

3D Mikrofonanordnungen

**Eine Übersicht nach Anwendungsbereichen,
Eigenschaften und Wiedergabesystemen**

Bachelorarbeit im Studiengang Audiovisuelle Medien

vorgelegt von

Florian Heidecker

Matr.-Nr.: 35940

am 22. Februar 2022

an der Hochschule der Medien Stuttgart

zur Erlangung des akademischen Grades

Bachelor of Engineering (B. Eng.)

Erstprüfer: Prof. Dr. Frank Melchior

Zweitprüfer: Prof. Oliver Curdt

Ehrenwörtliche Erklärung

Hiermit versichere ich, Florian Heidecker ehrenwörtlich, dass ich die vorliegende Bachelorarbeit mit dem Titel: „3D Mikrofonanordnungen – Eine Übersicht nach Anwendungsbereichen, Eigenschaften und Wiedergabesystemen“ selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

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

Florian Heidecker

Lübeck, den 22.02.2022

Kurzfassung

Neben der optimalen Wiedergabe von 3D-Audio stellen sich Tonschaffende die Frage, wie auditive Inhalte für diese produziert werden können. Die vorliegende Arbeit fasst unterschiedliche immersive Aufnahmemöglichkeiten in einer sinnvollen Kategorisierung zusammen, welche dem Leser Unterstützung zu individuellen Anwendungsentscheidungen leisten soll.

Unter Berücksichtigung tontechnischer Grundlagen werden Einblicke in die Hintergründe unterschiedlicher Mikrofonanordnungen ermöglicht, wonach dem Leser eine Kategorisierung dieser bezüglich ihrer Einsatzbereiche, Eigenschaften und der Wiedergabe ihrer Aufnahmen nahegebracht wird. Unterdessen werden die Themengebiete der Mikrofonprinzipien, psychoakustischer Grundlagen, stereofoner Aufnahmeverfahren und immersiver Lautsprecherwiedergabe behandelt, worauf eine Taxonomie und Beschreibung der 3D-Mikrofonanordnungen folgt.

Eine Auswertung erfolgt auch in einer der Thesis anhängenden Tabelle, welche eine Kategorisierung und Bewertung der Mikrofonanordnungen bezüglich ihrer Eigenschaften, Anwendungsbereiche und Wiedergabemöglichkeiten geben soll.

Abstract

In addition to the optimal rendering of 3D audio sound creators face the question of how to produce auditory content for such purpose. This paper summarizes different immersive recording options in a meaningful categorization, which should support the reader in making individual application decisions.

Considering sound engineering basics, insights into the background of different microphone arrangements are provided, after which the reader is introduced to a categorization of these with respect to their areas of application, characteristics and the reproduction of their recordings. Meanwhile, the topics of microphone principles, psychoacoustic fundamentals, stereophonic recording techniques and immersive loudspeaker reproduction are covered, followed by a taxonomy and description of 3D microphone arrangements.

An evaluation also takes place in a table attached to the thesis, which should give a categorization and evaluation of the microphone arrangements regarding their characteristics, application areas and reproduction possibilities.

Danksagung

An dieser Stelle möchte ich mich bei all denjenigen bedanken, die mich während der Anfertigung dieser Bachelorarbeit fachlich und persönlich unterstützt und motiviert haben.

Herzlichen Dank gebührt zuerst Professor Dr. Frank Melchior, welcher mir als Betreuer mit hilfreichen Anregungen und einem offenen Ohr stets zur Seite stand. Erst durch seine spannenden Vorlesungen und Workshops bin ich auf das Thema 3D Audio aufmerksam geworden, in welchem ich wertvolle neue tontechnische und persönliche Erfahrungen erlebte. Weiteren Dank gilt meinem Zweitprüfer Professor Oliver Curdt, bei dem ich vor allem in meinen Studioproduktionen sehr viel lernen durfte. Mein Bachelorstudium an der HdM war eine tolle Zeit, in der ich - gerade in Ton- und Filmstudios - viel neues lernen und mich kreativ audiovisuell ausleben durfte.

Größten Dank gilt auch meiner Familie, Freunden und Kommilitonen für die große Unterstützung und die Kraft, die sie mir geschenkt haben. Ein besonderes Dankeschön möchte ich dabei meiner Lebensgefährtin Céline, meinen Eltern und meiner Schwester Sophia aussprechen, die durch ihre persönliche Unterstützung und Ermutigung ein großes Stück zum Gelingen dieser Bachelorarbeit beigetragen haben.

Inhaltsverzeichnis

Ehrenwörtliche Erklärung	I
Kurzfassung	II
Abstract	III
Danksagung	IV
Inhaltsverzeichnis	V
Abbildungsverzeichnis	VIII
1. Einleitung	1
1.1 Motivation und Aufbau.....	1
2. Mikrofonprinzipien	3
2.1 Schallwandlung.....	3
2.1.1 Elektrodynamische Wandler.....	3
2.1.2 Elektrostatische Wandler.....	4
2.2 Schallempfang.....	4
2.2.1 Druckempfänger.....	4
2.2.2 Druckgradientenempfänger.....	5
2.2.3 Einfluss des Empfängerprinzips auf die Richtcharakteristik.....	5
2.2.3.1 Nierencharakteristik.....	6
2.2.3.2 Breite Niere.....	6
2.2.3.3 Hyper- und Superniere.....	7
3. Psychoakustische Wahrnehmung	7
3.1 Richtungs- und Entfernungshören.....	8
3.1.1 Außenohrübertragungsfunktion.....	8
3.1.2 Schalleinfall in Horizontalebene.....	8
3.1.3 Schalleinfall in Medianebene.....	8
3.1.4 Entfernungshören.....	9
3.2 Phantomschallquellen.....	10
3.2.1 Zeitdifferenzen.....	10
3.2.2 Pegeldifferenzen.....	10
3.2.3 Zusammenwirken von Zeit- und Pegeldifferenzen.....	10
3.2.4 Interaurale Kohärenz.....	11

3.3 Stereofone Lautsprecherwiedergabe.....	12
3.3.1 Sweetspot.....	13
3.3.2 Schallquellenposition.....	13
4. Stereofone Aufnahmeverfahren.....	14
4.1 Hauptmikrofonverfahren.....	14
4.1.1 Laufzeitstereofonie.....	14
4.1.2 Intensitätsstereofonie.....	15
4.1.3 Äquivalenzstereofonie.....	17
4.2 Einzelmikrofonverfahren.....	17
4.2.1 Anwendungsbereiche.....	18
4.2.2 Kombination mit Hauptmikrofonierung.....	19
5. Immersive Lautsprecherwiedergabe.....	20
5.1 Konventionelle Lautsprecherlayouts nach ITU.....	20
5.1.1 Surroundwiedergabe mit 3/2 Kanälen (0+5+0).....	21
5.1.2 Surroundwiedergabe mit 5/2 oder 3/4 Kanälen (0+7+0).....	23
5.1.3 Surroundwiedergabe mit Elevation (4+5+0).....	24
5.1.4 Wiedergabe in 22.2 (9+10+3).....	25
5.2 Alternative Lautsprecherlayouts.....	27
5.2.1 Wellenfeldsynthese.....	27
5.2.2 Reflexions-/ Transauralbasierte Soundbar.....	29
5.3 Wiedergabe unterschiedlichen Inhalts.....	30
5.3.1 Wiedergabe objektbasierten Inhalts.....	30
5.3.2 Lautsprecherwiedergabe und Decodierung von Ambisonics 1. Ordnung.....	30
6. 3D Mikrofonanordnungen und ihre praktischen Einsatzbereiche.....	32
6.1 Taxonomie mehrkanaliger 3D Mikrofonanordnungen.....	33
6.1.1 Lautsprecherbezogen oder Wiedergabeunabhängig.....	33
6.1.2 Aufbau.....	35
6.1.3 Eigenschaften.....	36
6.1.4 Wahrnehmung.....	37
6.1.5 Anwendungsbereiche.....	38
6.2 Lautsprecherbezogene Mikrofonanordnungen.....	40
6.2.1 Mikrofonarrays mit Anordnung 4+4+0 zur Wiedergabe über Sound System D (9.1).....	40
6.2.1.1 Williams-Umbrella.....	40

6.2.1.2 Twin Cube.....	41
6.2.1.3 ORTF-3D.....	42
6.2.1.4 ESMA-3D.....	42
6.2.1.5 Lee Rec-3D.....	43
6.2.1.6 Hamasaki Cube.....	44
6.2.2 Mikrofonarrays mit Anordnung 4+5+0 zur Wiedergabe über Sound System D (9.1).....	45
6.2.2.1 OCT-3D.....	45
6.2.2.2 Bowles-Array.....	46
6.2.2.3 2L-Cube.....	47
6.2.2.4 Spider Tree.....	47
6.2.2.5 Twin Cube (4+4+0) mit Ergänzung um M+000.....	48
6.2.2.6 Double-UFIX.....	48
6.2.2.7 PCMA-3D.....	49
6.2.3 Mikrofonarrays mit Anordnung 9+10+3 zur Wiedergabe über Sound System H (22.2).....	50
6.2.3.1 Hamasaki et al.....	50
6.2.3.2 Howie et al.....	51
6.2.4 Mikrofonarrays zur Wiedergabe über spezielle Lautsprecheranordnungen...52	
6.2.4.1 Au3Dio (4+6+0).....	52
6.3 Wiedergabeunabhängige Mikrofonanordnungen.....	52
6.3.1 Aufnahme in First-Order Ambisonics.....	53
6.3.1.1 Triple MS.....	53
6.3.1.2 Tetrahedral Array.....	54
6.3.2 Aufnahme in Higher-Order Ambisonics.....	55
6.3.2.1 Spherical Array.....	55
7. Diskussion.....	55
7.1 Einfluss der psychoakustischen Wahrnehmung auf Mikrofonanordnungen.....	55
7.2 Gegenüberstellung: Einzel- und Hauptmikrofonverfahren.....	56
7.3 Gegenüberstellung: Portabilität und Richtungswahrnehmung.....	57
8. Fazit.....	58
Literaturverzeichnis.....	60

Abbildungsverzeichnis

Abbildung 1: Richtcharakteristiken von Kugel bis Acht in Darstellung auf einem Polar-koordinatensystem und in dreidimensionaler Abbildung (Weinzierl, 2008).....	6
Abbildung 2: Darstellung eines Kopfbezogenen Koordinatensystem und dessen Horizontal- und Medianebene (Weinzierl, 2008).....	9
Abbildung 3: Lage und Ausdehnung eines Hörereignisses in Abhängigkeit vom Kohärenzgrad (Weinzierl, 2008).....	12
Abbildung 4: Aufstellung zweier Lautsprecher in Stereo-Standardaufstellung mit Basiswinkel $\alpha = 60^\circ$ (Weinzierl, 2008).....	13
Abbildung 5: Darstellung einer laufzeitstereofonischen Mikrofonanordnung zweier Nieren (Weinzierl, 2008).....	15
Abbildung 6: XY-Stereofonieranordnung (Weinzierl, 2008).....	17
Abbildung 7: Darstellung einer nativen Aufnahme in MS Stereofonie(Zotter& Frank, 2019)..	17
Abbildung 8: Darstellung der äquivalenzstereofonen Anordnungen ORFT und NOS (Weinzierl, 2008).....	17
Abbildung 9: Darstellung des Vector Base Amplitude Panning am Beispiel eines Lautsprecherdreiecks (Weinzierl, 2008).....	21
Abbildung 10: Darstellung einer 3/2-Lautsprecheranordnung nach ITU-R BS.775-1.....	23
Abbildung 11: Darstellung einer 3/4-Lautsprecheranordnung nach ITU-R BS.775-1.....	24
Abbildung 12: Darstellung der Anordnung nach Sound System D nach ITU-R BS.2051-2.....	25
Abbildung 13: Darstellung der Anordnung nach Sound System H nach ITU-R BS.2051-2.....	26
Abbildung 14: Alle möglichen Lautsprecherkanäle des Middle, Upper und Lower Layer nach ITU-R BS.2051-2.....	26
Abbildung 15: Abbildung des Huygens-Fresnel-Prinzip und der Wellenfeldsynthese (Weinzierl, 2008).....	29
Abbildung 16: Darstellung einer Decodierung von 2D-Ambisonicssignalen erster Ordnung auf symmetrische Lautsprecheranordnung (Zotter& Frank, 2019).....	31
Abbildung 17: Drauf- und Seitenansicht einer Anordnung des Williams Umbrella (Lee, 2021).....	40

Abbildung 18: Drauf- und Seitenansicht einer Anordnung des Twin Cubes (Lee, 2021).....	41
Abbildung 19: Drauf- und Seitenansicht der Anordnungen nach ORTF-3D in Indoor- und Outdoor-Version (Lee, 2021).....	42
Abbildung 20: Drauf- und Seitenansicht einer Anordnung des ESMA-3D (Lee, 2021).....	43
Abbildung 21: Drauf- und Seitenansicht von Anordnungen des Lee Rec-3D (Lee, 2021).....	44
Abbildung 22: Drauf- und Seitenansicht einer Anordnung des Hamasaki Cubes (Lee, 2021).....	45
Abbildung 23: Drauf- und Seitenansicht einer Anordnung des Twin Cubes (Lee, 2021).....	46
Abbildung 24: Drauf- und Seitenansicht einer Anordnung des Bowles Array (Lee, 2021).....	46
Abbildung 25: Drauf- und Seitenansicht einer Anordnung des 2L-Cubes (Lee, 2021).....	47
Abbildung 26: Drauf- und Seitenansicht einer Anordnung des Spider Tree (Lee, 2021).....	48
Abbildung 27: Drauf- und Seitenansicht einer Anordnung des Double-UFIX (Lee, 2021).....	49
Abbildung 28: Drauf- und Seitenansicht einer Anordnung des PCMA-3D (Lee, 2021).....	50
Abbildung 29: Drauf- und Seitenansicht von Anordnungen nach Howie et al. (Lee, 2021).....	51
Abbildung 30: Drauf- und Seitenansicht einer Anordnung des au3Dio (Lee, 2021).....	52
Abbildung 31: Darstellung einer Triple-MS-Anordnung (Zotter & Frank, 2019).....	53
Abbildung 32: Darstellung einer tetraederartigen Anordnung und ihrer Codierung in's Ambisonics B-Format (Zotter& Frank, 2019).....	54

1 Einleitung

1.1 Motivation und Aufbau

In Fachdiskussionen innerhalb der Audioszene ist in den letzten Jahren ein Wort aus dem Sprachgebrauch von Tonmeistern und Toningenieuren nicht mehr wegzudenken: 3D-Audio (auch immersive Audio). Es verspricht einem Hörer die auditive Wahrnehmung von Breite, Tiefe und Höhe und somit das Gefühl einer dreidimensionalen Umgebung aus Schall, welche den Hörer rundum umgibt und ihm ein lebendiges räumliches Klangerlebnis bereitet. An Popularität gewinnt das Thema auch durch eine zunehmende Verbreitung immersiver Wiedergabetechniken in etlichen Bereichen audiovisueller Medien, welche es durch Soundbars oder Kopfhörer auch bis ins Wohnzimmer der Endkonsumenten geschafft haben. Die auditive Erweiterung um eine weitere Dimension stellt dabei keine Errungenschaft der letzten Jahre dar, sondern hat sich in audiovisuellen Bereichen der Filmvertonung und ihren kommerziellen Lösungen (Dolby Atmos, AURO 3D und DTS:X) bereits etabliert. Auch in Anwendungen von Virtual oder Augmented Reality interaktiver Medien hat sich seit längerem der Einsatz von 3D-Audio in Form einer Binauralsynthese weitgehend durchgesetzt. Doch auch die flexible Anwendung in etlichen anderen Bereichen der audiovisuellen Medien, wie in Musik- oder Hörspielproduktionen, ist seit einiger Zeit auf dem Vormarsch.

Geht es in der allgemeinen Verbreitung von 3D Audio oft hauptsächlich um die optimale Wiedergabe in bestimmten Anwendungsbereichen, so stellt sich für den Tonschaffenden die Frage, wie der auditive Inhalt einer immersiven Wiedergabe produziert werden kann. Abhängig des Wiedergabeformates werden häufig in der Postproduktion einzelne Audiosignale räumlich verteilt und verhallt in einer 3D-Audio Mischung verarbeitet, wohingegen die Aufnahme zumeist im Stereo-Haupt- oder Einzelmikrofonverfahren durchgeführt wurde. Zur immersiven Aufnahme im Hauptmikrofonverfahren von Musik und Atmosphären wurden seit den frühen 2000er-Jahren durch Forscher und Toningenieure etliche Mikrofonanordnungen vorgestellt, welche in äquivalenten Einsatzbereichen zu Stereo- oder Surround-Hauptmikrofonen ihre Verwendung finden. Doch fehlt bis heute vielen Tonschaffenden ein Überblick über vorhandene Möglichkeiten

der 3D-Audio-Aufnahme und deren Eigenschaften und Einsatzbereiche, wonach einige wenige Mikrofonanordnungen einen deutlich häufigeren Einsatz erleben als andere. Aus dieser Problemstellung ergibt sich folgende Forschungsfrage:

Wie können 3D-Mikrofonanordnungen hinsichtlich der Wiedergabesysteme, Anwendungsbereiche und Eigenschaften kategorisiert werden?

Das Ziel dieser Thesis ist daher, einen Aufschluss über unterschiedliche Aufnahmemöglichkeiten für 3D-Audio in einer sinnvollen Kategorisierung zu geben, welche dem Leser Unterstützung zu individuellen Anwendungsentscheidungen leisten soll.

Begründet wird dies auf Grundlagen der Tontechnik, wonach Kapitel 2 sich mit den Mikrofonprinzipien der Schallwandlung und des Schallempfangs auseinandersetzt. Kapitel 3 beschäftigt sich mit psychoakustischen Grundlagen des Richtungs- und Entfernungshörens, der Phantomschallquellen und der stereofonen Lautsprecherwiedergabe. Stereofone Aufnahmeverfahren der Haupt- und Einzelmikrofonierung werden in Kapitel 4 behandelt. Kapitel 5 führt die immersive Lautsprecherwiedergabe durch konventionelle Lautsprecherlayouts nach ITU, über alternative Lautsprecherlayouts und die Wiedergabe unterschiedlichen Inhaltes ein. 3D Mikrofonanordnungen und ihre praktischen Einsatzbereiche werden in Kapitel 6 behandelt, welches sich mit deren Taxonomie und einer Beschreibung lautsprecherbezogener und wiedergabeunabhängiger Arrays beschäftigt. Abschließend werden in Kapitel 7 und 8 gewonnene Erkenntnisse diskutiert und ausgewertet. Eine Auswertung erfolgt auch in einer der Thesis anhängenden Tabelle, welche eine Kategorisierung und Bewertung der Mikrofonanordnungen bezüglich ihrer Eigenschaften, Anwendungsbereiche und Wiedergabemöglichkeiten geben soll.

2 Mikrofonprinzipien

Zwei Wandler sind im Grunde genommen Bestandteil nahezu jedes Schallwandlers, so Görne (2014). Er führt weiter aus: Als akustisch-mechanischer Wandler fungiert eine bewegte Membran, welche Schalldruck in mechanische Schwingungen umsetzt, woraufhin diese durch einen elektromagnetischen Wandler in elektrische Spannung umgesetzt werden. Ziel dieser Wandlung von Schallschwingungen in elektrische Wechselspannungen ist eine genaue Abbildung des akustischen Signals als elektrisches Signal.

2.1 Schallwandlung

Als Wandlerprinzip gilt die Umwandlung mechanischer in elektrische Schwingungen (Dickreiter, Dittel, Hoeg & Wöhr, 2008). Unterschieden wird prinzipiell zwischen elektrodynamischen Wandlern (Dynamisches Mikrofon) und elektrostatischen Wandlern (Kondensatormikrofon), welche in den folgenden Abschnitten aufgezeigt werden.

2.1.1 Elektrodynamische Wandler

Als Wandler eines dynamischen Mikrofons fungiert ein durch eine Membran bewegter Leiter, welcher sich zwischen den zwei Magnetpolen eines Dauermagneten in einem homogenen Magnetfeld befindet (Dickreiter, 2011). Jeder elektrodynamische Wandler ist dabei ein sogenannter Schnellewandler, bei dem die induzierte Spannung von der Bewegungsgeschwindigkeit der Membran abhängt. Unterschieden wird zwischen zwei Möglichkeiten der Spannungsinduzierung:

Beim Bändchenmikrofon fungiert eine Membran aus zwei gefalteten, direkt nebeneinanderliegenden und hauchdünnen Aluminiumbändchen zeitgleich als bewegte Spule (Dickreiter, 2011), welche als Leiter im Magnetfeld und somit im Luftspalt des Dauermagneten angeordnet sind (Görne, 2014).

Auch das Tauchspulenmikrofon funktioniert als elektrodynamischer Wandler, wobei als Leiter eine mit der Membran starr verbundene Spule fungiert (Dickreiter, 2011). Die Spule taucht dabei in der Frequenz der Membranschwingung in den Luftspalt eines Topfmagneten ein und induziert so eine Spannung (Dickreiter et al., 2008).

Zur Unterdrückung von Störspannungen durch magnetische Störfelder wird zumeist eine entgegengewickelte Kompensationsspule mit der Schwingspule in Reihe geschaltet.

2.1.2 Elektrostatische Wandler

Ein Elektrostatischer Wandler in Form eines Kondensatormikrofons ist ein Kondensator, welcher als Elektrode eine feste Platte und als Gegenelektrode eine bewegliche Membran führt (Dickreiter et al., 2008). Durch die Phantomvorspannung von zumeist 48V erhält der Kondensator eine konstante Ladung, sodass Schwingungen der Membran durch ein periodisches Verändern des Elektrodenabstandes eine Kapazitätsänderung hervorrufen (Dickreiter et al., 2008). Unterschieden wird bei Kondensatormikrofonen zwischen zwei Schaltungsvarianten: Der Hochfrequenz- und der Niederfrequenzschaltung. Als besondere Bauform eines Kondensatormikrofons gilt das Elektretmikrofon, welches als preiswerte Alternative lediglich eine Batterie zur Spannungsversorgung benötigt (Dickreiter et al., 2008).

2.2 Schallempfang

Das Empfängerprinzip des Mikrofons ist das Prinzip der Umwandlung akustischer Schwingungen in mechanische (Dickreiter et al., 2008). Die Membran kann in einer akustisch-mechanischen Konstruktion auf unterschiedliche Arten dem Schallfeld ausgesetzt sein (Görne, 2015), was maßgeblich die Richtcharakteristik eines Mikrofons bestimmt (Dickreiter, 2011). In einer Mikrofonkapsel befindet sich die eingespannte Membran, wobei die mechanische Bewegung durch unterschiedliche Bauformen - offen, in einem halboffenen oder geschlossenen Becher - in einem gewissen Schema erzwungen wird (Görne, 2015).

Unterschieden wird zwischen den verschiedenen Bauformen einer Mikrofonkapsel, dem Druckempfänger und dem Druckgradientenempfänger.

2.2.1 Druckempfänger

Das einfachste physikalische Prinzip eines Schallempfängers in einer Mikrofonkapsel ist der Druckempfänger (Görne, 2015). Eine elastisch eingespannte Membran wird als Teil eines luftdichten Bechers durch Druckschwankungen einer Schallwelle zum Schwingen gebracht (Dickreiter, 2011). Die Membran fungiert dabei als Abdichtung des Bechers, sodass dieser vollständig verschlossen wird. Somit findet eine schalldichte Abtrennung

zwischen dem Inneren der Kapsel und dem umgebenen Raum statt (Dickreiter et al., 2008). Dank einer kleinen Öffnung findet ein Druckausgleich statt, sodass im Inneren der Kapsel Normaldruck herrscht (Görne, 2015). Der Druckausgleich gilt dabei allerdings nicht für die raschen Druckschwankungen des Schalls, sodass es immer dann zu einer Membranbewegung kommt, sobald der Außendruck an der Membran vom Innendruck, dem Normaldruck in der Kapsel, abweicht (Dickreiter et al., 2008). Da sich durch die Beugung von Schallwellen aus rückwärtiger oder seitlicher Richtung ebenfalls Druckänderungen vor der Membran ergeben, nimmt der Druckempfänger diese aus jeglichen Richtungen in gleicher Intensität auf, was zu einer ungerichteten (kugelförmigen) Richtcharakteristik führt (Dickreiter et al., 2008). Mit einem Druckempfänger als Kondensatormikrofon lassen sich dank des linearen Frequenzgangs beliebig tiefe Frequenzen aufnehmen, allerdings existieren in hohen Frequenzen Einschränkungen in der ungerichteten Richtcharakteristik durch Druckstau vor der Membran (Görne, 2015).

2.2.2 Druckgradientenempfänger

Der Druckgradientenempfänger verfolgt in seiner Mikrofonkapsel ein anderes physikalisches Konzept, nämlich das einer im Schallfeld freien Membran (Görne, 2015), welche somit beidseitig dem Schalldruck ausgesetzt ist. Maßgeblich für die Auslenkung der Membran verantwortlich ist somit die Druckdifferenz zwischen Vorder- und Rückseite der Membran, sodass Schall senkrecht von vorne oder von hinten für eine maximale Auslenkung sorgt, wohingegen Schall von der Seite keine Differenz in Druck hervorruft (Dickreiter et al., 2008). Schalldruck von vorne und von hinten hat zeitlich und räumlich prinzipiell den gleichen Verlauf, allerdings eine entgegengesetzte Phase (Görne, 2015). In ein Polardiagramm eingetragen ergibt sich für den klassischen Druckgradientenempfänger somit die Richtcharakteristik einer Acht, welche in ihrem Übertragungsbereich weitgehend als frequenzunabhängig gilt (Dickreiter et al., 2008).

2.2.3 Einfluss des Empfängerprinzips auf die Richtcharakteristik

Die Empfindlichkeit von Mikrofonen verschiedener Empfängerprinzipien bezogen auf den Schalleinfallswinkel bezeichnet man als Richtungsmaß, welches als Darstellung aller Richtungen in einem Polarkoordinatensystem die Richtcharakteristik zeigt (Dickreiter, 2011). Die Formen einer Kugel und einer Acht gelten als Grundformen und gleichzeitig als Extreme der Richtcharakteristiken, so Görne (2015). Er beschreibt weiter, dass ohne weitere Hilfsmittel nicht mehr Richtwirkung (Acht) als mit dem offenen Gradientenempfänger und

nicht weniger Richtwirkung (Kugel) als mit dem Druckempfänger möglich wäre. Doch sind auch viele Richtcharakteristiken zwischen Kugel und Acht möglich, vor allem durch den Einsatz von Laufzeitgliedern, welche sich in Druckgradientenempfängern wiederfinden und Schall aus bestimmten Richtungen in bestimmter Weise verzögern sollen. Auch die Kombination verschiedener Richtcharakteristiken führt zu einer neuartigen Richtwirkung.

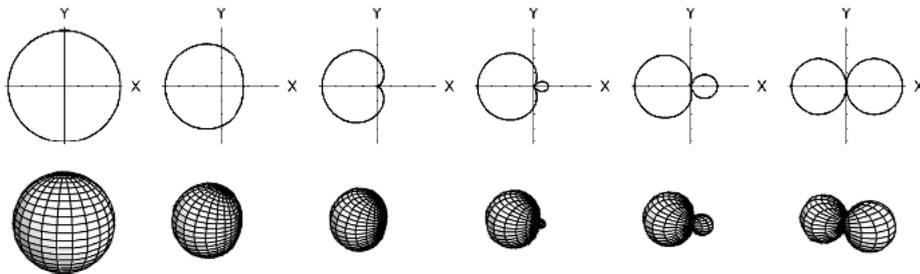


Abb. 1. Richtcharakteristiken von Kugel bis Acht in Darstellung auf einem Polarkoordinatensystem und in dreidimensionaler Abbildung

2.2.3.1 Nierencharakteristik

Die Nierencharakteristik gehört zu den meist verwendeten Richtwirkungen heutiger Studiomikrofone. Sie kann dabei auf zwei verschiedene Arten erreicht werden. Zumeist wird diese durch die Kombination eines klassischen Druckgradientenempfängers mit der Richtwirkung einer Acht und einem Laufzeitglied erreicht. Nutzt man das Laufzeitglied zur Verzögerung des rückwärtig einfallenden Schalls, sodass dieser zeitgleich die Vorder- und Rückseite der Membran erreicht, so findet keine Membranauslenkung statt und das Mikrofon hat für Schall aus dieser Richtung keine Empfindlichkeit (Dickreiter, 2011). Die Richtcharakteristik der Niere ist auch durch Kombination von Druckempfänger und Druckgradientenempfänger möglich, solange diese möglichst nah beieinander angeordnet sind. Legt man folgend Kugel- und Achtercharakteristik durch Addition übereinander, ergibt sich im Polarkoordinatensystem eine Nierencharakteristik (Dickreiter et al., 2008). Durch die Addition zweier Nieren in entgegengesetzter Richtung ergibt sich eine Kugel-, durch Subtraktion eine Achter- und durch Subtraktion einer kleinen Niere eine Hypernierencharakteristik (Dickreiter, 2011).

2.2.3.2 Breite Niere

Die Richtcharakteristik der breiten Niere gilt als Zwischenform zwischen Niere und Kugel, wobei sich ihr Einsatzgebiet eher als Alternative zur Kugel mit einer leicht größeren Richtwirkung wiederfindet (Dickreiter, 2011). Auch die breite Niere ist durch Addition von

Kugel- und Achtercharakteristik unter Berücksichtigung eines zu multiplizierenden Faktors möglich. Beim Einsatz eines Laufzeitgliedes in der Kapsel eines Druckgradientenempfängers ist die Verzögerung dementsprechend kleiner, sodass rückwärtig eintreffender Schall nicht zu keiner, sondern zu einer kleinen Auslenkung der Membran und so zu einer Dämpfung führt.

2.2.3.3 Hyper- und Superniere

Da möglichst gerichtete Mikrofontypen in ihrer Richtcharakteristik vornehmlich zwischen Niere und Acht liegen, kann die Superniere Schalleinfall von vorne (um die 0°) besser hervorheben, bringt dabei allerdings auch weniger Dämpfung für rückwärtig einfallenden Schall mit sich, so Dickreiter et al. (2008). Er führt weiter aus: Die Charakteristik einer Superniere lässt sich durch Einsatz eines Laufzeitgliedes im Druckgradientenempfänger mit geringerer Verzögerung oder durch Überlagerung von Kugel- und Achtercharakteristik mit einem größeren Anteil der Acht realisieren. Die Hyperniere hat die gleiche Bauform wie die Superniere und unterdrückt dabei den Diffusschall am Besten in Bezug auf alle Nierencharakteristiken (Görne, 2015). Dank der starken Abschattung des Diffusschalls finden sich häufig Super- und Hyperniere in der Bühnenanwendung wieder (Görne, 2015). Eine noch stärkere Richtwirkung ist nur durch noch größere Laufzeitbeeinflussung in Form eines Richtrohres möglich und findet seine Anwendung so gut wie ausschließlich im Filmton.

3 Psychoakustische Wahrnehmung

Unsere räumliche Wahrnehmung sei binaural, so Görne (2015). Er spricht damit den Zusammenhang unserer beiden Ohren mit der wahrgenommenen Einfallsrichtung und -distanz, der räumlichen Ausdehnung und der spektralen Frequenzwahrnehmung an. Durch diesen Zusammenhang ergeben sich starke Wahrnehmungsunterschiede, zum Beispiel bei Schallereignissen aus horizontaler oder vertikaler Ebene und bei der Lokalisation dieser. Diese Erkenntnisse lassen sich optimal vorerst auf eine stereofone und später auf eine immersive Lautsprecherwiedergabe und auf Aufnahmemöglichkeiten für diese übertragen.

3.1 Richtungs- und Entfernungshören

3.1.1 Außenohrübertragungsfunktion

Die genannten Unterschiede zwischen beiden Ohrsignalen, abhängig von der Einfallrichtung und -distanz, beschreibt mathematisch die Außenohrübertragungsfunktion (HRTF). Die HRTF beschreibt dabei die spektralen Unterschiede von Hörereignissen unterschiedlicher Einfallrichtung untereinander in Form von Kammfiltern, beeinflusst durch die Abschattung durch Kopf und Schulter und durch die Form des Gehörganges. Durch die Anwendung einer Fourier-Transformation auf die mathematische Funktion der HRTF erhält man die HRIR, die kopfbezogene Impulsantwort, welche, abhängig von der Einfallrichtung, die Laufzeitunterschiede zwischen beiden Ohrsignalen beschreibt. Durch den Vergleich der Eigenschaften genannter Ohrsignale in Frequenz und Laufzeit, abhängig der Schalleinfallrichtung, kann vor allem auf eine unterschiedliche Richtungswahrnehmung in horizontaler und vertikaler Ebene geschlossen werden.

3.1.2 Schalleinfall in Horizontalebene

Zur Orientierung in unserer Umgebung ist unsere räumliche Wahrnehmung auf das Richtungshören in der Horizontalebene, der Azimutebene, ausgelegt (Görne, 2015). Vergleicht man dabei die HRTF beider Ohrsignale bei Schallquellenpositionen frontal und seitlich, so sind sowohl Laufzeit- als auch spektrale Unterschiede zu erkennen, welche bei der Richtungswahrnehmung zusammenwirken (Dickreiter, 2011). Die Lokalisationsschärfe ist bei Schalleinfall aus horizontaler Ebene am größten, wenngleich diese auch von der Art des Signals abhängt, so Görne (2015). Er beschreibt die genaueste Lokalisation um etwa $\pm 1^\circ$ bei breitbandigen und impulsartigen Signalen und die am wenigsten genaue Lokalisation bei schmalbandigen Signalen mit etwa $\pm 5^\circ$ bis $\pm 10^\circ$ Lokalisationsschärfe.

3.1.3 Schalleinfall in Medianebene

Befindet sich eine Schallquelle in der Elevation, also höher oder tiefer der interauralen Achse, so verändern sich Laufzeit- und Pegeldifferenzen nicht wesentlich (Görne, 2015). Durch die geringen Unterschiede der beiden Ohrsignale zueinander bieten sich zur Schallquellenlokalisierung fast ausschließlich gelernte Klangfärbungsunterschiede an, abhängig des Einfallswinkels und der Form des Kopfes und der Gehörgänge (Dickreiter et

al., 2008). Somit hängt die Lokalisationsschärfe unter Anderem von der Übung des Gehörs und der Bekanntheit des Hörereignisses ab und stellt sich im Vergleich zur Horizontalebene als deutlich ungenauer dar: Bei einem unbekanntem Sprecher liegt sie zum Beispiel bei $\pm 15^\circ$ bis $\pm 20^\circ$, bei einem bekannten Hörereignis bei $\pm 10^\circ$ und bei weißem Rauschen bei $\pm 4^\circ$, so Dickreiter (2011).

3.1.4 Entfernungshören

Zur Auswertung der Entfernung einer Schallquelle bleiben für unsere Wahrnehmung keine einfachen Ohrsignalmerkmale mehr übrig, sodass hierfür die Erfahrung unseres Gehörs und der Bekanntheitsgrad eines Hörereignisses essentiell sind, so Dickreiter et al. (2008). Im freien Schallraum gelten die Pegelverhältnisse und die Dämpfung höherer Frequenzen als wichtigste Unterscheidungsmerkmale der Entfernung zur Schallquelle (Dickreiter, 2011). Bei kleineren Entfernungen führen vor allem Abschattungen und Beugungen durch den Kopf zu Klangfärbungsunterschieden (Dickreiter et al., 2008).

Anders verhält sich die Entfernungswahrnehmung innerhalb geschlossener Räume, welche sich durch die Auswertung des Verhältnisses von Direkt- und Diffusschall wesentlich verbessert (Dickreiter, 2011). Zur Entfernungswahrnehmung wichtig ist sowohl das Pegelverhältnis des Direktschalls zum Diffusschall, als auch die Nachhallzeit, weswegen auch in geschlossenen Räumen eine gewisse Hörerfahrung in der momentanen Raumakustik notwendig ist (Dickreiter et al., 2008). Da die Entfernungsabschätzung allerdings recht ungenau und sehr individuell ist, wird die Schallquelle meist näher eingeschätzt, als sie tatsächlich ist (Dickreiter, 2011).

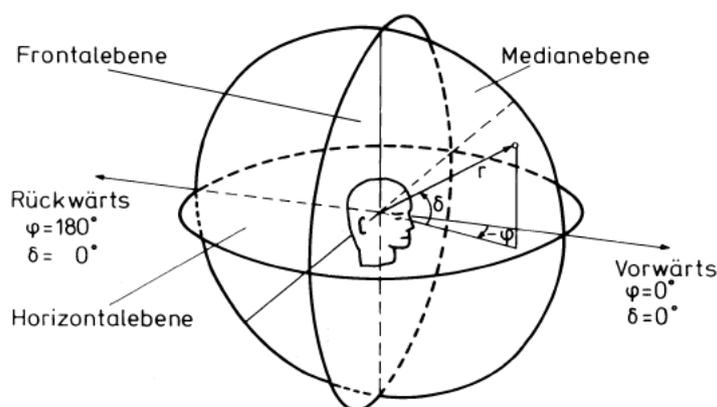


Abb 2. Darstellung eines kopfbezogenen Koordinatensystem und dessen Horizontal- und Medianebene

3.2 Phantomschallquellen

Eine Phantomschallquelle ist eine virtuelle Schallquelle, welche in ihrer räumlichen Wahrnehmung nur durch die Überlagerung mehrerer Schallereignisse aus unterschiedlichen Richtungen als Hörereignis existiert (Görne, 2015). Dabei entsteht ein frontal gelegenes Hörereignis bei stereofoner Lautsprecherwiedergabe, bei gleichzeitiger und gleichlauter Abstrahlung des gleichen Signals auf beiden Lautsprechern (Dickreiter, 2011). Soll das Hörereignis weiter auf einer Seite angeordnet sein, so ist dies durch zunehmende Pegel- und/oder Laufzeitunterschiede zu erreichen (Dickreiter, 2011). Gleiches gilt auch aufnahmeseitig, was zu den Aufnahmeprinzipien der Koinzidenz- und Laufzeitstereofonie führt.

3.2.1 Zeitdifferenzen

Zur räumlichen Ortung von Phantomschallquellen ist der Präzedenzeffekt, auch Gesetz der ersten Wellenfront genannt, essentiell, der die wahrgenommene Richtung als solche beschreibt, aus welcher der Schall zuerst den Kopf erreicht (Görne, 2015). Ist bei zwei Lautsprechern die zeitliche Verzögerung allerdings kleiner als 0,6 bis 1ms, so findet lediglich eine Verschiebung des Hörereignisses in Richtung der ersten Schallabstrahlung statt (Görne, 2015).

3.2.2 Pegeldifferenzen

Die gleiche Verschiebung der Phantomschallquelle zwischen den Lautsprechern kann auch durch Änderung der Pegelverhältnisse geschehen, wobei ab 0,6 bis 2dB eine Bewegung der Phantomschallquelle in Richtung des größeren Pegels wahrnehmbar ist und ab etwa 10 bis 20dB das Hörereignis räumlich auf dem Lautsprecher liegt, so Görne (2015).

3.2.3 Zusammenwirken von Zeit- und Pegeldifferenzen

Die Intensitätsstereofonie beschreibt zur Darstellung von Phantomschallquellen lediglich die Anwendung von Pegelunterschieden, die Laufzeitstereofonie nur die Anwendung von zeitlichen Unterschieden. In der Aufnahmetechnik ist zudem die Kombination beider Verfahren möglich, sodass sich Wirkung von Pegel- und Laufzeitdifferenzen zu einer gemeinsamen Wirkung addieren (Dickreiter, 2011). Sind bei mehrkanaligen Aufnahmeverfahren Richtmikrofone unterschiedlich ausgerichtet, Mikrofone distanziert angeordnet oder Trennkörper im Einsatz, spricht man von gemischter Aufnahme aus

Pegel- und Laufzeitdifferenzen gleichsinnig (Dickreiter et al., 2008). In gewissen Anordnungen kann so eine bessere räumliche Darstellung und eine bessere Lokalisationsschärfe der Hörereignisse erreicht werden.

3.2.4 Interaurale Kohärenz

Essentiell für die auditive räumliche Wahrnehmung gilt die Berücksichtigung der Ähnlichkeit beider Ohrsignale, so Weinzierl (2008). Er führt weiter aus: Das Ähnlichkeitsmaß wird häufig mit der normierten Kreuzkorrelationsfunktion der Ohrsignale erfasst, wobei die Ohrsignale als Leistungssignale mit endlicher Leistung betrachtet werden. In der Psychoakustik wird oftmals der Begriff des interauralen Kohärenzgrades verwendet, welcher ausschließlich Werte zwischen Null und Eins annehmen kann, wobei Null (0) als völlige Unähnlichkeit und Eins (1) als völlige Ähnlichkeit beider Ohrsignale gilt (Weinzierl, 2008). Ein hoher Kohärenzgrad spricht für eine relativ hohe Lokalisationsschärfe des Hörereignis, welches in der Mitte des Kopfes wahrgenommen wird, wohingegen ein kleiner Kohärenzgrad zu einem Zerfall des Hörereignis in zwei Signalanteile führt, welche sich nun jeweils einer Schallquelle zuordnen lassen (Weinzierl, 2008). Die Interaurale Kohärenz gilt somit als Maß der räumlichen Ausdehnung eines Hörereignisses in der psychoakustischen Wahrnehmung.

Die Ähnlichkeit beider Ohrsignale hängt im Wesentlichen von der Art des Schallfeldes, in dem sich ein Hörer bewegt, und von Signaleigenschaften und Abstrahlverhalten der Schallquelle ab. Doch besitzen selbst im freien Schallfeld Ohrsignale nie eine völlige Inkohärenz, da schon allein durch Übersprechen zwischen den Ohren teilweise gleiche Signale beide Ohren erreichen (Weinzierl, 2008). Ein diffuses Schallfeld sorgt durch Reflexion des Schalls, zum Beispiel an Wänden, für stark unähnliche Ohrsignale des Hörers im Raum. Diese Inkohärenz führt in der auditiven räumlichen Wahrnehmung zum Eindruck des „Eingehülltseins“, so Weinzierl (2008). Er beschreibt folgend, dass unser Gehör allerdings bei teilkohärenten Signalen stark kohärente Anteile herausfinden und diese jeweils scharf lokalisieren könne, so zum Beispiel auch einen Sprecher in lärmerfüllter Umgebung. Durch die Verwendung stärker oder weniger stark kohärenter Signale in einzelnen Kanälen in Aufnahme- oder Wiedergabesituationen, lässt sich folglich die wahrgenommene Breite der Aufnahme oder Wiedergabe beeinflussen. Aufnahmeseitig lässt sich die Ähnlichkeit der einzelnen Mikrofonsignale durch die Positionierung und

Ausrichtung der einzelnen Mikrofone und somit auch aus Stereo-Perspektive durch die Wahl des Aufnahmeverfahrens der Koinzidenz- oder Laufzeitstereofonie beeinflussen.

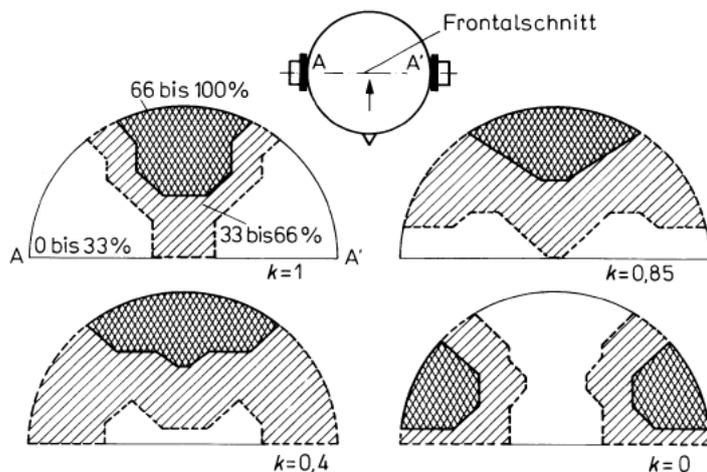


Abb 3. Lage und Ausdehnung eines Hörereignisses in Abhängigkeit vom Kohärenzgrad

3.3 Stereofone Lautsprecherwiedergabe

Das akustische Ziel, ein Schallfeld durch den Einsatz elektroakustischer Mittel in einen anderen Abhörraum zu übertragen, gilt schon lange als grundlegender Hintergrund einer mehrkanaliger Lautsprecherwiedergabe (Kuttruff, 2004). Verfolgt man das Prinzip allerdings konsequent, so bräuchte es möglichst viele sich auf einer Fläche befindliche Mikrofone, um das ursprüngliche Schallfeld abzutasten und auf der anderen Seite gleich viele Lautsprecher im Wiedergabesaal (Kuttruff, 2004). Vorig beschriebenes zeigt eine Wellenfeldsynthese auf, bei welcher möglichst viele Lautsprecher rund um das Auditorium auf Ohrhöhe angebracht werden und aufnahmeseitig mit gleicher Kanalanzahl und ähnlicher Anordnung abgetastet wird. Von diesem Prinzip der immersiven Beschallungs- und Aufnahmemöglichkeiten lassen sich vereinfachte Wiedergabesysteme und äquivalente Aufnahmesetups ableiten, wobei die konventionelle Stereofonie mit zwei Kanälen, welche unsere beiden Ohren repräsentieren sollen, das absolute Minimum darstellt (Kuttruff, 2004). Doch bei konventioneller Stereoanordnung mit zwei Lautsprechern bilden diese ein gleichseitiges Dreieck mit dem Hörer, was zu wesentlich anderen Verhältnissen führt als unser Hören im natürlichen Schallfeld (Dickreiter et al., 2008). Zu berücksichtigen sind dabei vor allem die optimale Hörposition und die Positionierung der Schallquellen zwischen den Lautsprechern.

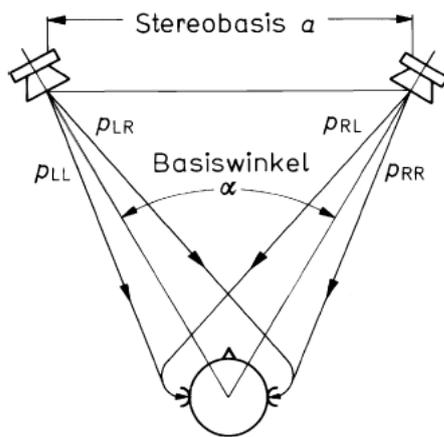


Abb. 4. Aufstellung zweier Lautsprecher in Stereo-Standardaufstellung mit Basiswinkel $\alpha = 60^\circ$

3.3.1 Sweetspot

Eine starke Einschränkung der stereofonen Lautsprecherwiedergabe ist die Begrenzung der optimalen Hörposition auf genau einen Punkt, welcher als sogenannter Sweetspot bezeichnet wird. Eine Erweiterung dieser optimalen Hörposition ist mit dem Einsatz von nur zwei Lautsprechern technisch nicht zu schaffen, sodass in den meisten Anwendungsbereichen die stereofone Beschallung zumindest aus technischer Perspektive nur einem Bruchteil des Auditoriums zufriedenstellende Ergebnisse liefern kann. Bewegt sich ein Hörer aus dem Sweetspot des Wiedergabesystems, so verschieben sich Phantomschallquellen und der Richtungseindruck verändert sich, was sich mit einer Pegel- oder Laufzeitänderung durch eine Entfernungs- und Richtungsverschiebung des Lautsprechers zum Hörer erklären lässt.

3.3.2 Schallquellenposition

Die Position einer Schallquelle in der stereofonen Wiedergabe kann sich entweder auf einem Lautsprecher oder auf einer Linie zwischen beiden Lautsprechern befinden, so beschreibt das Vector Base Amplitude Panning (kurz VBAP). Befindet sich die Schallquelle auf dem Lautsprecher, so spielt nur dieser das Hörereignis ab. Ist die Schallquelle allerdings mittig oder mittig-seitlich angeordnet, so ist die Positionierung des Hörereignisses einer Phantomschallquelle zuzuordnen, welche sich aus dem Pegel- oder Laufzeitverhältnis beider Lautsprecher zueinander ergibt. Von der Hörerposition werden dabei zwei Vektoren in Richtung der Lautsprecherpositionen gezogen, von welchen zwei weitere

Vektoren in Richtung des Hörereignisses zeigen. Aus dem Verhältnis letzterer Vektoren zueinander ergeben sich zu verrechnende Pegelfaktoren für beide Lautsprecher.

4 Stereofone Aufnahmeverfahren

Stereofone Aufnahmeverfahren haben die Absicht, ein Stereoklangbild des Aufnahmeraumes auf zwei Kanälen festzuhalten, um dem Hörer ein solches Klangbild zum Beispiel über Lautsprecher-Stereofonie wiederzugeben. Sie sind dabei allerdings nicht dazu ausgelegt, das Stereoklangbild eines jeden Hörers zu reproduzieren, sondern sind räumlich klar bezogen auf den Wiedergaberaum, was bei stereofoner Lautsprecherwiedergabe vom Hörer eine feste Position im Sweetspot und Kopfausrichtung zu den Lautsprechern erfordert (Dickreiter, 2011). Wird klassischerweise ein Mikrofonpaar zur Abbildung beider Kanäle verwendet, so spricht man von Hauptmikrofonverfahren, sind mehrere einzeln verteilte Mikrofone im Einsatz von Einzelmikrofonverfahren (Dickreiter, 2011).

4.1 Hauptmikrofonverfahren

Bei der Mikrofonierung mit Stereo-Hauptmikrofon, bestehend aus einem Mikrofonpaar, werden je nach Aufnahmeprinzip wiedergabeseitig entweder beide Aufnahmekanäle direkt Lautsprecherkanälen zugeordnet oder es sind Zwischenschritte der Codierung und Decodierung notwendig. Dies bringt verschiedene technische Anforderungen an die Aufnahmetechnik und Anordnung der Mikrofone mit sich und führt zu unterschiedlichen Klangeigenschaften der Aufnahmen, was verschiedene Prinzipien für verschiedene Aufnahmesituationen geeigneter oder ungeeigneter erscheinen lässt. Aus den Erkenntnissen der psychoakustischen Wahrnehmung bezüglich dem Einfluss von Pegel- und Laufzeitdifferenzen auf die Darstellung von Phantomschallquellen, lassen sich genannte Hauptmikrofonverfahren mit entsprechenden Mikrofonanordnungen ableiten: Die Laufzeit- und die Intensitätsstereofonie.

4.1.1 Laufzeitstereofonie

Sind zwei gleiche Mikrofone parallel und zur Schallquelle gerichtet aufgebaut, wobei der Abstand zu dieser größer ist als der Abstand beider Kapseln zueinander, so spricht man von Laufzeitstereofonie (Görne, 2015). Das Prinzip der reinen Laufzeitstereofonie ist dabei durch die Erkenntnis begründet, dass bei der Wiedergabe über zwei Lautsprecher ein

Hörereignis durch zeitliches Verzögern eines Kanals räumlich näher dem anderen Lautsprecher angeordnet ist. So gilt das Prinzip auch aufnahmeseitig in anderer Reihenfolge: Ist eine Schallquelle auf der Seite des einen Mikrofons angeordnet und somit näher zu diesem gelegen, so erreicht der Schall das eine Mikrofon früher als das andere, allerdings mit gleicher Intensität, da er auf beide nahezu im gleichen Winkel trifft (Görne, 2015). Bezeichnet wird das auf Laufzeitunterschiede gestützte Prinzip auch als AB-Aufbau, wobei durch den verwendeten Abstand beider Kapseln zueinander der Aufnahmewinkel beeinflusst werden kann. Unterschieden wird zwischen Klein-AB und Groß-AB, wobei Ersteres einen ungefähren Ohrenabstand und Letzteres einen größeren Abstand von bis zu mehreren Metern repräsentiert (Görne, 2015). Vergrößert sich der Kapselabstand, führt das zu einem kleineren Aufnahmewinkel und Schallquellen werden näher herangeholt (Görne, 2015). Dank ihrer klanglichen Vorzüge verwendet man häufig Kondensator-Druckempfänger mit der Richtcharakteristik einer Kugel für den Aufbau der Laufzeitstereofonie (Görne, 2015).

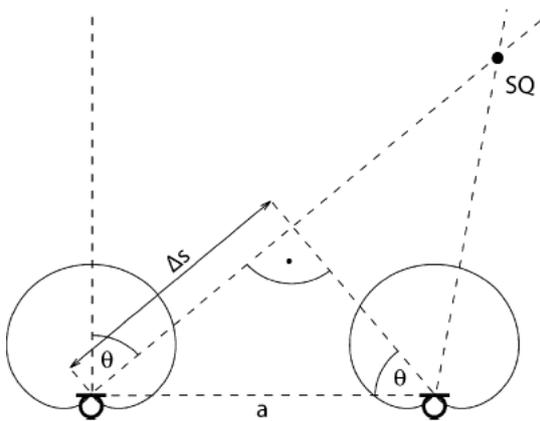


Abb. 5. Darstellung einer laufzeitstereofonischen Mikrofonanordnung zweier Nieren

4.1.2 Intensitätsstereofonie

Sind zwei Gradientenempfänger direkt übereinander gelegen, bilden allerdings durch Anwinkelung oder Wahl unterschiedlicher Kapseln unterschiedliche Richtungen ab, so spricht man von einer Koinzidenzmikrofonierung (Görne, 2015). Der Schall erreicht beide Mikrofone dabei zeitgleich, aber abhängig von der Einfallsrichtung mit unterschiedlicher Intensität (Pegel). Dieses Prinzip der Stereofonie kann dabei mit zwei unterschiedlichen

Mikrofonanordnungen durch jeweils anders gewählte Richtcharakteristiken entstehen, mit der XY- oder der MS-Anordnung.

Die XY-Anordnung in der Koinzidenzmikrofonierung beschreibt das Prinzip zweier gerichteter Mikrofonkapseln gleicher Richtwirkung (nahezu) am selben Ort, welche um einen Hauptachsenwinkel gegeneinander angewinkelt werden (Weinzierl, 2008). Der Aufnahmewinkel ist dabei über den Versatzwinkel beider Kapseln einstellbar, wobei ein größerer Versatzwinkel zu einem kleineren Aufnahmewinkel und somit zu einer größeren Abbildung führt (Görne, 2015). Der optimale Versatzwinkel ist dabei abhängig von der verwendeten Richtcharakteristik. So gilt in der XY-Anordnung für zwei Nieren ein optimaler Winkel von $\pm 66^\circ$, für Supernieren $\pm 58^\circ$, für Hypernieren $\pm 52^\circ$ und für zwei Achten $\pm 45^\circ$ (Görne, 2015). Letztere Anordnung zweier rechtwinklig angeordneten und auswärts gedrehten Achten bezeichnet eine besondere Art der XY-Anordnung, welche nach ihrem Erfinder als Blumlein-Anordnung bekannt ist. Dabei ergeben sich zwei gleichberechtigte Aufnahmebereiche vor und hinter den Mikrofonen, welche sich dabei durch den symmetrischen Aufbau ergeben (Görne, 2015).

Als weitere Stereoanordnung für die Aufnahme intensitätsstereofoner Signale gilt der MS-Aufbau, welcher durch die Anwendung zweier übereinander angeordneter, aber unterschiedlich gerichteter Kapseln ein Mitten- und ein Seitensignal widerspiegelt. Diese Codierung in zwei unterschiedliche Kanäle, welche nicht direkt dem linken oder rechten Kanal eines Lautsprecherpaars zuzuordnen ist, wird durch die Verwendung einer nach vorne gerichteten Kapsel beliebiger Richtcharakteristik für das Mittensignal und einer um 90° nach links gedrehten Acht erzielt (Görne, 2015). Eine Decodierung des Mitten- und Seitensignals für die stereofone Lautsprecherwiedergabe wird durch eine Summen- bzw. Differenzbildung dieser erzielt. Genanntes Prinzip der Codierung in raumbezogene Signale und der Rückwandlung für beliebige Lautsprecherkonfigurationen bezeichnet einen Workflow mit Ambisonics-Signalen (näher bezeichnet in Kapitel 5.3.2), eine szenenbasierte Darstellung des Schallfeldes.

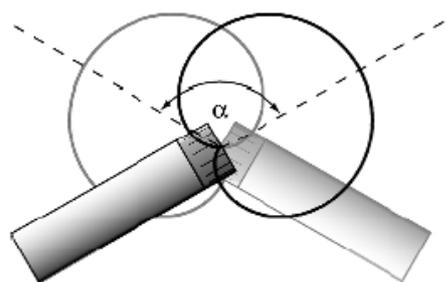


Abb. 6. XY-Stereofonieranordnung

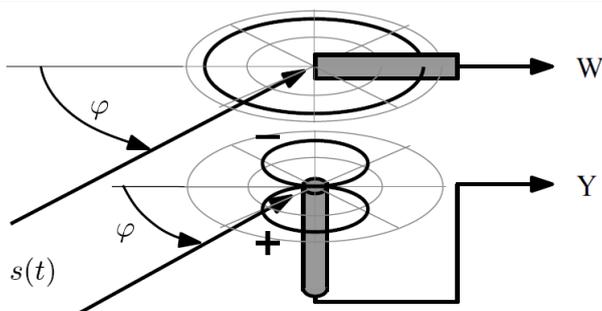


Abb. 7. Darstellung einer nativen Aufnahme in MS-Stereofonie

4.1.3 Äquivalenzstereofonie

Zwei unterschiedliche Mikrofonverfahren der Äquivalenzstereofonie bezeichnen die stereofone Aufnahme zweier Kanäle, welche räumlich getrennt und gegeneinander angewinkelt oder durch einen akustischen Trennkörper voneinander separiert sind und somit sowohl Laufzeit- als auch Pegeldifferenzen aufweisen (Görne, 2015). Die räumlich getrennte Anordnung führe dabei zu Laufzeit-, das gegeneinander Anwinkeln, ähnlich der Anwendung akustischer Trennkörper, zu Pegelunterschieden, so Görne (2015). Zwei populäre Anordnungen sind dabei ORTF und NOS, welche prinzipiell XY-Anordnungen mit räumlicher Trennung bezeichnen.

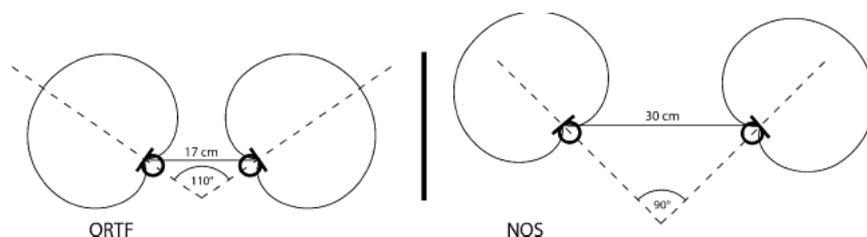


Abb. 8. Darstellung der äquivalenzstereofonen Anordnungen ORTF und NOS

4.2 Einzelmikrofonverfahren

Eine Stereomischung aus Aufnahmen im Einzelmikrofonverfahren setzt sich, im Gegensatz zu Aufnahmen der Hauptstereofonie, nicht schon im Aufnahmeraum an der Position der angeordneten Mikrofone zusammen, sondern bildet erst in der Tonregieeinrichtung durch verschiedene Pegel- und Abbildungsrichtungseinstellungen ein stereofones Klangbild (Dickreiter et al., 2008). Die Einzelmikrofonierung zielt dabei weniger auf eine korrekte klangliche akustische Abbildung des AufnahmeRaums ab. Vielmehr stehen andere

Verbesserungen bezogen auf die aufgenommenen Schallquellen im Vordergrund, wodurch die Mischung bei bestimmten Musikrichtungen durchaus von einer verbesserten Klangbalance, besserer Präsenz, Lokalisierbarkeit und Tiefenstaffelung profitiert (Dickreiter, 2011), vor allem wenn die Schallquellen sich in der Lautstärke stark unterscheiden oder diese noch im Nachhinein durch die Tonregie durch Klangbeeinflussung, Verhallung oder Verzögerung manipuliert werden sollen (Dickreiter et al., 2008).

Im Einzelmikrofonverfahren wird zu jeder Schallquelle unter Berücksichtigung einer möglichst stark gerichteten Charakteristik der Mikrofone und des jeweiligen Hallabstandes der Raumakustik ein möglichst geringer Mikrofonabstand vorausgesetzt, um den Direktschall- gegenüber dem Diffussschallanteil maximal zu halten. Alternativ können die Schallquellen auch durch distanzierte Anordnung oder den Einsatz von Trennwänden oder Kojen und den zusätzlichen Einsatz von Filtern im Frequenzbereich akustisch differenziert aufgenommen werden (Dickreiter et al., 2008). Daraus folgt das Ziel einer weitgehenden akustischen Trennung der Mikrofone (Dickreiter et al., 2008). Das Übersprechen zweier Einzelmikrofone muss mindestens um 12 bis 15dB gedämpft werden, um eine ausreichende akustische Entkopplung dieser zueinander zu erreichen (Dickreiter et al., 2008).

4.2.1 Anwendungsbereiche

Die Anwendung der Einzelmikrofonierung unterscheidet sich grundlegend zwischen der ausschließlichen Verwendung einzelner Mikrofone für einzelne Schallquellen und der Kombination mit verschiedenen Hauptmikrofonverfahren (näher beschrieben in Abs. 4.2.2). Prinzipiell ist eine Einzelmikrofonierung unter anderem bei Aufnahmen jeglicher Musikrichtungen möglich und findet seine Anwendung vor allem in Aufnahmestudios. Durch eine räumliche Trennung verschiedener Instrumentengruppen lässt sich das Übersprechen noch weiter reduzieren, sodass in der stereofonen Mischung eine möglichst große Flexibilität erreicht wird. Auch die Anwendung sogenannter Ansteckmikrofone verringert das Übersprechen und sorgt für eine klare Präsenz des aufgenommenen Signals. Diese werden zur Aufnahme häufig am Steg von Saiteninstrumenten, in Gussrahmen von Flügeln oder bei Filmtonaufnahmen für körpernahe Sprachaufnahmen eingesetzt. Die vielen Möglichkeiten der beeinflussbaren Klangbalance, Präsenz, Lokalisation und Tiefenstaffelung, sowie der

einzelnen Bearbeitung der Schallquellensignale im Frequenzbereich und durch Verzögerung im Nachhinein, finden vorzugsweise in der Pop- und Rock-Musik eine große Bedeutung, sodass sich in diesen Musikrichtungen die Mischung häufig erst in der Postproduktion ergibt. Traditionsbedingt steht bei Aufnahmen älterer Musikrichtungen aus der sogenannten E-Musik häufig die Raumakustik mehr im Vordergrund, sodass diese zumeist außerhalb eines Tonstudios in einer Kirche oder einem Konzertsaal stattfinden und zur Repräsentation der Raumakustik vermehrt auf das Verfahren der Hauptmikrofonierung gesetzt wird. Doch auch viele Aufnahmen akustischer Instrumente werden im Studio unter Berücksichtigung der Einzelmikrofonierung oder einer Kombination mit der Hauptmikrofonierung gemacht, was in der Mischung - dank Hallraumsimulationen auf Basis von Raumimpulsantworten aus Kirchen oder Konzertsälen - den eigentlichen Aufnahmen nicht wesentlich nachsteht und eine deutlich größere Flexibilität bietet. Auch in Sprachaufnahmen für Hörspiele oder in der Filmtonaufnahme hat die Einzelmikrofonierung und somit auch die Trennung der einzelnen Kanäle in der Wiedergabe eine große Bedeutung.

4.2.2 Kombination mit Hauptmikrofonierung

Werden Hauptmikrofon- und Einzelmikrofonverfahren kombiniert, spricht man von einer Hauptmikrofonierung mit zusätzlichen Stützmikrofonen. Die Hauptmikrofonierung nimmt dabei das akustische Klangfeld möglichst korrekt auf, wozu einzelne Mikrofone z.B. für einzelne Instrumentengruppen oder den Chor hinzugemischt werden können. So lässt sich auch bei einer Hauptmikrofonierung on Location (z.B. in einer Kirche) im Nachhinein die Mischung bezüglich der Klangbalance und der Präsenz einzelner Instrumentengruppen zueinander beeinflussen, ohne dabei die Akustik der Aufnahmeumgebung zu verlieren. Wichtig zu beachten sind dabei allerdings die verschiedenen Mikrofonabstände bezogen zur Hauptmikrofonierung, welche in der Nachbearbeitung durch eine Verzögerung der einzelnen Signale Beachtung finden muss, um etwaige Kammfiltereffekte zu verhindern. Ähnliches gilt bei der Verwendung von Stereostützmikrofonierungen, welche für die schlechte Ausblendung von Instrumentengruppen in zweiter Reihe bekannt sind, und so ein größeres Übersprechen aufweisen.

5 Immersive Lautsprecherwiedergabe

Die matrizierte und diskrete Audiowiedergabe im Surround-Format findet seine Ursprünge in der frühen Kinofilmvertonung, für welche stets nach Möglichkeiten gesucht wurde, die räumliche Tonwiedergabe zu erweitern. So wurde auch im Jahre 1940 in den Disneystudios für den legendären Animationsfilm „Fantasia“ mit einem aufwändigen achtkanaligen Verfahren experimentiert (Görne, 2015). Doch erst 1977 setzte sich das Surround-Format in Form von Dolby Stereo (auch bekannt als Dolby Surround) durch und wurde mit dem Film „Star Wars“ durch Dolby und Lucasfilm eingeführt (Görne, 2015). Eine matrizierte Surroundwiedergabe wie 4:2-Matrix-Surround (z.B. in Form von Dolby Stereo oder DTS-Stereo) bildet eine MS-Matrix-ähnliche Schaltung aus vier Kanälen, bestehend aus L(Left), C(Center), R(Right) und S(Surround), wobei die linken und rechten Surroundkanäle in einem zusammengefasst werden (Görne, 2015). Diese Matrizierung bietet die Vorteile einer geringen Datenmenge und der Stereokompatibilität, verfügt allerdings über eine schlechte Kanaltrennung von nur 3dB (Görne, 2015).

5.1 Konventionelle Lautsprecherlayouts nach ITU

Die diskrete Surroundwiedergabe über getrennt übertragene Kanäle stellt eine konsequente Weiterentwicklung der Matrixschaltungen dar und findet sich nach dem definierten Standard der ITU-Norm ITU-R BS.775-1 aus dem Jahre 1994 in Surroundformaten von Dolby, DTS und anderen Codec-Herstellern wieder (Görne, 2015). Genannte ITU-Norm beschreibt dabei den populären und mittlerweile klassischen Standard des 5.1-Surround mit 3/2 Kanälen, aufgeteilt in drei Frontkanäle, bestehend aus einem linken, einem Center und einem rechten Lautsprecher, und zwei links- und rechts-hinten angeordneten Surroundlautsprechern (Ls und Rs), optional ergänzt durch einen separaten Kanal für tieffrequente Effekte (LFE). In nachfolgenden ITU-Normen, wie der aktuellen ITU-R BS.2051-2 aus dem Jahre 2018, wird die Kanalanzahl auf horizontaler Ebene (Hörebene) zur verbesserten Darstellung seitlicher oder rückseitiger Schallquellen durch weitere Lautsprecher ergänzt. Auch kommen die Ebenen der Elevationsdarstellung hinzu, welche im Hinblick auf weitere Wiedergabekanäle zu einer anderen Sichtweise auf die Anordnung führen: Alle Wiedergabekanäle werden bezüglich ihrer vertikalen

Anordnung in Bottom-, Middle- und Upper-Layer aufgeteilt, sodass vor allem das Middle-Layer oftmals vergleichbar mit der typischen 3/2-Anordnung eines 5.1-Surroundsetups bleibt. Nach dem Konzept des Vector Base Amplitude Panning werden alle räumlich angeordneten Lautsprecherkanäle zur besseren Darstellung von Phantomschallquellen durch sogenanntes Triangulieren in Dreiecke aus jeweils drei Kanälen aufgeteilt. Demnach können sich Phantomschallquellen entweder auf einem Lautsprecher, auf der Achse zwischen zwei Lautsprechern oder innerhalb eines Lautsprecherdreiecks befinden, wonach jeweils entweder ein, zwei oder drei Lautsprecher gleichzeitig spielen. Doch auch bei allen diskreten immersiven Wiedergabeformaten bleibt die Stereokompatibilität stets essentiell, sodass sich auch in der ITU-Norm ein definierter Standard für einen Downmix wiederfindet - unter Berücksichtigung eines bestimmten Mischungsverhältnis.

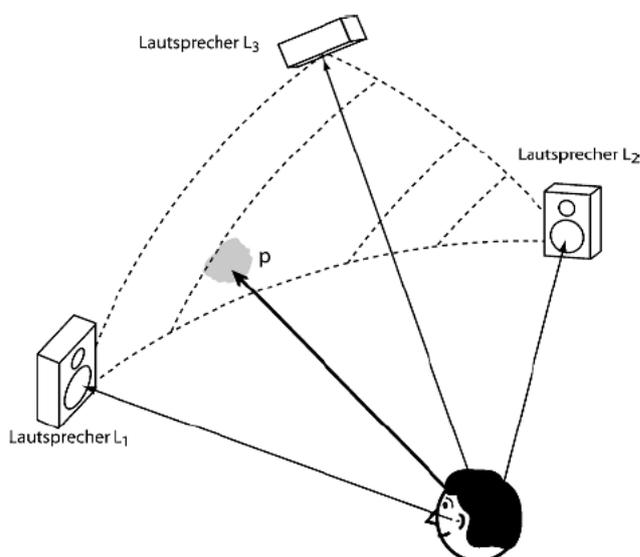


Abb. 9. Darstellung des Vector Base Amplitude Panning am Beispiel eines Lautsprecherdreiecks

5.1.1 Surroundwiedergabe mit 3/2 Kanälen (0+5+0)

Die Anordnung in 5.1-Surround ergänzt die stereofone Wiedergabe um einen Centerlautsprecher zwischen dem linken und rechten Kanal und um zwei seitliche Surroundlautsprecher L_s und R_s , sodass drei Frontlautsprecher zwei seitlichen Kanälen entgegenstehen (3/2). Insgesamt befinden sich also fünf Lautsprecher auf der horizontalen Hörebene, welche auch als Middle Layer bezeichnet wird (0+5+0). Bei der Konfiguration der Frontlautsprecher ist die Anordnung in einem gleichseitigen Dreieck mit der Hörerposition entscheidend (Görne, 2015). Die Frontkanäle L, C und R liegen mit den

beiden Surroundkanälen Ls und Rs - bei korrekter Anordnung - auf einer Kreisform, worin Ls und Rs in einem Toleranzbereich von $\pm 100^\circ$ bis $\pm 120^\circ$ aufgestellt werden (Görne, 2015). Lässt sich die Kreisform aufgrund räumlicher Gegebenheiten des Abhörortes nicht realisieren, so werden Lautsprechersignale entsprechend verzögert (Görne, 2015). Ist beispielsweise der Centerlautsprecher einen Meter zu nah angeordnet, muss das Signal um 2,94ms verzögert werden (Görne, 2015). Speziell für die Filmbeschallung wurde in die Standards ITU-R BS.775-1 und SMPTE RP-173 ein spezieller Übertragungskanal für den Bassbereich von 20Hz bis etwa 80/120Hz übernommen, der LFE, welcher der Mischung häufig einen bedrohlichen Charakter ohne lokalisierbaren Ort verleiht, aber nicht mit einem Subwoofer zur Wiedergabe tiefer Frequenzen verwechselt werden sollte (Görne, 2015).

Das Prinzip der Zweikanal-Stereofonie des klaren Referenzpunktes, oder auch „sweet spot“, behält auch bei der Mehrkanal-Stereofonie seine Gültigkeit, wenn auch durch den zusätzlichen Centerlautsprecher Vorteile einer Verbreiterung der Stereohörfläche und einer höheren Lokalisationsschärfe entstehen und somit die Verschiebungen der Phantomschallquellen von vorne deutlich geringer sind (Görne, 2015). Allerdings entstehen auch neue Probleme, vor allem in der Darstellung seitlicher Phantomschallquellen, sodass diese sehr instabil und empfindlich in Bezug zum Abhörpunkt sind (Görne, 2015). Bewegt sich der Hörer nur leicht aus dem Referenzpunkt heraus oder dreht seinen Kopf, wird die Lokalisationsschärfe deutlich geringer und seitliche Phantomschallquellen werden verschwommen und instabil, was zu gravierenden Richtungsverzerrungen führen kann (Görne, 2015). Auf das Panning von Hörereignissen zwischen den Front- und Surroundlautsprechern mithilfe dem Einsatz von Phantomschallquellen sollte daher grundsätzlich verzichtet werden (Görne, 2015). Auch gibt es bei der Surroundwiedergabe mit fünf Lautsprechern die Gefahr eines „Lochs“ zwischen frontseitiger und seitlicher Abbildung, welches es bei großen Abhörräumen durch den Einsatz weiterer Surroundlautsprecher zu schließen gilt (Görne, 2015).

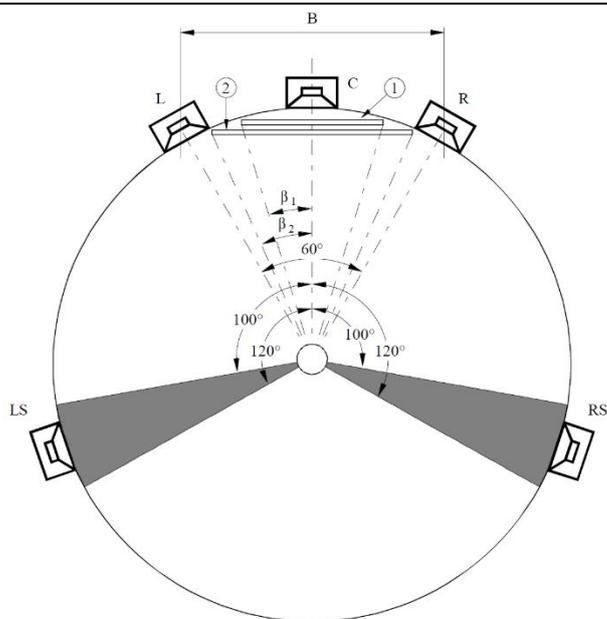


Abb. 10. Darstellung einer 3/2-Lautsprecheranordnung nach ITU-R BS.775-1

5.1.2 Surroundwiedergabe mit 5/2 oder 3/4 Kanälen (0+7+0)

Die Surroundwiedergabe mit sieben horizontal angeordneten Lautsprechern ist auf zwei verschiedene Arten möglich, mit fünf Front- zu zwei Surroundkanälen oder mit drei Front- zu vier Surroundkanälen. Erstes Verfahren ergänzt die Anordnung um zwei weitere seitliche Lautsprecher, welche rechts und links weiter hinten angeordnet sind (Lss und Rss). Diese Anordnung hält an der dreikanaligen Frontanordnung nach 5.1 fest, sorgt aber mit ihrer erweiterten seitlichen Wiedergabe für eine bessere Abbildung seitlicher Phantomschallquellen. Das zweite diskrete Surroundverfahren des 7.1 verfügt, wie auch das 5.1-Surround, über zwei seitlich angeordnete Surroundkanäle, wobei die Front durch zusätzliche Centerlautsprecher davon abweicht. Die frontseitige Wiedergabe in dem von Sony etablierten Format besteht somit aus insgesamt fünf diskreten Kanälen: Die stereofone Anordnung des linken und rechten Kanals (L und R), sowie der mittig angeordnete Centerkanal (C) bleiben bestehen und werden durch zwei zusätzliche Lautsprecher zwischen den stereofonen Kanälen und dem Centerkanal ergänzt. Sie bilden damit einen linken und einen rechten Zusatz zum Centerkanal (LC und RC), was vor allem bei der Wiedergabe frontseitiger Hörereignisse in besonders großen Abhörräumen die Lokalisationsschärfe gewährleisten soll. Somit ist die Surroundwiedergabe in 7.1

Elevation entstehende Neigung zum Abhörpunkt, welche in einem Toleranzbereich von 30° bis 55° liegt.

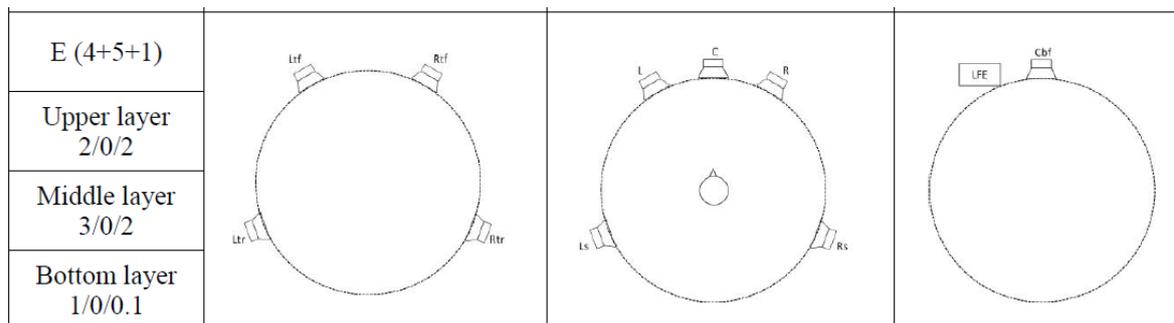


Abb. 12. Darstellung der Anordnung nach Sound System D-Format (9.1) (links nach rechts = oben nach unten)

5.1.4 Wiedergabe in 22.2 (9+10+3)

Zur Erweiterung der immersiven Wiedergabemöglichkeiten werden im Format des 22.2 sowohl Bottom-, Middle- und Upper Layer durch weitere diskrete Kanäle auf neun zu zehn zu drei ergänzt. Das Middle Layer besteht aus der typischen Anordnung aus linkem, rechten und Centerkanal (LCR) und den beiden Surroundkanälen back left und back right (BL und BR), front- und rückseitig unterstützt durch weitere Centerkanäle front left center, front right center (FLc und FRc) und back center (BC). Auch die obere Lautsprecherebene verfügt über eine ähnliche Anordnung, sodass die typische Surroundanordnung mit linkem (TpFL), mittigen (TpFC) und rechten Kanal (TpFR) mit zwei zusätzlichen Surroundkanälen (TpBL und TpBR) auch oberhalb der Hörachse, unter Berücksichtigung des Elevationswinkels, ihre Bedeutung behält. Zusätzlich werden in der oberen Lautsprecherebene links und rechts im 90°-Winkel (TpSiL und TpSiR), hinten mittig (TpBC) und vertikal direkt über der Hörerposition (TpC) weitere Kanäle angeordnet. Für eine unterseitige Beschallung sorgen drei im Lower Layer angeordnete Frontlautsprecher, welche eine tiefergelegte Abbildung eines linken, mittigen und rechten Kanals darstellen (BtFL, BtFC und BtFR). Auf der gleichen Ebene befinden sich auch die beiden Kanäle für tieffrequente Effekte (LFE1 und LFE2). Wie auch bei zweidimensionalen Surroundformaten ist auch bei vertikalen Anordnungen jeweils auf einen gleichen Lautsprecherabstand zum Referenzpunkt der Hörerposition zu achten, sodass sich die Wiedergabekanäle bei optimaler Anordnung auf einer Kugelform wiederfinden. Dies führt auch zu einer

jeweiligen Neigung hin zum Referenzpunkt bezogen auf die Elevation und zu einer jeweiligen Signalverzögerung bei Nichteinhalten der Abstände.

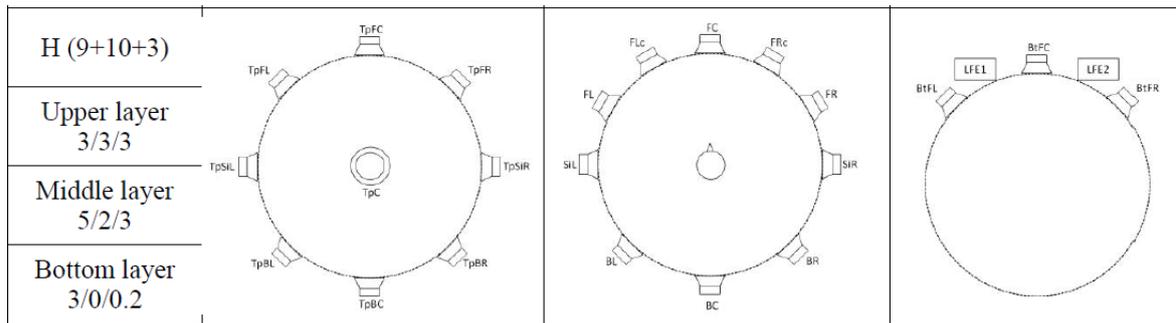


Abb. 13. Darstellung der Anordnung nach Sound System H-Format (22.2) (links nach rechts = oben nach unten)

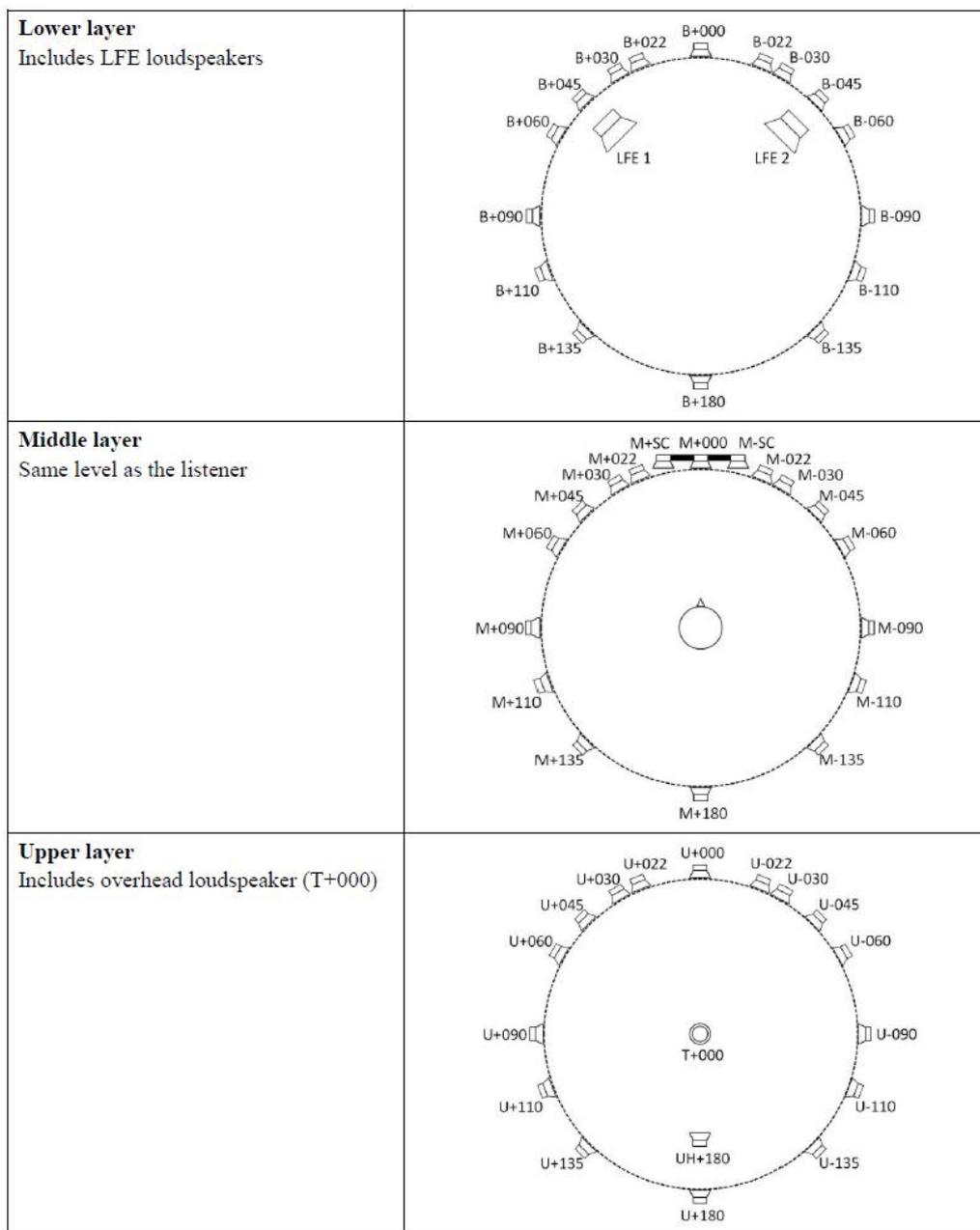


Abb. 14. Alle möglichen Lautsprecherkanäle des Middle, Upper und Lower Layer nach ITU-R BS.2051-2

5.2 Alternative Lautsprecherlayouts

Auch gibt es neben der diskreten Lautsprecherwiedergabe, welche auf einen Abhörpunkt hin definiert ist, noch unterschiedliche alternative Möglichkeiten der immersiven Lautsprecherwiedergabe. So zielt beispielsweise die Wellenfeldsynthese, anlehnd an das Prinzip des sogenannten akustischen Vorhangs, auf eine realistische Nachbildung eines Schallfeldes ab (Dickreiter et al., 2008). Auf eine Nachbildung einer diskreten immersiven und Lautsprecherwiedergabe zielt der Einsatz einer entsprechend ausgerüsteten Soundbar ab, welche durch Reflexion an Wänden, Decke und Boden des Abhörortes die Schallquellenpositionen auf vertikaler horizontaler Ebene imitiert.

5.2.1 Wellenfeldsynthese

Einen akustischen Raum mit realistischen und ortstabilen virtuellen Schallquellen im gesamten Hörfeld zu erzeugen, gilt als Zielsetzung der Wellenfeldsynthese (Dickreiter et al., 2008). Das auf dem Huygensschen Prinzip basierende physikalische Modell wurde zuerst an der Technischen Universität Delft im Jahre 1988 entwickelt und dort als Demonstrationssystem aus 160 Lautsprechern das erste Mal realisiert (Dickreiter et al., 2008). Erzeugt eine Quelle demnach eine Wellenfront, auch in Form sich ausbreitenden Schalls, so stellt diese die Summe unendlich vieler Einzelquellen dar, sodass im Umkehrschluss eine mit der ursprünglichen Welle identische neue Wellenfront synthetisiert werden kann, indem alle Elementarwellen sich überlagern (Dickreiter et al., 2008). In der Theorie kann dadurch jede Schallquelle durch den Einsatz unendlich vieler Lautsprecher nachgebildet werden (Dickreiter et al., 2008). Das Prinzip des akustischen Vorhangs zeigt genau diese Wellenfrontnachbildung im Bild einer dicht nebeneinander durchlöcherten Wand zwischen Schallquelle und Publikumsraum, hinter welcher Schallereignisse perfekt zu hören sind. Werden die Löcher durch linear verbundene Mikrofone und Lautsprecher ersetzt, so führt dies zu gleichen Hörereignissen. Im Beschallungskonzept der Wellenfeldsynthese findet zumeist allerdings keine lineare Verbindung von Lautsprechern mit entsprechenden Mikrofonen im Sinne des akustischen Vorhangs statt. Die Lautsprecher werden vielmehr als horizontales Line Array an den Wänden rund um den Hörort angeordnet (Görne, 2015), in neueren Konzepten auch horizontal in der Elevation. Für eine optimale Wiedergabe von beispielsweise musikalischem Inhalt sollte bereits in der Aufnahme auf eine Mikrofonierung der

trockenen Signale der Schallquellen ohne entsprechende Raumanteile geachtet und anschließend die akustischen Eigenschaften des Aufnahmeraums in Form einer Raumimpulsantwort erfasst werden (Dickreiter et al., 2008). So werden im Sinne einer objektorientierten Übertragung die trockenen Audiosignale mit der Raumimpulsantwort des Aufnahmeraumes oder eines anderen Ursprungs gefaltet und einzelne Lautsprechersignale entsprechend berechnet (Dickreiter et al., 2008). Die Vorteile einer immersiven Beschallung im Sinne der Wellenfeldsynthese liegen vor allem in dem theoretisch fehlerfrei synthetisierten Schallfeld und die dadurch korrekte Richtungsartung im ganzen Hörraum, welche somit nicht mehr auf einen Sweet Spot beschränkt bleibt (Görne, 2015). Auch können Phantomschallquellen in Form virtueller Punktschallquellen, im Gegensatz zum Virtual Base Amplitude Panning diskreter Systeme, frei im Abhörraum platziert und vom Hörer präzise lokalisiert und eingeschränkt umwandert werden (Dickreiter et al., 2008). Auch synthetisierte ebene Schallwellen mit unendlich großer Entfernung zur virtuellen Quelle wandern in ihrer Richtungswahrnehmung bei Bewegung des Hörers mit (Dickreiter et al., 2008). So erreicht die Wellenfeldsynthese in vielen Fällen ihr Ziel eines realistisch wahrgenommenen Raumeindrucks durch Einhüllung des Hörers (Dickreiter et al., 2008). Doch auch die Wellenfeldsynthese kommt in der Praxis nicht ohne Einschränkungen und nachteilige Effekte aus, welche sich häufig durch den Aufbau ergeben: So führt der Abstand der Array-Lautsprecher zueinander zu einer Grenzfrequenz des Spatial Aliasing, über welcher durch eine physikalisch inkorrekte Synthetisierung des Schallfeldes räumliche und spektrale Fehler auftreten (Dickreiter et al., 2008). Auch führen an den Rändern begrenzter Arrays entstehende Beugungswellen zu bemerkbaren Klangverfärbungen fokussierter Schallquellen (Dickreiter et al., 2008). Die Übertragung ausgedehnter Schallquellen ist mit dem Ansatz der Faltung trockener Signale mit ihren Raumimpulsantworten kaum praktikabel, da beispielsweise ein Chor oder ein Klavier beinahe unmöglich ohne Raumelemente aufgenommen werden können (Dickreiter et al., 2008). Wegen des großen Aufwand im Aufbau eines solchen Systems und wegen einer wachsenden erforderlichen Übertragungskapazität wurden bis heute nur wenige Abhörräume nach dem Konzept der Wellenfeldsynthese ausgestattet, finden allerdings vereinzelt ihre sinnvolle Anwendung in Kinos oder Konzertbeschallungen.

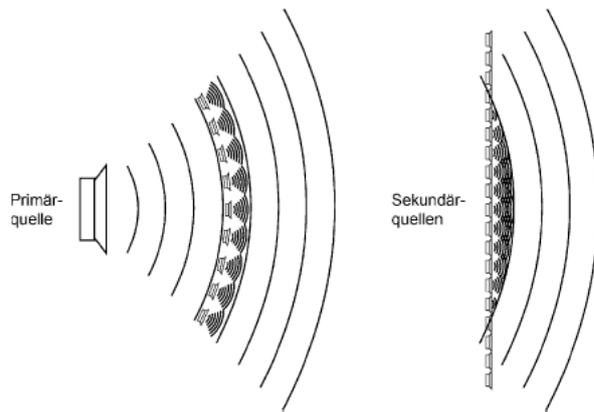


Abb. 15. Abbildung des Huygens-Fresnel-Prinzip und der Wellenfelsynthese

5.2.2 Reflexions-/ Trausaualbasierte Soundbar

Eine gute Alternative zur diskreten immersiven Wiedergabe - gerade im Homebereich - stellen mittlerweile Soundbars da, welche auf zwei verschiedenen Prinzipien beruhen können. Die Soundbars verfügen dabei über unterschiedlich angeordnete Lautsprecherarrays, welche nebeneinander nach vorne, seitlich, nach oben oder nach unten abstrahlen. Nach diesen Abstrahlverhalten führen zumeist erste oder zweite Reflexionen an Wänden, Decke und Boden zu einer Art Simulation räumlich angeordneter Schallquellen. Eine korrekte Darstellung der durch Reflexionen erworbenen Hörereignisse bedarf dabei ganz anderer akustischer Voraussetzungen an einen Abhörraum als bei einer diskret-räumlichen Beschallung, sodass im Gegensatz zu einer optimalen Tonstudioakustik eine gewisse Schallhärte der Wände und Decken in einer symmetrischen Anordnung zum Abhörpunkt gewährleistet sein muss. Durch eine raumakustische Einmessung an der Hörerposition können mithilfe kanalspezifischer Pegel- oder Laufzeitkorrekturen gewisse raumsymmetrische Unterschiede in Grenzen ausgeglichen werden. Eine weitere mögliche Bauart einer immersiv-beschallenden Soundbar stellt die Verwendung einer auf den Hörer angepassten transauralen Wiedergabe dar. Durch Beamforming erreichen beide Ohren binaurale Signale, welche durch zusätzliche Crosstalk-Cancellation-Filter weiter voneinander unterschieden werden. Wird die transaurale Wiedergabe über eine Soundbar durch Positionstracking des Hörers ergänzt, ist eine korrekte binaurale Wahrnehmung auch bei Bewegung des Hörers im Raum gewährleistet, solange dieser sich nicht von der Soundbar abwendet.

5.3 Wiedergabe unterschiedlichen Inhalts

Für die Wiedergabe auf diskreter immersiver Lautsprecherwiedergabe vorgesehener Inhalt muss nicht zwangsläufig bereits auf eine diskrete Kanalordnung ausgelegt sein, sondern kann auch in anderer Art „gespeichert“ werden. So ist auch eine Wiedergabe objektbasierten Inhalts oder von Ambisonics-Signalen auf diskreten Systemen möglich, wobei es jeweils einer weiteren Verarbeitung oder Decodierung bedarf.

5.3.1 Wiedergabe objektbasierten Inhalts

Eine Wiedergabe objektbasierten Inhalts ist prinzipiell - unter Berücksichtigung unterschiedlich erforderlichen Processings - auf allen Anordnungen immersiver Wiedergabe möglich. Objektbasiert bedeutet in diesem Kontext eine Aufteilung der Mischung in einzelne Audio-Elemente, welchen einzeln beispielsweise richtungs- oder lautstärkenbestimmende Metadaten zugeordnet werden. In der Wiedergabe über verschiedene Lautsprecheranordnungen oder Kopfhörer berücksichtigt ein Renderer die in den Metadaten der Audio-Objekte hinterlegten Richtungs- und Lautstärkeinformationen und wendet, nach dem jeweiligen Wiedergabelayout, unterschiedliche Signalverarbeitungsprozesse an. So werden für eine binaurale Wiedergabe zwei Ohrsignale berechnet, wohingegen in einer diskreten immersiven Wiedergabe die Signale der einzelnen Lautsprecher errechnet werden und dabei auch unkonventionelle räumliche Anordnungen berücksichtigt werden können. Auch eine gemischte Wiedergabe objektbasierten und kanalbasierten Contents ist in vielen Fällen möglich und findet seine Anwendung auch in kommerziellen Systemen, beispielsweise in Dolby Atmos, wo eine Kinomischung oftmals in ein kanalbasiertes Bed aus Dialogen, Musik und Atmos und objektbasierte Sound Effekte aufgeteilt wird.

5.3.2 Lautsprecherwiedergabe und Decodierung von Ambisonics 1. Ordnung

Sprechen wir im Bezug immersiver Wiedergabe oder auch Mikrofonierung über Ambisonics, so ist dabei eine szenenbasierte Repräsentation des Schallfeldes gemeint, welche in der Mikrofonierung durch native Anordnungen oder eine Encodierung entsteht und zur Wiedergabe auf Lautsprechern einer kanalabhängigen Decodierung bedarf. Um im Prinzip der szenenbasierten Darstellung in einem vereinfachten Rahmen zu bleiben wird im folgenden hauptsächlich auf die Bestandteile und Decodierung einer First-Order

Ambisonics-Darstellung eingegangen, wohingegen die Decodierung aus Higher-Order Ambisonics im Prinzip ähnlicher aber in der Anwendung aufwändigerer Berechnungen bedarf. Eine vollständige Audioszene in FOA teilt sich, für die Darstellung verschiedener Schallrichtungen, in die Bestandteils-Signale W , X , Y und Z auf. In einer korrekten Ambisonics-Szene enthält W dabei die Richtungsinformation einer Kugelcharakteristik, X die einer Achtercharakteristik nach vorne und hinten gerichtet, Y die einer Achtercharakteristik nach links und rechts gerichtet und Z die einer Achtercharakteristik nach oben und unten gerichtet. First-Order Ambisonics bildet somit alle drei Dimensionen eines Schallfeldes ab, verglichen mit Higher-Order Ambisonics allerdings noch in einer recht geringen räumlichen Auflösung. Soll eine Ambisonics-Szene nun über ein einfaches symmetrisches Lautsprecherlayout wiedergegeben werden, errechnen sich im Zuge der Decodierung die einzelnen Lautsprechersignale jeweils aus einer Addition oder Subtraktion aus den Signalen von X , Y und Z mit dem W -Signal. Wächst die Ordnung der szenenbasierten Repräsentation oder die Kanalanzahl der diskreten Wiedergabe an, so wächst auch der Aufwand für die Berechnung der einzelnen Lautsprechersignale, welcher sich nun nicht mehr auf einfache Matrixoperationen beläuft, sondern eher einem Rendering nahe kommt. Neben der Lautsprecherwiedergabe decodierter Ambisonicssignale ist auch eine recht einfache Realisierung der immersiven Kopfhörerwiedergabe möglich, welche auf einer ähnlichen Berechnung der Ohrsignale, abhängig der winkelabhängigen kopfbezogenen Impulsantworten, beruht.

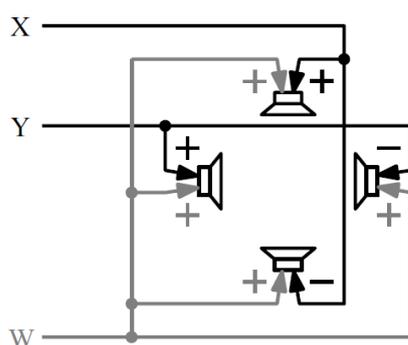


Abb. 16. Darstellung einer Decodierung von 2D-Ambisonicssignalen erster Ordnung auf symmetrische Lautsprecheranordnung

6 3D Mikrofonanordnungen und ihre praktischen Einsatzbereiche

Aus den Erkenntnissen einer stereophonen Mikrofonierung, ausgelegt für eine zweikanalige Repräsentation des aufgenommenen Schallfeldes in einer definierten optimalen Abhörposition, kam bereits früh der Wunsch nach äquivalenten Mikrofonanordnungen für mehrkanalige immersive Wiedergabesysteme auf. Neben einer auf Einzelmikrofonierung basierenden Mehrkanal-Mischung bestehen auch Möglichkeiten mehrkanaliger Mikrofonanordnungen, ganz nach Beispiel der stereophonen Hauptmikrofonierung (Görne, 2015). Auf eine Wiedergabe in 5.1-Surround ausgelegte Mikrofonverfahren verfolgen dabei in der Praxis eine systematische Einteilung in Haupt-, Front- oder Surroundmikrofonverfahren (Dickreiter, 2011). Diese Einteilung, auch in verschiedenen sinnvolle Anwendungsgebiete, behält auch bei weiterentwickelten Aufnahmeverfahren zur immersiven Wiedergabe auf 4+5+0 oder 9+10+3 Kanälen ihre Bedeutung. Auch bestehen bei Mikrofonverfahren der immersiven Aufnahme - vergleichbar zu verschiedenen Aufnahmeverfahren der Stereofonie - unterschiedliche Differenzen zwischen den Signalen der einzelnen Kanäle, welche aufgrund unterschiedlicher Anordnungen durch Pegel-, Laufzeit- oder Pegel- und Laufzeitdifferenzen entstehen (Dickreiter, 2011). Angesprochene verschiedene Aufnahmeverfahren beinhalten auch Mikrofonanordnungen, welche unabhängig eines bestimmten Wiedergabesystems zu betrachten sind und somit einer anderen Abbildung des aufgenommenen Schallfeldes entsprechen.

Zur Einordnung verschiedener 3D Mikrofonanordnungen in Bezug auf Wiedergabe, Aufbau, Eigenschaften, Wahrnehmung und Anwendungsbereiche ist in dieser Arbeit, nach einer Beschreibung der Taxonomie der Mikrofonanordnungen in genannte Kategorien, eine Übersichtstabelle der vorhandenen Möglichkeiten immersiver Aufnahmen enthalten, welche den Leser bestenfalls zu einer etwaigen Anwendungsentscheidung bestimmter möglicher Arrays befähigen soll. Sieht der Leser für seine Anwendung nach der Übersichtstabelle eine oder mehrere Mikrofonanordnungen als potentiell geeignet an, so findet er in den folgenden Kapiteln weitere Beschreibungen und Illustrationen der einzelnen möglichen Aufnahmeverfahren.

6.1 Taxonomie mehrkanaliger 3D Mikrofonanordnungen

Folgende Taxonomie mehrkanaliger 3D Mikrofonanordnungen beschäftigt sich mit der Einordnung, Kategorisierung und Bewertung unterschiedlich möglicher Mikrofonierungsverfahren zur immersiven Aufnahme eines Schallfeldes und berücksichtigt dabei eine Bewertung der Arrays bezüglich ihrer Anwendung und der Wiedergabe ihrer Aufnahmen. Die Taxonomie erfolgt dabei in Kategorien einer lautsprecherbezogenen oder lautsprecherunabhängigen Wiedergabe, des Aufbaus der Arrays, deren Eigenschaften, der Hörerwahrnehmung der Aufnahmen und der sinnvollen Anwendungsbereiche der Mikrofonanordnungen. In diesem Kapitel wird beschrieben, nach welchen Gesichtspunkten die Arrays kategorisiert werden, in welchem Zusammenhang Kategorien zueinander stehen und wie eine kategoriebezogene Bewertung erfolgt. Kommt es bei der Einordnung oder Bewertung bestimmter Mikrofonanordnungen oder -verfahren zu Unklarheiten oder Missverständlichkeiten, so werden diese im Kapitel 7 der Diskussion in dieser Arbeit nochmals angesprochen. Als Ergebnis der Auseinandersetzung mit einer Taxonomie mehrkanaliger 3D Mikrofonanordnungen enthält diese Arbeit anschließend eine Übersichtstabelle dieser, welche dem Leser eine erste Vorstellung über Aufwand und Anwendung der unterschiedlichen Arrays im Vergleich geben soll.

6.1.1 Lautsprecherbezogen oder Wiedergabeunabhängig

Wie bereits in Kapitel 5 bezüglich verschiedener diskreter immersiver Wiedergabeformate angesprochen und der in dessen Unterkapitel 5.3.2 beschriebenen Wiedergabe szenenbasierten Contents über diese, gibt es wiedergabeseitig einen prinzipiellen Unterschied von Aufnahmen aus Verfahren koinzidenten oder räumlich beanstandeten Aufbaus. So gelten verschiedene Aufnahmeanordnungen, abhängig ihres Aufbaus und ihrer Kanalanzahl oder -anordnung, als lautsprecherbezogen auf in der ITU-Norm definierte Wiedergabesysteme, oder als wiedergabeunabhängig im Sinne einer Decodierung einer Ambisonics-Szene auf eine beliebige Lautsprecheranordnung.

Die Anordnung der Aufnahmekanäle innerhalb eines lautsprecherbezogenen Mikrofonarrays ist dabei, genau wie bei den Konzepten der diskreten immersiven Wiedergabe, aufgeteilt in Upper-, Middle- und Lower Layer, wobei dies nur in manchen

Fällen durch die tatsächliche vertikale Mikrofonposition begründet ist, in anderen Fällen durch die Wahl einer entsprechenden Neigung und Richtcharakteristik. Einem bestimmten diskreten Wiedergabesystem zuordnen lassen sich lautsprecherbezogene Mikrofonanordnungen dabei unter Berücksichtigung der Kanalanzahl und -anordnung. Hat ein Mikrofonarray dabei 4+4+0 Kanäle, so handelt es sich um ein Surroundmikrofonverfahren mit ergänzter vierkanaliger Elevationsdarstellung, welches im Sound System D der ITU-Norm ITU-R BS 2051-2 den Kanälen L, R, Ls, Rs, Ltf, Rtf, Ltr und Rtr zugeordnet werden kann. Mikrofonarrays mit 4+5+0 Kanälen können dabei ebenfalls dem Sound System D der ITU-Norm zugeordnet werden, wobei diese im Middle Layer über einen zusätzlichen Aufnahmekanal zur diskreten Center-Darstellung verfügen. Ein Mikrofonarray mit 4+6+0 Kanälen ist dabei allerdings nur bedingt über das Sound System D der ITU-Norm wiederzugeben, da das Middle Layer über einen Aufnahmekanal mehr verfügt als das Wiedergabesystem beinhaltet, sodass zur korrekten Wiedergabe ein zusätzlicher Lautsprecherkanal oder ein Rendering notwendig wird. Auch gibt es Mikrofonarrays mit noch mehr Kanälen im Upper-, Middle und Lower Layer, aktuell bis zu 9+10+3 Kanäle, welche dem diskreten Wiedergabeformat des Sound System H der ITU-Norm und somit neun Lautsprechern des Upper Layer, zehn Lautsprechern des Middle Layer und drei frontal angeordneten Lautsprechern des Lower Layer zugeordnet werden.

Gilt es der Lautsprecherwiedergabe von Aufnahmen aus Mikrofonanordnungen szenenbasierten Prinzips, so bedarf es vor einer diskreten Kanaluordnung einer Decodierung auf das jeweilige Wiedergabeformat. Die Ambisonics-Signale bilden somit keine direkten Lautsprechersignale ab, sondern beschreiben das Schallfeld als eine Art Audio-Szene, wobei Bestandteilssignale zumeist in Form einer Achtercharakteristik mehrere Seiten gleichzeitig beschreiben. In einem nativen Mikrofonarray von Ambisonics erster Ordnung bildet ein Mikrofon der Kugelcharakteristik das W-Signal, eine Acht nach links und rechts ausgerichtet das Y-Signal, eine Acht nach vorne und hinten ausgerichtet das X-Signal und eine Acht nach oben und unten ausgerichtet das Z-Signal ab. Eine Mikrofonanordnung der Higher Order Ambisonics ergänzt die Szene um weitere Bestandteils-Signale und somit um eine höhere räumliche Auflösung. Aufnahmen in Ambisonics erster oder höherer Ordnung können prinzipiell, mit Einschränkung einer geeigneten Decodierung, flexibel auf allen immersiven Wiedergabemöglichkeiten gespielt

werden und sind somit sowohl für Binauralsynthese, als auch für eine Lautsprecherwiedergabe, beispielsweise über Sound System D oder H, geeignet.

6.1.2 Aufbau

Geht es um Unterschiede im Aufbau von 3D Mikrofonanordnungen, so ist häufig für den Anwender vorerst die physikalische Anordnung der Mikrofone zueinander und daraus folgend ein Vergleich der Portabilität der Arrays essentiell. Die Mikrofone können dabei in drei verschiedenen Konfigurationsarten zueinander stehen: horizontal und vertikal beabstandet (HVS = horizontally & vertically spaced), horizontal beabstandet und vertikal koinzident (HSVC = horizontally spaced & vertically coincident) oder horizontal und vertikal koinzident (HVC = horizontally & vertically coincident) (Lee, 2021). Physikalische Anordnungen der horizontal und vertikal beabstandeten Konfiguration zielen dabei eher auf eine wahrnehmungsbezogene räumliche Reproduktion des Schallfeldes ab, wohingegen Konfigurationen der horizontal und vertikal koinzidenten Anordnung auf eine physikalisch korrekte Reproduktion in Form einer Audioszene abzielen. Bei einer horizontal beabstandet und vertikal koinzidenten Anordnung der Mikrofone sind die Aufnahmekanäle des Upper-Layers auf der gleichen vertikalen Ebene wie die des Middle-Layers angeordnet, zumeist aber mit nach oben stärker gerichteten Charakteristiken. Die beabstandete oder koinzidente Anordnung in einem 3D Mikrofonarray hat einen großen Einfluss auf das Verhältnis der einzelnen Mikrofonsignale zueinander und deren Eigenschaften, und so auch auf die Wahrnehmungsunterschiede in der Wiedergabe über Lautsprecher, sowie dadurch indirekt auch auf die möglichen Bereiche ihrer optimalen Anwendung. Doch in der praktischen Anwendung spielt auch häufig die Portabilität der Arrays eine entscheidende Rolle, sodass in genannter Übersichtstabelle dieser Arbeit abhängig der Anordnung der Mikrofone zueinander - eine Bewertung in vier Stufen erfolgt. So gelten Arrays der horizontal und vertikal koinzidenten Anordnung prinzipiell als zumeist sehr portabel, wohingegen Arrays horizontal und vertikal beabstandeter Anordnung als weniger portabel gelten. Doch auch für horizontal beabstandete und vertikal koinzidente Anordnungen gibt es teilweise kommerzielle, recht portable Lösungen, so auch für ORTF-3D, welches so zum Beispiel auch in Windkörben für einen einfachen Außeneinsatz nutzbar wird. Bei der Kategorisierung verschiedener Konfigurationen geht es folgend auch um die vertikale Orientierung und Richtcharakteristik in der Elevation angeordneter Mikrofone, welche einen großen Einfluss

auf kanalübergreifende Eigenschaften, sowie auf die Wahrnehmungskomponenten der über Lautsprecher wiedergegebenen Aufnahmen haben. Unterschieden wird dabei zwischen nach oben oder von der Schallquelle weg gerichteter Richtmikrofone und nach vorne gerichteter Richtmikrofone oder Mikrofone der Richtcharakteristik einer Kugel, wobei alle Aufnahmekanäle des Upper Layer der gleichen Richtcharakteristik entsprechen. Der Einsatz nach oben oder rückseitig gerichteter Mikrofone von Charakteristik einer Niere oder Superniere führt dabei zu einem zunehmend geringen Direktschallanteil in den Aufnahmekanälen des Upper Layer, sodass im Mischungsverhältnis zwischen Middle und Upper Layer ein größerer Spielraum besteht, als bei größerem Direktschallanteil. Zu diesem führt der Einsatz von Mikrofonen mit Kugelcharakteristik oder nach vorne gerichteter Mikrofone mit Nierencharakteristik, welche durch ihre verbesserte vertikale und ihre bessere tieffrequente spektrale Abbildung in der Aufnahme besser für musikalische Schallquellen in der Elevation geeignet, allerdings gleichzeitig in der Wiedergabe anfälliger für eine vertikale Verschiebung horizontaler Schallquellen oder für Unterschiede in der Klangfärbung sind. Bei Aufnahmeanordnungen der Ambisonics-Domäne findet sich in der Vertikalen prinzipiell der Einsatz gerichteter Mikrofone wieder, welche hierbei allerdings keine Repräsentation des Upper Layer bilden und somit für den Vergleich des Übersprechens und der Korrelation zum Middle Layer keine Bedeutung haben.

6.1.3 Eigenschaften

Wie in der vorangegangenen Kategorisierung der Mikrofonanordnungen beschrieben, stehen die Eigenschaften eines Arrays und dessen Mikrofone in Verbindung zu jeweiligen Abstands- und Richtungsverhältnissen in deren Aufbau. Folgend lassen sich aus den Eigenschaften der Aufnahmekanäle Rückschlüsse auf die Wahrnehmungskomponenten eines Hörers bei der Wiedergabe über Lautsprecher ziehen. Das kanalübergreifende Übersprechen beschreibt dabei den Direktschallanteil auf Aufnahmekanälen, welche nach Konzeption ausschließlich oder hauptsächlich diffusen Schall aufnehmen sollen. Ist das Übersprechen auf vielen Kanälen vergleichsweise groß, so führt dies in der Wahrnehmung des Hörers durch interaurale Kammfiltereffekte zu stärkeren Verfärbungen des Klangbildes. Eine weitere Eigenschaft stellt die Korrelation zwischen allen Aufnahmekanälen dar, welche in der Aufnahme ebenfalls stark von Richtcharakteristik und Ausrichtung der Mikrofone abhängt. Sind die Aufnahmekanäle eines Arrays dabei

stark inkohärent, so führt dies in der Wiedergabe zu einer deutlich stärkeren Umhüllungswahrnehmung des Hörers und so zu einem räumlicheren Eindruck. Verfügbar sind Anordnungen der horizontal und vertikal beabstandeten Aufnahmekanäle zumeist über ein geringes Übersprechen und eine kleinere Korrelation zwischen den Kanälen, so sind die Kanäle horizontal und vertikal koinzidenter Anordnungen prinzipiell ähnlicher zueinander, was in der Wiedergabe zu weniger umhüllenden und somit weniger räumlich wahrgenommenen Klangeigenschaften führt.

6.1.4 Wahrnehmung

Die Wahrnehmungskomponenten eines Hörers bei Wiedergabe der Aufnahmen immersiver Mikrofonanordnungen teilen sich auf in drei Bereiche, welche sowohl räumliche als auch spektrale Unterschiede berücksichtigen: Die Umhüllung des Hörers und somit die Wahrnehmung der Räumlichkeit in der Wiedergabe, die Schallquellenlokalisierung in der Elevation und die Klangfarbenunterschiede in Bezug zum Spektrum der Originalschallquelle. Die Umhüllungswahrnehmung bildet eine Art Maß dafür, wie räumlich die Wiedergabe einer oder mehrerer Schallquellen vom Hörer empfunden wird, was eine erste Einschätzung bezüglich möglicher Anwendungsbereiche erlaubt: Eine Aufnahme einer sphärischen Atmo bedarf eines großen Räumlichkeitseindrucks in der Wiedergabe und so einer hohen Umhüllung des Hörers, wohingegen eine Aufnahme eines kleinen Ensembles mit einem Sänger einen stärkeren Frontbezug mit klarer Schallquellendarstellung und nicht zwingend eine hohe Umhüllung des Hörers benötigt. Die Lokalisierung von Schallquellen in der Elevation ist nicht mit allen Mikrofonanordnungen ohne weiteres möglich, wobei Arrays, welche dies ermöglichen, häufiger Probleme einer Klangverfärbung durch Kammfiltereffekte zwischen Aufnahmekanälen des Middle und Upper Layer aufweisen. Trotz allem weisen sie beispielsweise durch den Einsatz von Mikrofonen mit Kugelcharakteristik eine bessere Darstellung tiefer Frequenzen auf, welche in Verbindung mit der besseren vertikalen Schallquellenlokalisierung zum Beispiel entscheidend für Aufnahmen einer Kirchenorgel oder eines Chores auf einer Empore sind. Mikrofonanordnungen im Upper Layer mit stärker nach oben oder nach hinten gerichteten Aufnahmekanälen bilden dagegen vertikale Reflexionen im Aufnahmeraum sehr gut ab, punkten dabei allerdings nicht mit ihrer vertikalen Schallquellenlokalisierung. Ist durch Übersprechen über viele Aufnahmekanäle der Direktschallanteil in seitlichen oder oberen Mikrofonsignalen höher, so findet sich in

der Wiedergabe das gleiche Signal, durch kanalübergreifende Zeitdifferenzen verzögert - auf mehreren Wiedergabekanälen nahezu gleichzeitig wieder, was an der Hörerposition zu stärkeren Klangverfärbungen durch Kammfiltereffekte in der interauralen Domäne führen kann. Bei Wiedergabe von Ambisonicsaufnahmen und einer Hörerposition direkt im Sweetspot ist - nach dem Prinzip der physikalischen Nachbildung der Audioszene, am wenigsten Klangverfärbung wahrzunehmen, wohingegen typischerweise eine Hörerposition außerhalb des Sweetspots ebenfalls zu stärkeren Kammfiltereffekten führt.

6.1.5 Anwendungsbereiche

Aus der Auswertung der Konfigurationen, Eigenschaften und Wahrnehmungskomponenten lassen sich den einzelnen 3D Mikrofonanordnungen bestimmte Anwendungsbereiche zuordnen, welche durch jeweiliges Array bedient werden können. Grundsätzlich unterschieden wird dabei zwischen der Aufnahme von Musik, von Atmos von 360°-Content und von Live- und Broadcast-Anwendungen, welche jeweils unterschiedlichen Aufnahmeeigenschaften entsprechen. Wichtige Eigenschaften stellen dabei sowohl bei Musikaufnahmen kleiner Ensembles, als auch zum Beispiel bei der Aufnahme einer TV Show eine gute Lokalisation frontaler Schallquellen und eine zumeist klare Aufteilung in Kanäle zur Übertragung von Direkt- oder Diffusschall dar. Eine Orchesteraufnahme bedarf ebenfalls einer guten Schallquellendarstellung mit zusätzlich räumlicherer Darstellung. Das Mikrofonarray stellt dabei eine für klassische Musikaufnahmen typische Hauptmikrofonierung, vergleichbar mit einer Laufzeit- oder Koinzidentstereofonen Anordnung. Aufnahmeanordnungen mit nach oben oder hinten gerichteten Mikrofonen im Upper Layer werden dabei häufig im Sinne einer typischen Hauptmikrofonierung oberhalb eines Orchesters oder einer anderen Schallquelle angeordnet, sodass eine gute horizontale Abbildung mit einer Abbildung vertikaler Reflexionen an Decke und Wänden optimal verbunden werden kann. Zu berücksichtigen ist dabei jeweils auch der Hallabstand. Soll hingegen eine Kirchenorgel oder eine andere musikalische Schallquelle in der Elevation aufgenommen werden, bedarf es beispielsweise der Verwendung von Mikrofonen einer Kugelcharakteristik und der somit besseren Elevationsdarstellung in Verbindung mit der Aufnahme tieferer Frequenzen. Geht es um die Aufnahme von Geräuschen in Form einer Atmo kommt es darauf an, ob das Schallfeld mehr direkte Schallquellen mit potentieller Lokalisation dieser beinhaltet oder ob es sich um eine eher sphärische Geräusch-Atmosphäre handelt. Einige Arrays sind speziell zur

Aufnahme möglichst räumlicher Klanganteile ausgelegt, bieten hingegen eine schlechtere oder diffuse Schallquellenlokalisierung, was sie prinzipiell für Atmoaufnahmen sehr geeignet macht. Sind Schallquellen rund um das Mikrofonarray angeordnet, spricht man von einer 360°-Aufnahme, welche sich gerade in Bezug auf Virtual Reality-Anwendungen großer Beliebtheit erfreut. Prinzipiell sind solche Aufnahmen auch gut mit einigen Anordnungen der horizontal und vertikal beabstandeten Mikrofone zu realisieren, scheitern allerdings häufig an praktischen Gegebenheiten der Portabilität, sodass horizontal und vertikal koinzidierte Anordnungen durch ihre beinahe physikalisch korrekte Reproduktion des Schallfeldes und der deutlichen portableren Anordnungen für dieses Anwendungsgebiet zumeist von Vorteil sind.

3D Microphone Arrays

Categorization and Rating of the Arrays, concerning field of application and playback of their recordings

Speaker related	Configuration			Characteristics		Perception			Field of application												
	Phys. Layout	Portability	vertical mic-orientation	interch. cross-talk	interch. correlation	Listener envelopment	Localisation of elevated sound	Sound coloration	Music-recording	Small en-semble	Orchestra/choir	Church organ	Ambience-recording	with more direct sounds	spherical ambience	360°-recording	Sound sources around array	Live / Broadcast	Live-recording/ TV show		
Sound System D (9.1)	Williams Umbrella (4+4+0)	-	UfUd	+	--	++	-	+					X								
	Twin Cube (4+4+0)	-	FfOd	+	+	+	-	+		X											
	ORIT-3D (4+4+0)	++	UfUd	--	-	++	+	-						X		(X)					
	ESMA-3D (4+4+0)	+	UfUd	-	-	+	+	-						X							
	Lee Rec-3D (4+4+0)	+	UfUd	--	-	++	+	-						X							
	Hamasaki Cube (4+4+0)	-	UfUd	--	-	++	+	+	--					X							
	OCT-3D (4+5+0)	-	UfUd	--	-	+	+	+	-	X	(X)			(X)							X
	Bowles-Array (4+5+0)	-	UfUd	-	+	-	-	+	+	X	X										
	2L-Cube (4+5+0)	-	FfOd	++	+	+	-	--	++		X		X								X
	Spider Tree (4+5+0)	-	UfUd	+	+	+	+	++	+	X											X
	Twin Cube+M+000 (4+5+0)	-	FfOd	+	+	+	+	-	+		X										(X)
	Double-UFIX (4+5+0)	-	UfUd	++	+	+	+	--	++	(X)	X		X								
	PCMA-3D (4+5+0)	+	UfUd	--	-	+	+	+	-	X	X										X
	Au3dio (4+6+0)	+	UfUd	+	-	-	-	+	+	X											X
	Hamasaki et al. (9+10+3)	--	UfUd	+	+	--	++	++	+		X										
	Howie et al. (9+10+3)	--	FfOd	+	+	--	++	++	+		X										
	Independent* First order Ambisonics	Triple MS (B-Format)	++	/	-	++	--	+	--		X	(X)			(X)						X
tetrahedral Array (A-Format)		++	/	+	++	--	+	--		X	(X)			(X)						X	
Spherical Arrays (A-Format)		++	/	+	++	-	+	-		(X)	(X)			X	(X)					(X)	

Rating: + + = most + = more - = less -- = least (X) = Array conditionally suitable for application area

UfUd = Upward or backward facing unidirectional microphones in upper layer FfOd = Frontal facing unidirectional or omnidirectional microphones in upper layer

HVS = horizontally and vertically spaced HSVC = horizontally spaced, vertically coincident HVC = horizontally and vertically coincident

*Independent of playback formats

6.2 Lautsprecherbezogene Mikrofonanordnungen

6.2.1 Mikrofonarrays mit Anordnung 4+4+0 zur Wiedergabe über Sound System D (9.1)

6.2.2.1 Williams-Umbrella

Die Mikrofonanordnung des Williams-Umbrella stellt eine horizontal und vertikal beanstandete Konfiguration aus vier im Middle Layer angeordneten Nieren und vier im Upper Layer angeordneten Achtern dar, welche durch Michael Williams im Mai 2013 auf der 134. Convention of the Audio Engineering Society vorgestellt wurde. Die vier Mikrofone im Middle Layer der Richtcharakteristik Niere sind dabei in einem kleinen Quadrat mit Abständen von jeweils 35cm angeordnet, wobei die Mikrofone jeweils in diagonale Richtungen von $+45^\circ$, -45° , $+135^\circ$ und -135° zeigen (Lee, 2021). Das Upper Layer bildet mit seinen vier vertikal ausgerichteten Mikrofonen der Richtcharakteristik Acht ebenfalls ein Quadrat einer Kantenlänge von 52cm, wobei die Anordnung im Vergleich zum Middle Layer um 45° auf die Positionen 0° , $+90^\circ$, -90° und 180° gedreht und einen Meter vertikal nach oben versetzt wurde (Lee, 2021). Durch die horizontal und vertikal beanstandete Anordnung in der Diagonalen, liegt der Vorteil dieses Arrays gerade auch in der guten diagonalen Abbildung von Schallquellen, welches durch aus den Abständen begründeten stärkeren Zeit- und Pegeldifferenzen zwischen den Kanälen in seinen Aufnahmen ein recht räumliches Abbild des Aufnahmeraums schafft. Vertikal angeordnete Schallquellen können durch den Einsatz der Achtermikrofone im Upper Layer allerdings nicht nur über dem Array, sondern auch zwischen Upper und Middle Layer lokalisiert werden, was zu einer unstabilen Abbildung von Schallquellen führen würde, welche sich in der Elevation etwa unter $26,5^\circ$ befinden (Lee, 2021).

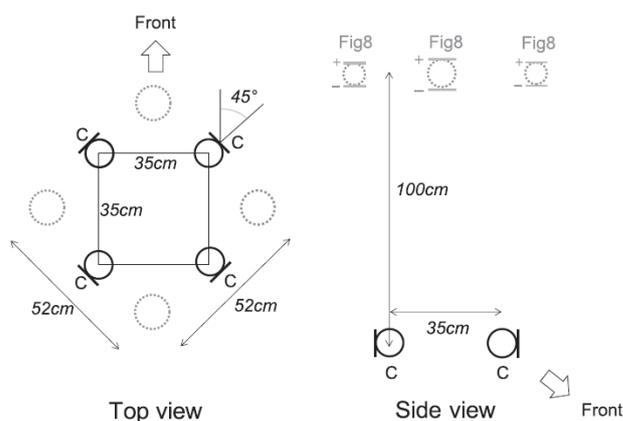


Abb. 17. Drauf- und Seitenansicht einer Anordnung des Williams Umbrella

6.2.2.2 Twin Cube

Das Array des Twin Cubes stellt eine Anordnung von acht Mikrofonen in einem großen Würfel dar, dessen Kantenlängen mit 2m im Vergleich zu vorigen Quadratanordnungen größere Dimensionen einnimmt. In einer Kooperation mit Sennheiser entwickelte Gregor Zielinsky sein 3D-Mikrofonarray eines Würfels aus bis zu neun MKH 800 Twin Doppelmembranmikrofonen, welche basierend auf der A-B Hauptmikrofonierung horizontal und vertikal weit versetzt wurden (Eskow, 2016). Der Einsatz der Doppelmembranmikrofone, welche sowohl im Upper-, als auch im Middle Layer nach vorne ausgerichtet werden, führt zu unterschiedlichen möglichen Richtcharakteristiken im Nachhinein der Aufnahme: Für jeden Aufnahmekanal können durch das Mischungsverhältnis in der zweikanaligen Aufnahme zweier entgegengesetzter Nieren Richtcharakteristiken von Kugel bis hin zur Acht erreicht werden. Haben die Aufnahmekanäle des Upper Layer dabei allerdings die Richtcharakteristik einer nach vorne zeigenden Niere oder einer Kugel, so führt das resultierende kanalübergreifende Übersprechen vermutlich zu unstabilerer Phantomschallquellenabbildung in der Elevation (Lee, 2021). Auch macht die große Dimension der Mikrofonanordnung dieses nicht für alle Anwendungen nützlich: Besonders die großen Mikrofonabstände im Sinne möglichst großer Zeitdifferenzen zwischen allen Kanälen machen das Array vor allem zur Aufnahme ausgebreiteter Schallquellen, wie z.B. einem großen Orchester, geeignet (Lee, 2021).

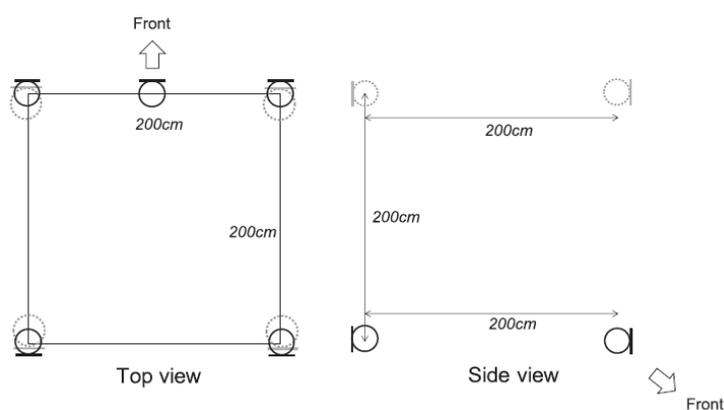


Abb. 18. Drauf- und Seitenansicht einer Anordnung des Twin Cubes

6.2.2.3 ORTF-3D

Die Anordnung in ORTF-3D, entwickelt durch Helmut Wittek und Günther Theile, bildet durch seine horizontal beabstandete, aber vertikal koinzidente Konfiguration von acht

Aufnahmekanälen in der Praxis eine optimale immersive Aufnahmemöglichkeit von Atmosphären. Als gemischtes Aufnahmeverfahren aus Laufzeit- und Koinzidenzstereofonie ist ORTF (Office de Radiodiffusion Télévision Francaise) ein beliebtes Prinzip der zweikanaligen Aufnahme, bestehend aus zwei Nieren im Abstand von 17cm und einem Versatzwinkel im Bezug zur Front von jeweils 45° (Lee, 2021). Aus dem Konzept einer Surroundanordnung des ORTF-Prinzips heraus, in der in einer rechteckigen Anordnung eine gleiche Kombination von Mikrofonen nach hinten ausgerichtet wurde, ergänzte Schoeps diese um zwei weitere ORTF-Paare, welche an gleichen Positionen jeweils schräg nach oben ausgerichtet wurden. ORTF-3D gibt es in der Version eines quadratischen Indoor-Sets mit einer Kantenlänge von 18cm und der eines rechteckigen Outdoor-Sets mit 20cm Weite zu 10cm Tiefe, sodass dieses auch in einen entsprechenden Windkorb passt (Lee, 2021). Wie auch in der stereofonen Aufnahme im ORTF-Prinzip, verfügen auch die acht Kanäle des ORTF-3D über eine sehr gute Signaltrennung und somit über geringes kanalübergreifendes Übersprechen (Wittek & Theile, 2017).

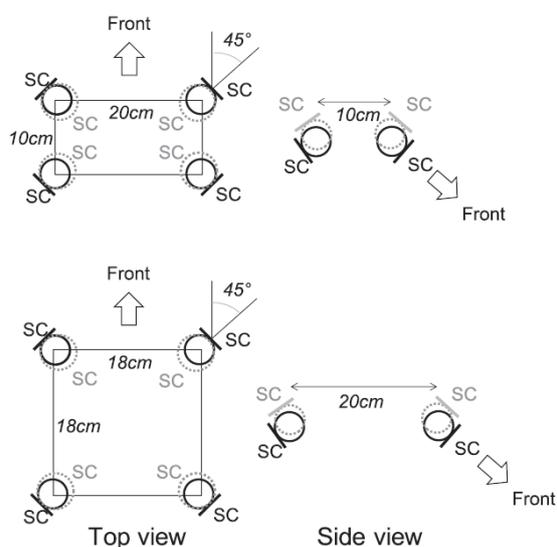


Abb. 19. Drauf- und Seitenansicht der Anordnungen nach ORTF-3D in Indoor- und Outdoor-Version

6.2.2.4 ESMA-3D

ESMA (Equal segment microphone array) wurde ursprünglich von Michael Williams zur 360° -Darstellung eines Schallfeldes entwickelt, wobei die Anordnung der Mikrofone im Aufnahmebereich strikt der Anordnung der Lautsprecher im Abhörbereich entsprechen sollte (Lee, 2021). Ziel dieser diskreten Verbindung ist dabei die mehrkanalige Aufnahme ohne entstehende Überlappungen oder Abbildungslöcher (Lee, 2021). Bei der Anordnung in

ESMA-3D besteht das Middle Layer aus vier diagonal quadratisch angeordneten Nieren mit einem Abstand von 50cm zueinander und Versatzwinkeln von $+45^\circ$ und $+90^\circ$. Die Aufnahmekanäle des Upper Layers befinden sich auf gleicher vertikalen Ebene wie die des Middle Layers, sind allerdings mit stärkerer Richtwirkung einer Superniere nach oben gerichtet.

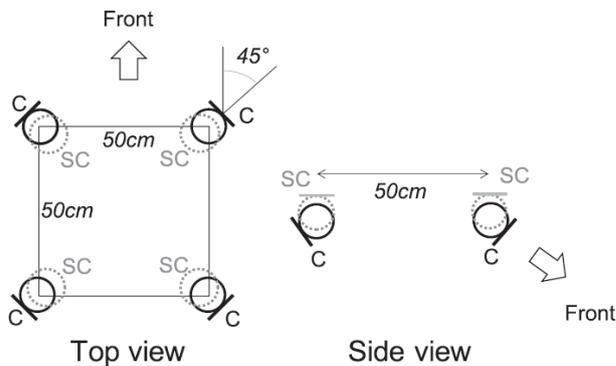


Abb. 20. Drauf- und Seitenansicht einer Anordnung des ESMA-3D

6.2.2.5 Lee Rec-3D

Aus seinen Erkenntnissen aus der Auswertung früherer 3D-Mikrofonarrays heraus, entwickelte Hyunkook Lee mit seinem Lee Rec-3D eine horizontal beanstandete (1m Weite zu 0,5m Tiefe) und vertikal koinzidente Mikrofonanordnung mit 4+4+0 Kanälen (Lee, 2021). In der immersiven Wiedergabe aus Arrays einer Weite von 1m oder 2m gab es laut Lee keinen signifikanten Unterschied der wahrgenommenen Breite oder Einhüllung, sodass er mit 1m den kleinstmöglichen Abstand in der Weite des Arrays übernahm (Lee, 2021). So verglich er auch die Eigenschaften von Arrays der Tiefendimensionen von 0,5m, 1m und 2m, wobei er auch hier keinen signifikanten Unterschied der wahrgenommenen Tiefe oder Einhüllung erkannte (Lee, 2021). Für das Middle Layer bietet das Lee-Rec-3D Array drei mögliche Optionen: vier nach hinten gerichtete Nieren, vier seitwärts gerichtete Achten oder zwei seitwärts gerichtete Achten zu zwei nach hinten gerichteten Nieren zur Repräsentation der jeweiligen Kanäle vorne und hinten (Lee, 2021). Das Upper Layer besteht jeweils aus vier nach oben gerichteten Supernieren, welche sich vertikal auf gleicher Ebene des Middle Layers befinden.

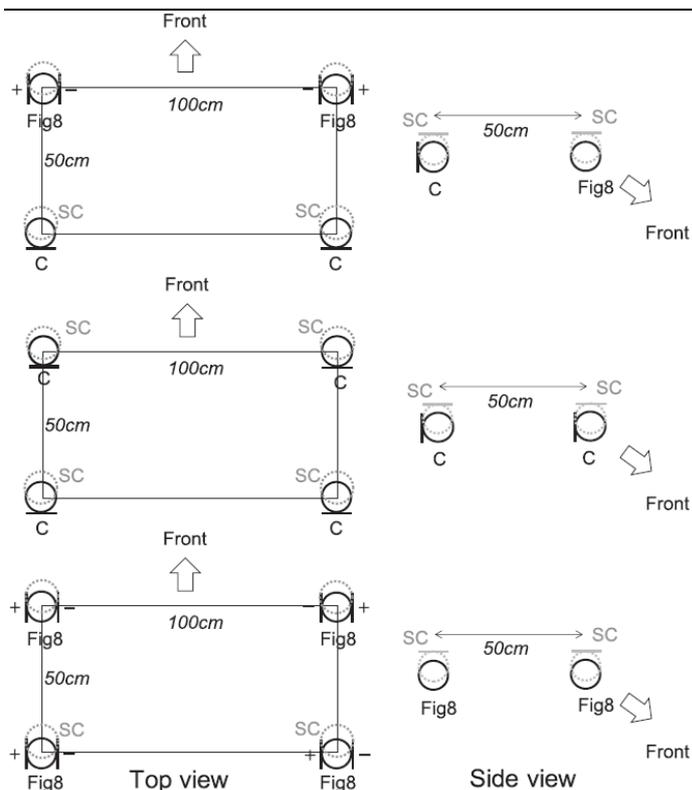


Abb. 21. Drauf- und Seitenansichten von Anordnungen des Lee Rec-3D

6.2.2.6 Hamasaki Cube

Die 3D-Mikrofonanordnung des Hamasaki Cube, vorgestellt im Mai 2015 durch Kimio Hamasaki und Wilfried Van Baelen auf der 138. Convention of the Audio Engineering Society, bildet eine Formation aus acht Aufnahmekanälen in einem Würfel großer Dimensionen mit 2m Kantenlänge. Das Array basiert dabei auf auf der populären vierkanaligen Aufnahmetechnik des Hamasaki Square zur Aufnahme dekorrelierter Atmosphären, welche für die immersive Mikrofonanordnung das Middle Layer bildet (Lee, 2021). Das Quadrat des Middle Layer besteht dabei aus vier seitlich ausgerichteten Mikrofonen der Achter-Charakteristik, welche für die Front- und Rückkanäle durch ihren frontal liegenden Nullpunkt ein deutlich größeres Verhältnis erster Reflexionen zu Direktschallanteilen haben und so eine höhere Einhüllungswahrnehmung in der Wiedergabe der Aufnahmen erreichen (Lee, 2021). Das Upper Layer besteht aus vier nach oben gerichteten Supernieren, welche vertikal direkt über den Mikrofonen des Middle Layers angeordnet sind und Reflexionen von der Decke in der Elevation abbilden sollen. Der große Mikrofonabstand gilt natürlich als großer Nachteil im Bezug auf die Portabilität, wobei dieser auch den Vorteil einer vollen kanalübergreifenden Dekorrelation über 100Hz mit sich bringt (Lee, 2021).

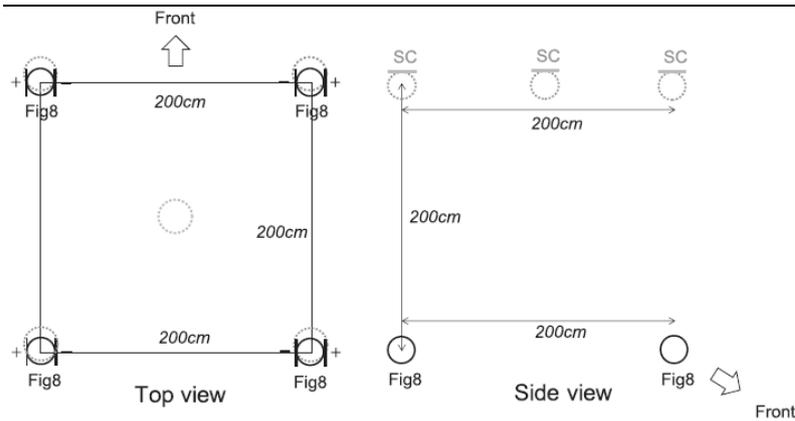


Abb. 22. Drauf- und Seitenansicht einer Anordnung des Hamasaki Cubes

6.2.2 Mikrofonarrays mit Anordnung 4+5+0 zur Wiedergabe über Sound System D (9.1)

6.2.3.1 OCT-3D

Das durch Helmut Wittek und Günther Theile vorgestellte Array des OCT-3D stellt eine horizontal und vertikal beabstandete Anordnung von Mikrofonen dar, welche sich in fünf Kanäle des Middle Layer und vier Kanäle des Upper Layer aufteilen. Der Unterschied zu Anordnungen mit 4+4+0 Kanälen ist die Ergänzung um einen weiteren Kanal zur Abbildung der Front, diskret verbunden zum Center-Kanal einer Surround-Wiedergabe. Das Middle Layer besteht dabei aus einer typischen fünfkanaligen Anordnung nach OCT-Surround, entwickelt durch Wittek, wobei sich das Upper Layer aus vier 1m direkt oberhalb befindlichen nach oben gerichteten Supernieren besteht (Lee, 2021). Die Anordnung in OCT-Surround bildet ein typisches Verhältnis aus drei Frontkanälen zu zwei Surroundkanälen (3/2), wobei alle Kanäle auf eine möglichst große Kanaltrennung untereinander ausgelegt sind, um ein Übersprechen zwischen diesen zu verhindern. So werden links und rechts durch nach außen zeigende Supernieren und links-hinten und rechts-hinten durch nach hinten gerichtete Nieren abgebildet. Zusätzlich sind der Centerkanal leicht nach vorne und die beiden Surroundkanäle leicht nach außen versetzt.

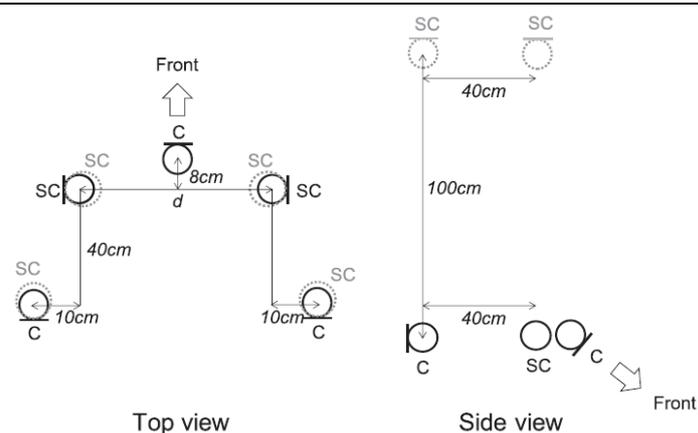


Abb. 23. Drauf- und Seitenansicht einer Anordnung des OCT-3D

6.2.3.2 Bowles-Array

Das Bowles-Array umfasst eine Kombination von Nieren- und Kugelmikrofonen im Middle Layer mit Supernieren zur Abbildung des Upper Layer, welche 2015 durch David Bowles vorgestellt wurde. Die vier Kanäle links, rechts, links-hinten und rechts-hinten werden dabei durch Mikrofone der Richtcharakteristik einer Kugel abgebildet, wohingegen der Centerkanal durch eine nach vorne gerichtete Niere aufgenommen wird. Die Kanäle des Upper Layer bestehen aus vier um 60° geneigten nach oben gerichteten Supernieren, welche vertikal über den Kanälen des Middle Layers angeordnet sind. Die horizontalen und vertikalen Abstände zwischen den Aufnahmekanälen können, nach Bowles, je nach Größe des aufgenommenen Ensembles verändert werden, was gerade der Aufnahme musikalischer Schallquellen zugutekommen kann. Um ein vertikales Übersprechen und etwaige Kammfiltereffekte zu verhindern, entschied sich Bowles bei den Mikrofonen zur Aufnahme des Upper Layers gegen die Kugelcharakteristiken, wie sie im Middle Layer zum Einsatz kommen.

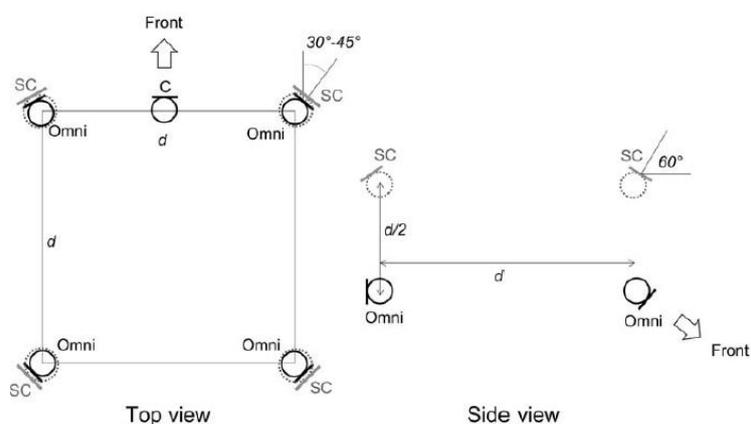


Abb. 24. Drauf- und Seitenansicht einer Anordnung des Bowles Array

6.2.3.3 2L-Cube

Die horizontal und vertikale Anordnung des 2L-Cube stellt ein würfelförmiges Array dar, welches durch den Decca- oder Mercury-tree inspiriert und von Morten Lindberg entwickelt wurde (Lindberg, o. D.). Der Würfel aus Aufnahmekanälen setzt sich dabei aus acht Kugelmikrofonen mit einem nach Ensemblegröße variierenden Abstand von 40cm bis 120cm und einem etwas weiter vorne angeordneten Kugelmikrofon zur Abbildung des Centerkanals zusammen (Lindberg, o. D.). In Abhängigkeit des Raumes, der aufzunehmenden Musik und der Instrumente variiert Lindberg zwischen zwei Mikrofonen der Marke DPA (4003 und 4041). Bei der Aufnahme großer Orchester wird das Array häufig in die Mitte einer kreisförmigen Formation der Musiker gebracht und erzielt somit eine 360° Abbildung der Schallquellen (Lee, 2021). Die eigentliche Mischung entsteht dabei zumeist schon bei der Aufnahme durch unterschiedliche Abstände der Musiker zum Array. Die Wahl der Mikrofone mit Kugelcharakteristik beruht dabei bei der Aufnahme klassischer Musik hauptsächlich in der deutlich besseren Abdeckung tiefer Frequenzen im Aufnahmespektrum und weniger in räumlichen Gegebenheiten (Lee, 2021).

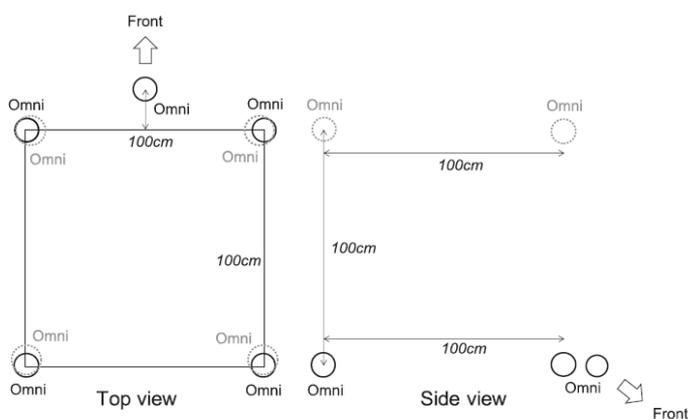


Abb. 25. Drauf- und Seitenansicht einer Anordnung des 2L-Cubes

6.2.3.4 Spider Tree

Anders als bei quadratisch- oder würfelförmigen Mikrofonanordnungen sind die Aufnahmekanäle des Spider Trees, entwickelt durch Mick Sawaguchi, in Surroundanordnung vom Mittelpunkt des Arrays aus angeordnet, was zu einer fünfeckigen Formierung des Middle Layers des 3D-Arrays führt. Die Mikrofone haben dabei einen jeweiligen Abstand von 90-110cm zum Mittelpunkt (Lee, 2021). Ähnlich wie beim Array des 2L-Cubes befinden sich auch beim Spider Tree die Musiker eines Ensembles

oder Orchesters rund um die Mikrofonanordnung des Middle Layers. Mit der Aufnahme der Vertikalen experimentierte Sawaguchi durch einige unterschiedliche Anordnungen der Upper Layer-Kanäle. Bei der klassischen Musikaufnahme eines Sextetts befanden sich die Aufnahmekanäle des Middle Layer zum Beispiel in der Mitte zwischen den Musikern, wohingegen das Upper Layer einer quadratische Anordnung nach oben gerichteter Nieren entsprach, welche leicht hinter den Musikern angeordnet waren (Sawaguchi, 2017).

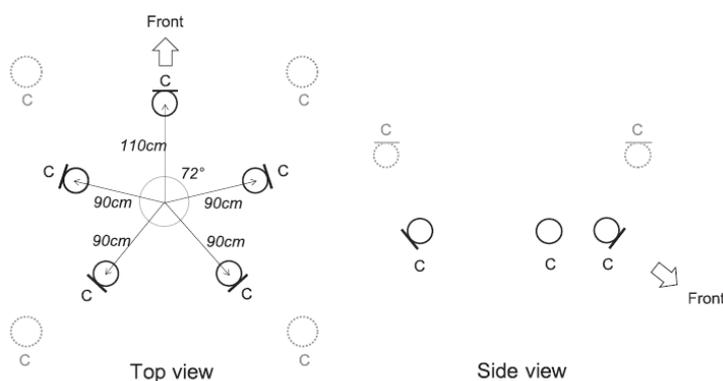


Abb. 26. Drauf- und Seitenansicht einer Anordnung des Spider Tree

6.2.3.5 Twin Cube (4+4+0) mit Ergänzung um M+000

Die Anordnung des Twin Cube mit 4+5+0 Kanälen setzt sich prinzipiell aus der in Abschnitt 6.2.2.2 beschriebenen Anordnung des Twin Cube mit 4+4+0 Kanälen und einem zusätzlichen Aufnahmekanal zur Abbildung des Centerkanals (M+000) zusammen. Die Ergänzung um einen mittig angeordneten weiteren frontalen Kanal führt in der Wiedergabe typischerweise zu einer besseren Lokalisation frontaler Schallquellen und frontal angeordneter Phantomschallquellen zwischen den Kanälen.

6.2.3.6 Double-UFIX

Double-UFIX wurde als neunkanalige 3D-Mikrofonanordnung durch Florian Camerer mit dem Ziel einer Aufnahmemöglichkeit von Atmosphären on Location zur Wiedergabe über ein AURO-3D Setup (Sound System D) entwickelt (Camerer, 2020). Das Array wurde dabei strikt nach dem Konzept des optimalen Stereo-Aufnahmewinkels in Bezug zu den Winkeln jeweiliger horizontaler Lautsprecherpaare in der Wiedergabe konzipiert (Camerer, 2020), was zu einer klar nachvollziehbaren diskreten Verbindung zwischen Aufnahme- und Wiedergabekanälen führt. Dies zeigt sich allerdings auch in einer etwas unregelmäßigeren Anordnung des Middle und Upper Layers, beispielsweise im Vergleich zu würfelförmigen Arrays. Nach dem Konzept der horizontal und vertikal beabstandeten

Anordnung der Mikrofone zueinander und mit deren weiteren Aufnahmewinkeln, wird die Dekorrelation zwischen den Kanälen hauptsächlich durch kanalübergreifende Zeitdifferenzen erreicht. Camerer (2020) entscheidet sich daher bei allen neun Aufnahmekanälen für die Richtcharakteristik einer weiten Niere, was - verglichen mit einer Niere - dem Spektrum der Aufnahme zusätzliche tiefe Frequenzen beschafft (Lee, 2021). Der optimale Aufnahmewinkel der Mikrofonpaare könnte dabei allerdings durch ein aus der geringen Richtwirkung der weiten Nieren wirkendes kanalübergreifendes Übersprechen beeinflusst werden (Lee, 2021). Die vertikale Dekorrelation ist durch den vertikalen Abstand zwischen Middle und Upper Layer von 25-30cm etwa bis hinunter zu 500Hz gegeben, unter dieser Grenzfrequenz ist die vertikale Abbildung nur begrenzt möglich (Camerer, 2020). Einen Vorteil, gerade für Außenaufnahmen von Atmosphären, bietet die Möglichkeit einer Anordnung in zwei übereinander angeordneten Windkörben von CINELA, wobei Camerer (2020) das Setup als für eine Person schwer zu transportieren beschreibt, wonach er nach weiteren Transportierungsmöglichkeiten beispielsweise in Form eines auf Rollen angeordneten Setups sucht.

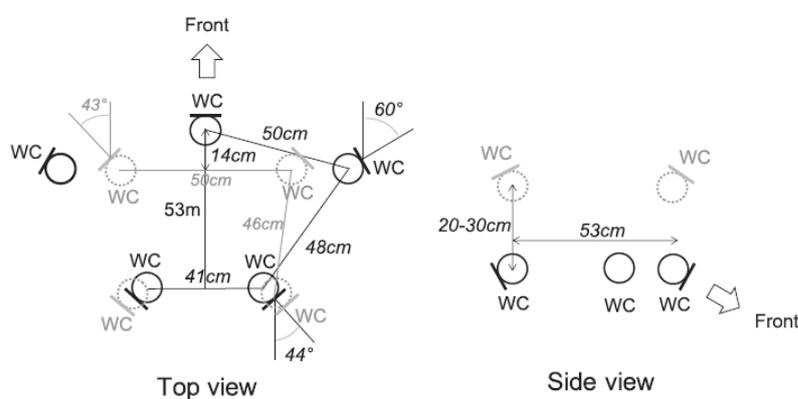


Abb. 27. Drauf- und Seitenansicht einer Anordnung des Double-UFIX

6.2.3.7 PCMA-3D

Das Mikrofonarray des PCMA-3D ergänzt die 2011 durch Hyunkook Lee auf der 130. Convention of the Audio Engineering Society vorgestellte Surroundanordnung des PCMA (Perspektive control Microphone array) mit 0+5+0 Kanälen um vier weitere Höhenkanäle zur Wiedergabe über das Sound System D (9.1). Ähnlich des Konzepts des Twin Cubes kommen für die Front- und Rückkanäle jeweils Doppelmembranmikrofone zum Einsatz, welche in der Postproduktion der Aufnahmen noch in ihrer virtuellen Richtcharakteristik beeinflusst werden können (Lee, 2021). Zur Anordnung der Aufnahmekanäle des Upper Layers beruft sich Lee (2021) auf Erkenntnisse der räumlichen Wahrnehmung, wonach ein

vertikaler Abstand zwischen Middle und Upper Layer keinen signifikanten Unterschied in der wahrgenommenen Räumlichkeit aufweist. Auch lieferte ein Downmix von 3D auf 2D bei vertikal koinzidenten Anordnungen durch weniger Klangverfärbung ein besseres Ergebnis (Lee, 2021). In der Konfiguration des PCMA-3D empfiehlt Lee (2021) die Verwendung vier nach oben gerichteter Supernieren, um das kanalübergreifende Übersprechen möglichst gering zu halten. Um das Übersprechen so gut wie vollständig zu verhindern, können die Mikrofone des Middle und Upper Layer auch so geneigt werden, dass ihre Richtcharakteristiken genau gegeneinander zeigen (Lee, 2021).

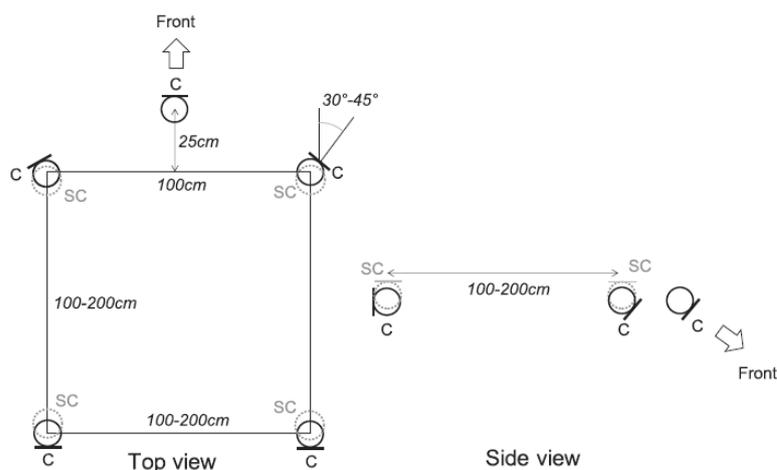


Abb. 28. Drauf- und Seitenansicht einer Anordnung des PCMA-3D

6.2.3 Mikrofonarrays mit Anordnung 9+10+3 zur Wiedergabe über Sound System H (22.2)

6.2.4.1 Hamasaki et al.

In der Entwicklung immersiver Aufnahmemöglichkeiten wurden unter anderem auch wenige Möglichkeiten einer Hauptmikrofonierung mit diskretem Routing auf das Wiedergabeformat des Sound System H (22.2) mit 9+10+3 Kanälen vorgestellt. Kimio Hamasaki, Toshiyuki Nishiguchi, Koichiro Hiyama und Kazuho Ono lieferten dabei 2004 auf der 116. Convention of the Audio Engineering Society erste Ansätze eines Arrays zur immersiven Aufnahme großer Orchester in einem solchen Format, welches durch seine große Kanalanzahl eine strikte Trennung zwischen Direkt- und Diffusschallaufnahme erlaubte. Durch diese Trennung in eine Haupt- und eine Diffusmikrofonierung nennt man Arrays dieser oder ähnlicher Anordnung auch Main-/ Ambience Array. In der Anordnung des Middle Layer wurde eine Art Line-Array aus fünf Supernieren mit gleichen Abständen

zwischen den Mikrofonen direkt vor dem Orchester angeordnet (Lee, 2021). Zur Aufnahme des Diffusfallschalls wurden 13 seitwärts gerichtete Achtermikrofone im gesamten Aufnahmeraum verteilt, um durch große Abstände der Kanäle zueinander eine starke Dekorrelation dieser zu erzielen, wobei eine genaue Anordnung der Abstände durch Hamasaki et al. (2004) nicht beschrieben wurde (Lee, 2021).

6.2.4.2 Howie et al.

Will Howie, Richard King, Denis Martin und Florian Grond stellten im Mai 2017 auf der 142. Convention of the Audio Engineering Society ihre Lösungen zur experimentellen Aufnahme in 9+10+3 Kanälen vor, welche Hamasaki et al.'s Ansatz von 2004 aufgriffen und erweiterten. Sie definierten somit auch genaue Abstände zwischen einzelnen Mikrofonkanälen und Middle und Upper Layer, sodass z.B. das Main-Array fünf frontal aufgerichteter Supernieren Mikrofonabstände von 2,16m zugeordnet bekam. Acht seitwärts gerichteten Achtermikrofonen stehen acht nach oben gerichteten Supernieren zur Aufnahme des Diffusfallschalls entgegen, welche weit im Zuschauerraum des Konzertsaals verteilt sind (Lee, 2021). Die drei Frontkanäle des Lower Layer bilden in der Aufnahme drei weit voneinander entfernte Mikrofone ab, die auf dem Boden der Bühne platziert werden (Lee, 2021). Der große vertikale Abstand von 3,67m zwischen Middle und Upper Layer ist dabei deutlich höher als bei anderen horizontal und vertikal beabstandeten Anordnungen (Lee, 2021). Eine weitere Möglichkeit der immersiven Aufnahme in 9+10+3 Kanälen stellten Howie et al. 2017 vor, welche als Main-Array die Anordnung eines Decca Tree in Ergänzung um weit im Aufnahmeraum verteilte Mikrofone zur Diffusfallschallaufzeichnung beinhaltet.

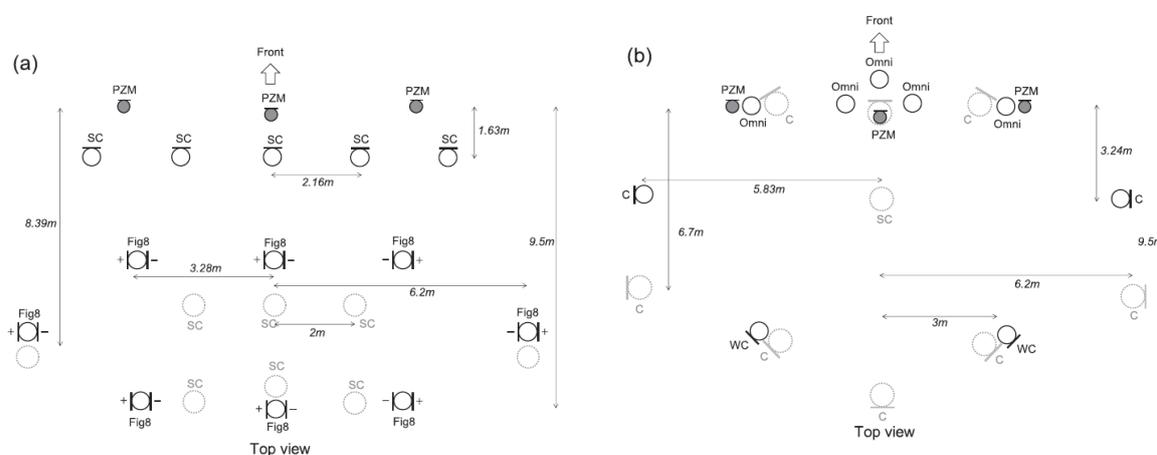


Abb. 29. Drauf- und Seitenansicht von Anordnungen nach Howie et al.

6.2.4 Mikrofonarrays zur Wiedergabe über spezielle Lautsprecherkonfigurationen

6.2.5.1 Au3dio (4+6+0)

Die Anordnung des au3Dio wurde durch Christian Viada im Jahre 2017 auf der International Conference on Spatial Audio als 4+6+0 Array vorgestellt (Lee, 2021). Sechs Nierenmikrofone in einer hexagonalen Anordnung mit jeweils 62cm Abstand zur Mitte bilden dabei das Middle Layer, wobei die vorne und hinten liegenden Mikrofone durch nach oben zeigende Achtermikrofone auf gleicher vertikaler Ebene ergänzt werden (Lee, 2021). Kleine Ensembles können, ähnlich wie beim Spider Tree, rund um das Array direkt vor den Mikrofonen angeordnet werden, was den Aufnahmekanälen des Upper Layer eine maximale Reduzierung des Direktschallanteils ermöglichen würde (Lee, 2021). Auch eine Kombination der Nieren und Acht-Mikrofone zu vertikalen MS-Paaren wäre möglich, wobei das Array in diesem Fall über der Schallquelle angeordnet werden sollte (Lee, 2021).

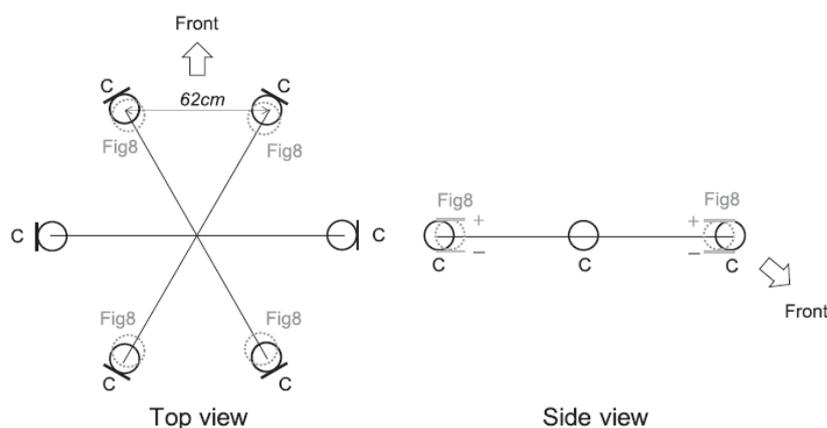


Abb. 30. Drauf- und Seitenansicht einer Anordnung des au3Dio

6.3 Lautsprecherunabhängige Mikrofonanordnungen

Schaut man sich Mikrofonanordnungen einer horizontal und vertikal koinzidenten Form an, so sind die Verhältnisse der Aufnahmekanäle durch ausschließliche Pegeldifferenzen geprägt und nicht wie bei beabstandeten Anordnungen auch durch Zeitdifferenzen. Diese Art der Stereophonie wird, wie in Kapitel 4.1.2 bereits erwähnt, als Koinzidenz- oder Intensitätsstereophonie bezeichnet, deren Weiterentwicklungen zu koinzidenten Arrays zur Aufnahme von Surround -oder 3D-Inhalten führten. First Order Ambisonics ist dabei ein

Format zur Darstellung einer dreidimensionalen Audioszene aus vier Signalen (W, X, Y und Z). Diese Ambisonics-Signale können entweder in bestimmten Anordnungen diskret aufgenommen oder aus anderen Anordnungen konvertiert werden. Geht es um die Aufnahme in einem Ambisonics-Format höherer Ordnung, kommen weitere Anteilssignale hinzu, welche jeweils mehrere Richtungen des Schallfeldes beschreiben.

6.3.1 Aufnahme in First-Order Ambisonics

6.2.3.1 Triple MS

Aus der Koinzidenzstereofonie des MS entwickelten Zotter und Frank (2019) eine doppelte Anordnung des Seitensignals, wobei ein Achtermikrofon dabei um 90° gedreht wurde, um eine Repräsentation des Schallfeldes in der zweidimensionalen Ambisonics-Domäne zu ergründen. Um daraus ein dreidimensionales FOA-Array zu machen, wurde das Array um eine weitere Acht ergänzt, welche nach oben ausgerichtet wurde (Zotter & Frank, 2019). Diese Anordnung bildet dabei nahezu das Einzige native Array zur diskreten Aufnahme der Signalkomponenten W, X, Y und Z dar, wobei ein Kugelmikrofon W, eine frontal ausgerichtete Acht X, eine seitlich ausgerichtete Acht Y und eine senkrecht ausgerichtete Acht Z direkt abbilden. Die Domäne der Ambisonics-Signale wird dabei auch als B-Format bezeichnet. Die Anordnungen des Doppel- und Dreifach-MS sind auch durch portable Outdoor-Lösungen im Windkorb, beispielsweise von Schoeps, flexibel einsetzbar und finden ihre Anwendung in vielen Bereichen.

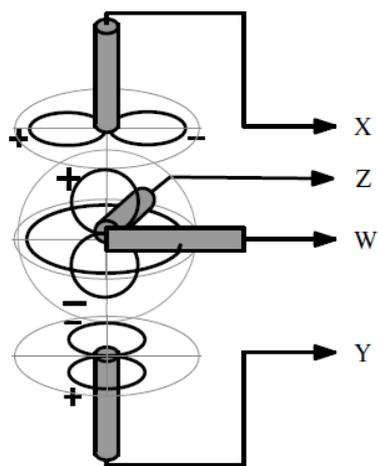


Abb. 31. Darstellung einer Triple-MS-Anordnung.

6.2.3.2 Tetrahedral Arrays

Auch aus einer Anordnung von vier koinzident angeordneten Nieren in einer tetraederartigen Form können durch eine Matritzierung die vier Ambisonics-Signale gewonnen werden, sodass in der Aufnahme keine unterschiedlichen Richtcharakteristiken mehr vonnöten sind. In der Encodierung müssen für das W-Signal alle Signale addiert, für X das Signal der nach hinten gerichteten Niere von dem der nach vorne gerichteten abgezogen, für Y das rechte vom linken Signal abgezogen und für Z das untere vom oberen Signal abgezogen werden (Zotter & Frank, 2019). So wird das sogenannte A-Format in das B-Format und somit in die dreidimensionale Ambisonics-Domäne gewandelt, aus welcher es zur Wiedergabe über Lautsprecher erst auf das entsprechende Lautsprecherlayout dekodiert werden muss. Auch das Array aus vier Nieren erfreut sich in der Anwendung durch seine unschlagbare Portabilität großer Beliebtheit, beispielsweise in kommerziellen Lösungen von Sennheiser, Oktava oder Soundfield. Doch auch wenn die Mikrofonanordnungen so wenig versetzt wie möglich gebaut werden, so können diese gerade für hohe Frequenzen nicht mehr als koinzident betrachtet werden, was zu veränderten Richtwirkungen und einem Verlust der Präsenz hoher Frequenzen im Diffusfeld gleichkommt (Zotter & Frank, 2019). Daher werden durch high-shelf Filter hohe Frequenzen typischerweise um 3dB angehoben (Zotter & Frank, 2019).

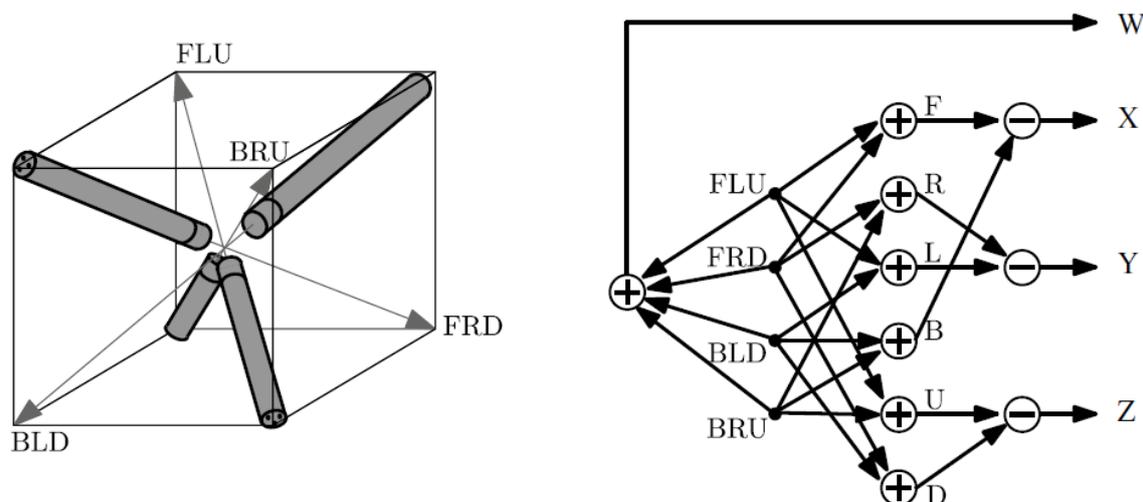


Abb. 32. Darstellung einer tetraederartigen Anordnung und ihrer Codierung in's Ambisonics B-Format

6.3.2 Aufnahme in Higher-Order Ambisonics

6.2.3.1 Spherical Arrays

Für die Aufnahme einer höher aufgelösten Audioszene im Sinne des Higher-Order Ambisonics ist eine Erweiterung der Teilsignale um weitere Richtungsinformationen notwendig. Da höhere Richtwirkungen in mehrere Richtungen als eine Achtercharakteristik als einzelne Mikrofonkanäle physikalisch nicht zu realisieren sind, werden notwendige Richtcharakteristiken virtuell gebildet, weshalb keine native Mikrofonanordnung zur Aufnahme von HOA bestehen kann. Beim Aufbau eines sphärischen HOA-Arrays wird daher auf eine möglichst hohe Anzahl kleinerer Mikrofone geachtet, welche beispielsweise in der Schale einer schallharten Kugel angebracht werden. In einer auf der FOA-Matrix der Encodierung von A- nach B-Format aufbauenden mathematischen Funktion werden die einzelnen Ambisonics-Signale aus ihren virtuellen Richtcharakteristiken heraus berechnet. Auch für sphärischen Arrays gibt es einige wenige kommerzielle Lösungen, beispielsweise von Zylia oder mhAcustics.

7 Diskussion

7.1 Einfluss psychoakustischer Wahrnehmung auf Mikrofonanordnungen

Unsere beiden Ohren gelten stets als Referenz zum Aufbau immersiver Aufnahme- oder Wiedergabemöglichkeiten, so zeigt es die Entwicklung dieser anhand psychoakustischer Prinzipien. Nach Görne (2015) sei unsere räumliche Wahrnehmung binaural, was sich also auch auf eine Abhörsituation einer immersiven Wiedergabe bezieht. Wichtig anzumerken ist dabei die Subjektivität der immersiven Wahrnehmung - gerade auch bei unterschiedlichen Abhörpositionen im Raum -, welche sich aus interauralen Laufzeit- und Pegeldifferenzen ergibt. Generell gilt nochmals zu erwähnen, dass eine Richtungswahrnehmung in horizontaler Ebene auf zeitlichen und spektralen Unterschieden zwischen den Ohrsignalen beruht, wohingegen diese bei Schalleinfall in der Medianebene beinahe ausschließlich auf gelernten Klanfärbungsunterschieden aufbaut und wir so über eine deutlich geringere vertikale Lokalisationsschärfe von Schallquellen verfügen. Aus diesen Erkenntnissen heraus wurde auch eine Lautsprecherwiedergabe in der Elevation konzipiert, wonach bei einer Anordnung von 4+5+0 nach ITU die Horizontale

einer höheren Auflösung verfügt als die Vertikale. In einigen Anordnungen von Mikrofonen in immersiven Arrays zeigen sich ähnliche Ansätze mit dem Ziel einer kompakteren Positionierung, in denen das Middle und Upper Layer vertikal auf einer Ebene angebracht werden. Nach Lee (2021) hat dabei ein vertikaler Abstand zwischen Middle und Upper Layer keinen signifikanten Einfluss auf die räumliche Wahrnehmung in der Wiedergabe der Aufnahmen des Arrays. Doch genauso wichtig zum Wahrnehmungseindrucks des „Eingehülltseins“ ist die Berücksichtigung der interauralen Kohärenz, welche die Ähnlichkeit beider Ohrsignale beschreibt. Sie hängt in der immersiven Wiedergabe somit stark von der Ähnlichkeit der Lautsprechersignale zueinander ab, was äquivalent auf die Aufnahmekanäle einer 3D Mikrofonanordnung zu übertragen ist. Mikrofon Arrays zur Aufnahme diffuser Atmosphären verfügen daher zumeist - für eine maximale „Einhüllungswahrnehmung“ - über stark inkohärente Signale, welche durch eine horizontale Beabstandung und die Wahl stärkerer Richtcharakteristiken mit seitlicher Ausrichtung erreicht werden. Nach diesem Prinzip erzeugt eine Wiedergabe einer Ambisonics-Aufnahme typischerweise eine weniger räumliche Wahrnehmung, bringt allerdings durch ihre beinahe physikalisch korrekte Abbildung des aufgenommenen Schallfeldes eine deutlich geringere Klangverfärbung mit sich. Ein Vergleich zwischen Mikrofonanordnungen mit diskreter Verbindung zu typischen immersiven Wiedergabeformaten (z. B. 4+5+0 oder 9+10+3) und Anordnungen der Ambisonics-Domäne gilt allerdings prinzipiell als schwierig, da eine optimale Decodierung einer symmetrischen Lautsprecheranordnung bedarf (Lee, 2021).

7.2 Gegenüberstellung: Einzel- und Hauptmikrofonverfahren

Auch in der immersiven Wiedergabe kann sowohl die Aufnahme einer Hauptmikrofonierung aus einer 3D Mikrofonanordnung, als auch eine immersive Mischung einzelner Mikrofonensignale abgespielt werden, wonach auch Einzel- und Hauptmikrofonierung für jeweilige Anwendungen konkurrieren. Einen Sonderfall stellt dabei eine Beschallung in Wellenfeldsynthese dar, bei der trockene Einzelmikrofonensignale mit einer Raumimpulsantwort gefaltet und einzelne Lautsprechersignale berechnet werden (Dickreiter et al., 2008). Typische Anwendungen einer Hauptmikrofonierung stellen hingegen Musikaufnahmen „on location“ dar, beispielsweise Orchesteraufnahmen in einem Konzertsaal oder einer Kirche, oder Aufnahmen von Atmosphären. Kommt dabei

zum Beispiel ein Chor dazu, gelten Stützmikrofone in der Regel aufgrund der Textverständlichkeit als unverzichtbar (Dickreiter et al., 2008): Ein erster Schritt in Richtung Einzelmikrofonverfahren. Findet eine Aufnahme allerdings nicht in einem Aufnahmeraum besonderer Akustik, sondern in einem „trockenen“ Studio statt, so gehen bereits viele Qualitäten einer Hauptmikrofonierung, welche sich durch diffuse Reflexionen an Wänden und Decke ergeben, verloren. In diesem Fall ist vorzugsweise auf eine strikte Einzelmikrofonierung einzelner (musikalischer) Schallquellen zu setzen, welche in einer anschließenden Mischung im Raum verteilt werden können.

7.3 Gegenüberstellung: Portabilität und Räumlichkeitswahrnehmung

Wie vorher erläutert, hängt die Wahrnehmung von Räumlichkeit bei der Wiedergabe von Aufnahmen aus 3D Mikrofonanordnungen stark von Abstand, Ausrichtung und somit der Inkohärenz der einzelnen Kanäle ab. Doch sind große Dimensionen der Mikrofonarrays andererseits auch eine große Bürde in Bezug auf Portabilität. Die portabelsten Lösungen immersiver Aufnahmemöglichkeiten stellen (meist kommerzielle) Anordnungen im Sinne von Ambisonics erster oder höherer Ordnung dar, welche eine Audioszene so realistisch wie möglich reproduzieren (Lee, 2021). Doch führt genau diese koinzidente Anordnung der Mikrofone nicht zu Laufzeit-, sondern beinahe ausschließlich zu Pegeldifferenzen und so mit einer höheren Korrelation der Kanäle zu einer geringeren „Einhüllungswahrnehmung“, sowohl horizontal als auch vertikal. Horizontal und/oder vertikal beabstandete Mikrofonanordnungen werden daher auch als wahrnehmungsbezogene Arrays bezeichnet (Lee, 2021). Diese verfügen zwar über eine schlechtere Portabilität, lassen sich allerdings deutlich flexibler konfigurieren und so auf verschiedene Anwendungsbereiche besser anpassen. Doch in der Praxis wiegt der Vorteil eines unaufwendigen Aufbaus häufig höher, sodass Ambisonics-Systeme in vielen Aufnahmebereichen ihren Einsatz finden, obwohl deren Eigenschaften nicht immer das Optimum darstellen.

8 Fazit

Im Rahmen dieser Thesis wurde das Thema einer Übersicht von 3D Mikrofonanordnungen nach Anwendungsbereichen, Eigenschaften und Wiedergabesystemen unter Berücksichtigung tontechnischer Grundlagen bearbeitet, was einen möglichst vollständigen Einblick in die Hintergründe dieser ermöglicht. Aus den Erkenntnissen einer Kategorisierung der Arrays bezüglich ihrer Einsatzbereiche, Eigenschaften und Wiedergabe wurde im Rahmen dieser Arbeit eine übersichtliche Tabelle erstellt, welche dem Leser Aufschluss über für die jeweilige Anwendung sinnvoller Mikrofonanordnungen gibt.

Eine Anordnung von Mikrofonen zur stereofonen oder immersiven Aufnahme baut prinzipiell auf der Wahl aus verschiedenen Richtcharakteristiken für einzelne Aufnahmekanäle auf, welche durch verschiedene Empfängerprinzipien beeinflusst werden. Die Erkenntnisse aus der psychoakustischen Wahrnehmung des Richtungs- und Entfernungshören, sowie der Darstellung von Phantomschallquellen gelten als Grundsätze zur Konzipierung stereofoner und immersiver Aufnahmeverfahren, welche auf Basis von Zeit- und/oder Pegeldifferenzen ausgelegt sind. Immersive Wiedergabeformate über Lautsprecherkanäle sind innerhalb der ITU-Normen ITU-R BS.775-1 und ITU-R BS.2051-2 zusammengefasst und bilden so die Basis der meisten Lautsprecheranordnungen von 5.1 bis 22.2 (oder mehr), abgesehen weniger alternativen Konzept wie der Wellenfeldsynthese und einer Reflexions- oder Transauralbasierten Soundbar. Neben der Lautsprecherwiedergabe objektbasierten Inhalts, gilt auch das Thema der Decodierung aus der Ambisonics-Domäne durch einfache bis aufwendige Matrixoperationen, abhängig der Ordnung, als wichtige Grundlage zur Wiedergabe von Aufnahmen entsprechender Arrays. Andere 3D-Mikrofonanordnungen basieren hingegen auf einer diskreten Verbindung zwischen den Kanälen der Aufnahme und Wiedergabe über Formate wie 4+5+0 oder 9+10+3, was unter Berücksichtigung des Formates oder der Ordnung eine Einordnung in lautsprecherbezogen und Wiedergabeunabhängig erlaubt. Die Einordnung des Aufbaus in horizontal und vertikal beabstandete und/oder koinzidente Systeme gab Aufschluss über die Dimensionen einzelner Arrays und folgend über ihrer Portabilitätseigenschaften. Die vertikale Ausrichtung und Richtcharakteristik der Aufnahmekanäle des Upper Layers lieferte wichtige Eigenschaften des kanalübergreifenden Übersprechen und Korrelation,

welche die Wahrnehmungskomponenten der „Einhüllung“, der Lokalisationsschärfe vertikaler Schallquellen und der Klangverfärbung beeinflussen. Aus den gewonnenen Eigenschaften und aus Empfehlungen der Entwickler einzelner Arrays heraus, wurden sinnvolle und bedingt sinnvolle Anwendungsbereiche einzelner Anordnungen erörtert.

Abschließend gilt es anzumerken, dass im Themengebiet der Aufnahme und Wiedergabe in 3D-Audio noch weiterer Forschungsbedarf besteht, z. B. in Bezug der Aufnahme eines Bottom Layers. Zwar können aktuell verbreitete Formate wie 4+5+0 oder 9+10+3 durch immersive Aufnahmen von Musik und Atmosphären abgedeckt werden, doch es fehlt bislang an weiter Verbreitung genannter Hauptmikrofontechniken in der Audio-Community.

Literaturverzeichnis

Bowles, D. (2016, 29. November). *A microphone array for recording music in surround-sound*

with height channels. Academia. Abgerufen am 1. Dezember 2021, von

https://www.academia.edu/30160204/A_microphone_array_for_recording_music_in_surround_sound_with_height_channels

Camerer, F. (2020, Oktober). *Designing a 9-channel location sound microphone from scratch*.

Cornell University. <https://arxiv.org/pdf/2010.05877.pdf>

Dickreiter, M. (2011). *Mikrofonaufnahme: Aufnahmeräume, Instrumente, Mikrofone, Stereo- &*

Surroundaufnahmen (4., neu bearb. u. erw. Aufl. 2011 Aufl.). S. Hirzel Verlag.

Dickreiter, M., Dittel, V., Hoeg, W. & Wöhr, M. (2008). *Handbuch der Tonstudioteknik* (7.

Aufl., Bd. 1). K. G. Saur Verlag.

Eskow, G. (2016, 11. Februar). *Recording the Orchestra in 9.1*. MIX: Music Production, Live

Sound, Sound for Picture. Abgerufen am 1. Dezember 2021, von

<https://www.mixonline.com/live-sound/recording-orchestra-91-426704>

An Exceptionally Successful 2016 For Unamas And Mick Sawaguchi. (2017, Januar).

Merging: Audio for the Networking Age. [https://www.merging.com/news/press-](https://www.merging.com/news/press-releases/an-exceptionally-successful-2016-for-unamas-and-mick-sawaguchi)

[releases/an-exceptionally-successful-2016-for-unamas-and-mick-sawaguchi](https://www.merging.com/news/press-releases/an-exceptionally-successful-2016-for-unamas-and-mick-sawaguchi)

Görne, T. (2014). *Tontechnik: Hören, Schallwandler, Impulsantwort und Faltung, digitale*

Signale, Mehrkanaltechnik, tontechnische Praxis: Hören, Schallwandler, tontechnische

Praxis. (4., aktualisierte Aufl.). Carl Hanser Verlag.

Howie, W., King, R., Martin, D. & Grond, F. (2017, Mai). *Subjective Evaluation of Orchestral*

Music Recording Techniques for Three-Dimensional Audio. Will Howie.

Abgerufen am 1. Dezember 2021, von [http://www.willhowie.com/wp-](http://www.willhowie.com/wp-content/uploads/2017/05/WH.AES_142FullPaper.FINAL_.pdf)

[content/uploads/2017/05/WH.AES_142FullPaper.FINAL_.pdf](http://www.willhowie.com/wp-content/uploads/2017/05/WH.AES_142FullPaper.FINAL_.pdf)

Kuttruff, H. (2004). *Akustik: Eine Einführung* (1. Aufl.). S. Hirzel Verlag GmbH.

Lee, H. (2021, 19. Februar). *Multichannel 3D Microphone Arrays: A Review*.

Audio Engineering Society. Abgerufen am 1. Dezember 2021, von

<https://www.aes.org/e-lib/browse.cfm?elib=21013>

Lindberg, M. (o. D.). *3D Recording with the "2L-cube"*. 2L: The nordic sound.

Abgerufen am 1. Dezember 2021, von <http://www.2l.no/artikler/2L-VDT.pdf>

Recommendation ITU-R BS.775-1. (1992–1994). [Multichannel stereophonic sound system

with and without accompanying picture]. Internationale Fernmeldeunion.

https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.775-1-199407-S!!PDF-E.pdf

- Recommendation ITU-R BS.2051-2*. (2018). [Advanced sound system for programme production]. Internationale Fernmeldeunion. https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.2051-2-201807-1!!PDF-E.pdf
- Theile, G. (2001, Juni). *Natural 5.1 Music Recording Based on Psychoacoustic Principles*. Hauptmikrofon.de. Abgerufen am 1. Dezember 2021, von https://hauptmikrofon.de/theile/2001-1_Natural-5.1-Music-Recording_19th-AES-Conference-2001.pdf
- Theile, G. (2011). *Principles in Surround Recordings with Height*. Hauptmikrofon.de. Abgerufen am 1. Dezember 2021, von https://www.hauptmikrofon.de/doc/Auro3D_E_0311.pdf
- Weinzierl, S. (2008). *Handbuch der Audiotechnik*. Springer-Verlag Berlin Heidelberg.
- Wittek, H. & Theile, G. (2017, Oktober). *Development and Application of a Stereophonic Multichannel Recording Technique for 3D Audio and VR*. Schoeps. Abgerufen am 1. Dezember 2021, von https://schoeps.de/fileadmin/user_upload/user_upload/Wittek_AES_143_9869_002_.pdf

Zotter, F. & Frank, M. (2019). *Ambisonics: A Practical 3D Audio Theory for Recording, Studio*

Production, Sound Reinforcement, and Virtual Reality (1. Aufl.) [E-Book]. Springer

Open.