiviert. Als erstes Argument erhält das Skript die Szenenbeschreibung im ASD-Format wie z.B. in Listing 6.3 gezeigt ist.

# Abbildungsverzeichnis

2.1	Geometrisches Schema 2-Kanal-Stereosystems mit den Einheitsvektoren $\mathbf{e}_l$ und $\mathbf{e}_r$ in Richtung der Lautsprecher bzw. $\mathbf{e}_p$ in Richtung der Phantomschallquelle [2]. . . . .	4
2.2	Beziehung zwischen Phantomschallquellenwinkel $\theta$ und Amplitudenverhältnis $g_l/g_r$ des <b>Tangens-Gesetz</b> ( gestrichelte Linie) bzw. <b>Sinus-Gesetz</b> (durchgehende Linie) [2]. . . . .	4
2.3	Direkter Vergleich zwischen Einfluss der Laufzeit- bzw. Pegeldifferenz auf die Wahrnehmung der Hörereignisrichtung ( $0\% \rightarrow \theta = 0^\circ$ , $100\% \rightarrow \theta = \theta_0$ ) [4]. . . . .	5
2.4	Blumlein Mikrofonanordnung: Die beiden Linien stellen das typische Aufnahmeverhalten von zwei um $90^\circ$ versetzte Druckgradienten-Mikrofone abhängig vom Einfallswinkel $\alpha$ dar [1]. . . . .	6
2.5	Beziehung zwischen Aufnahme- $\alpha$ und Wiedergabewinkel $\psi$ der Schallquelle [1]. . . . .	6
2.6	5.1-Surround Wiedergabeschema [3]. . . . .	10
2.7	5.1-Surround Aufnahmeschema [1]. . . . .	10
2.8	Zirkuläre Ambisonics Lautsprecheranordnung mit Deutung der Aufnahmeverhaltens der verwendeten Druckgradientenmikrofone, Vergleich Abbildung 2.4 [3]. . . . .	11
2.9	Vergleich zwischen zirkulären und sphärischen Moden [3]. . . . .	11

2.10	Geometrische Deutung in Polarkoordinaten für die Ortskoordinaten $(x, y)$ und die räumliche Frequenz $\beta$ [7]. . . . .	13
2.11	Berechnung der winkel- bzw. radiusabhängigen Anteile der <i>Fourier</i> -Koeffizienten bis zur dritten Ordnung. Oben: Realteil $\Re \{ \}$ der winkelabhängigen Funk- tion $a_v(\theta)$ . Mitte: Radiusabhängige Bessel-Funktion $b_v(r)$ . Unten: Realteil $\Re \{ \}$ des Produkts [7]. . . . .	14
2.12	Deutung des Huygensschen Prinzips als Grundlage der Wellenfeldsyn- these [1]. . . . .	16
2.13	Geometrische Hilfe zur Berechnung des Schalldrucks $P(\mathbf{x}, \omega)$ ausgelöst durch eine Quelle $Q(\mathbf{x}, \omega)$ innerhalb eines Volumens $V$ . Der Vektor $\xi$ zeigt auf den Rand $\partial V$ des Volumens mit dem Normalenvektor $\mathbf{n}$ , der Vektor $\mathbf{x}$ auf den zu bestimmenden Schalldruck [7]. . . . .	17
2.14	Beziehung zwischen Aufnahme und Wiedergabe im WFS-System [9]. . .	19
2.15	Links: Veranschaulichung zweier am Kopf ankommenden Wellen [1]. Rechts: Verarbeitung der Signale durch das menschliche Gehör (adap- tiert von [3]). . . . .	21
2.16	Messschema zur Bestimmung von HRTFs [1]. . . . .	22
2.17	Abbildung eines Kunstkopfes zur Messung von HRTFs [10]. . . . .	23
3.1	Darstellung des Spiegelquellenprinzips[11] . . . . .	25
3.2	Konstruktion der Spiegelquellen für einen rechteckigen Raum[11] . . . .	26
4.1	Flussgraphen zur Bestimmung der Lautsprecher bzw. Kopfhörersignale. Source signal - Quellensignal; angle of incidence - Einfallswinkel; delay - Verzögerung; weight - Gewichtung; [14]. . . . .	29
4.2	Einfluss einer Punktquelle auf die Lautsprechersignale eines Wellenfeldsynthese- Systems im Soundscape Renderer. . . . .	31
4.3	Einfluss einer planaren Welle auf die Lautsprechersignale eines Wellenfeldsynthese- Systems im Soundscape Renderer. . . . .	31

6.1	Audio-Szene eines Konzerts in der katholischen Kirche in Eichenau: Die Quellen des Orchesters bzw. Chors sind hervorgehoben, die planaren Hallquellen in einem Kreis angeordnet. . . . .	40
6.2	Umsetzung einer 5.1 Surround Audio-Szene im SSR. . . . .	42
6.3	Umsetzung einer für einen Lautsprecherring konzipierten 48-Kanal-wav-Datei. . . . .	43
6.4	Geometrische Hilfe zur Berechnung der Spiegelquellen $SQ_{x/y}(x_s, y_s)$ durch Spiegelung der Quelle $Q(x, y)$ an der Wänden $W_{  x}(x' = 0, y')/W_{  y}(x', y' = 0)$ . . . . .	44
6.5	Darstellung der durch den virtuellen Raum erzeugten Spiegelquellen. .	46
6.6	Geometrische Hilfe zur Beschreibung einer Bewegung entlang einer Kurve zwischen Anfangsvektor $\vec{a}$ und Endvektor $\vec{b}$ . Zum Zeitpunkt $t$ befindet sich die Quelle um $\vec{x}(t)$ von $\vec{a}$ entfernt bei $\vec{r}(t)$ . . . . .	48
6.7	Darstellung der möglichen Bewegungen der Schallquellen. Links: Linear. Mitte: Zirkular. Rechts: Spiralförmig. . . . .	49
6.8	Zusammenhang zwischen den Bewegungen der Basisquelle und der Spiegelquellen. . . . .	53
6.9	Algorithmus zum setzen der Spiegelquellen. Wand $  x \rightarrow$ Wand parallel zur x-Achse. . . . .	54

# Tabellenverzeichnis

- 5.1 Übersicht über die vorhandenen Hörbeispiele des aktuellen WFS-Systems  
des LMS. Eine Beschreibung der einzelnen Projekte ist im Text zu finden. 36



# Literaturverzeichnis

- [1] R. Rabenstein, “Vorlesungsskriptum mensch maschine schnittstelle,” Lehrstuhl für Multimediakommunikation und Signalverarbeitung ( LMS ), 2011.
- [2] R. Rabenstein and S. Spors, *Handbook of Speech Processing*. Springer, December 2007, ch. 53 Sound Field Reproduction.
- [3] Blauert, Jens, and R. Rabenstein, “Schallfeldsynthese mit lautsprechern i - beschreibung und bewertung,” in *9. ITG Fachtagung Sprachkommunikation*, Bochum, October 2010, paper 42.
- [4] S. Weinzierl, Ed., *Handbuch der Audiotechnik*. Berlin, Heidelberg: Springer, 2008.
- [5] V. Pulkki, “Spatial sound generation and perception by amplitude panning techniques,” Ph.D. dissertation, Helsinki University of Technology, August 2001.
- [6] (2011, August) 5.1. [Online]. Available: <http://de.wikipedia.org/wiki/5.1>
- [7] R. Rabenstein and J. Blauert, “Schallfeldsynthese mit lautsprechern ii - signalverarbeitung,” in *9. ITG Fachtagung Sprachkommunikation*, Bochum, October 2010.
- [8] S. Spors, R. Rabenstein, and J. Ahrens, “The theory of wave field synthesis revisited,” in *Audio Engineering Society (AES) 124th Convention*, Amsterdam, May 2008, p. Paper Number: 7358.
- [9] M. Fischer, “Studienarbeit: Eine audiovisuelle demonstrationsumgebung für die vielkanal-lautsprecher-wiedergabe,” Lehrstuhl für Multimediakommunikation und Signalverarbeitung ( LMS ), 2009.

- [10] (2011, August) Binaurale tonaufnahme. [Online]. Available: [http://de.wikipedia.org/wiki/Binaurale\\_Tonaufnahme](http://de.wikipedia.org/wiki/Binaurale_Tonaufnahme)
- [11] P. M. Krejci, “Entwicklung eines mehrkanaligen systems für eine virtuelle raumakustik,” Master’s thesis, Universität für Musik und darstellende Kunst Graz, 2001.
- [12] S. McGovern, “The image-source reverberation model in an n-dimensional space,” in *DAFx-11 Proceedings*, September 2011, pp. 19–23.
- [13] M. Geier, J. Ahrens, and S. Spors, *Introduction to the SoundScape Renderer (SSR)*, 2011. [Online]. Available: <https://dev.qu.tu-berlin.de/projects/ssr/files>
- [14] M. Geier and S. Spors, “Conducting psychoacoustic experiments with the soundscape renderer,” in *9.ITG Fachtagung Sprachkommunikation*, 2010.
- [15] J. B. Allen and D. A. Berkley, “Image method for efficiently simulating small-room acoustics,” Juni 1978.
- [16] M. Geier, J. Ahrens, and S. Spors, “Object-based audio reproduction and the audio scene description format,” *Organized Sound*, vol. 15(3), pp. 219–227, 2010.
- [17] —, “Asdf: Ein xml format zur beschreibung von virtuellen 3d-audioszenen,” in *34rd German Annual Conference on Acoustics (DAGA)*, 2008.
- [18] M. Geier, S. Spors, and S. Weizierl, *Adaptive Multimedia Retrieval. Identifying, Summarizing, and Recommending*. Springer, 2010, ch. The future of audio reproduction: Technology - Formats - Applications, pp. 1–17.