

**AUFNAHME- UND WIEDERGABEVERFAHREN FÜR
SURROUND SOUND**

DIPLOMARBEIT

von
Markus Nußbaumer

durchgeführt am
Institut für Elektronische Musik
der
Universität für Musik und darstellende Kunst Graz
im Rahmen des interuniversitären Diplomstudiums
Elektrotechnik - Toningenieur

Graz, im November 2000

Betreuer: Univ.-Ass. Dipl.-Ing. Alois Sontacchi
Begutachter: O.Univ.-Prof. Mag. Dipl.-Ing. Dr. Robert Höldrich

An dieser Stelle möchte ich mich bei all jenen bedanken, die mir bei der Erstellung dieser Arbeit behilflich waren.

Besonders bedanken möchte ich mich bei meinem Betreuer Herrn Univ.-Ass. Dipl.-Ing. Alois Sontacchi. Er gab während dieser gesamten Arbeit immer wieder wertvolle Anregungen sowie rasche Hilfestellung und stand stets als Ansprechperson für mich zur Verfügung. Vielen Dank dafür, Luis.

Herrn O.Univ.-Prof. Mag. Dipl.-Ing. Dr. Robert Höldrich danke ich für sein Engagement und sein Interesse an dieser Arbeit, wodurch für mich immer ein zusätzlicher Ansporn vorhanden war. Herzlichen Dank!

Herrn Ao.Univ.-Prof. Dipl.-Ing. Dr.tech. Gerhard Graber danke ich, dass er für die Zusammenstellung der Hörversuchsreihe und die Durchführung derselben das Tonstudio zur Verfügung stellte.

Bei der Fa. Schoeps GmbH. bedanke ich mich für das zur Verfügung stellen des Prototypen der OCT – Anordnung.

Dem Ensemble Tanguango danke ich, dass es sich für die Aufnahmen zur Verfügung gestellt hat, und wünsche den Musikern sowie dem Leiter des Ensembles, Klaus Johns, alles Gute für ihre weiteren Wege.

Allen Versuchspersonen, welche durch ihr Mitwirken die Durchführung der Versuchsreihe ermöglichten, sei ebenfalls ein Dank ausgesprochen.

Weiters möchte ich mich bei Roland Voraberger bedanken, der mir bei so manchen Schwierigkeiten mit der Hard- und Software im neu gestalteten Studio behilflich war, - Danke Roli.

Bei meinen Freunden und Kollegen bedanke ich mich, dass sie immer ein offenes Ohr für meine Probleme hatten bzw. mir Aufmunterung, Abwechslung und Zerstreuung in schwierigen Momenten waren. Danke ;-).

Mein größter Dank gilt aber meinen Eltern, die mir mit ihrem Vertrauen und ihrer Unterstützung immer beiseite standen.

ZUSAMMENFASSUNG

In dieser Diplomarbeit werden verschiedene Aufnahme- und Wiedergabeverfahren für Surround Sound behandelt. Neben einem Überblick über das räumliche Hören, Ambisonic und kommerzielle Übertragungsformate und Wiedergabekriterien liegt das Hauptaugenmerk auf der Konfiguration 3/2 (bzw. 5.1) und den dazugehörigen Mikrofonierungstechniken. Möglichkeiten der Konfiguration werden erläutert und Hinweise für die Praxis gegeben. Mit den schwerpunktmäßig behandelten Mikrofonierungstechniken Optimized Cardioid Triangle (OCT) nach Theile und Multichannel Microphone Array (MMA) nach Williams wurden Aufnahmen gemacht. Ausschnitte aus den Aufnahmen wurden einander in einem Hörversuch gegenübergestellt und auf verschiedene Parameter hin untersucht. Die Ergebnisse und die Auswertung des durchgeführten Hörversuchs werden dokumentiert.

ABSTRACT

This thesis deals with different recording- and reproduction-methods for surround sound. There is an overview concerning spatial hearing, ambisonics and commercially used formats for distribution and criteria for reproduction for multichannel material, whereas special attention is given to the 3/2 (respectively 5.1) – stereo - format and related microphone – techniques. Possibilities of the format are explained and instructions for the practical usage are given. Recordings have been made with the microphone – techniques OCT (optimized cardioid triangle, developed by Theile) and MMA (multichannel microphone array, developed by Williams), which are explained in detail. Parts of these recordings have been taken into a listening test, and the two techniques have been examined concerning various parameters. The results and the analysis of the listening test are documented.

INHALTSVERZEICHNIS

1	<i>Einleitung</i>	1
2	<i>Räumliches Hören</i>	2
2.1	Historisches - Ein Rückblick über das räumliche Hören	2
2.2	Umgebungskontext	3
2.2.1	Direkter und indirekter Schall	3
2.2.2	Koordinatensystem	3
2.3	Aufbau des Ohres	4
2.4	Wahrnehmung von Azimuth, Elevation und Entfernung	5
2.4.1	Lokalisation und Lokalisationsunschärfe	5
2.4.2	Entfernungshören im freien Schallfeld	6
2.4.3	Richtungshören in der Medianebene	7
2.4.4	Interaurale Zeitdifferenzen	10
2.4.5	Interaurale Pegeldifferenzen	11
2.4.6	Gesetz der ersten Wellenfront / Haaseffekt	12
2.4.7	Außenohrübertragungsfunktionen	13
2.4.8	Bewegungen von Kopf und Quelle	13
3	<i>Verschiedene Surround – Wiedergabekonfigurationen</i>	15
3.1	Allgemeines	15
3.2	Wiedergabekonfiguration für das Format 3/2 bzw. 5.1	16
3.2.1	Allgemeines	16
3.2.2	Möglichkeiten des Formats	16
3.2.3	Wiedergabeanordnung und Hörbedingungen	19
3.2.4	Einstellung der Wiedergabeanlage	23
3.2.5	Richtlinien für Aufzeichnung, Programmaustausch und Archivierung	24

3.2.6	Mikrofonaufnahmetechniken: Haupt- Stütz- und Raummikrofonie	26
3.2.7	Wiedergabe- und Übertragungsformate nach Dolby	63
3.3	Wiedergabekriterien nach THX (vgl. [77])	67
3.3.1	Digitale Tonspurformate bei THX	68
3.4	Wiedergabekonfiguration für Ambisonic	69
	Allgemeines	69
	Theorie von Ambisonic	69
3.4.1	Kompatibilität des Systems	75
3.4.2	Systemerweiterung und Speichermedium	78
3.4.3	Ambisonic Aufnahmen	79
4	Hörversuch	80
4.1	Allgemeines	80
4.2	Auswahl der Mikrofonverfahren	81
4.3	Aufnahmen	83
4.4	Erstellen der Versuchsreihe	84
4.5	Durchführung des Hörversuchs	85
4.5.1	Der Abhörraum	85
4.5.2	Versuchspersonen	86
4.5.3	Fragestellung und Versuchsablauf	87
4.5.4	Auswertung der Versuchsergebnisse	88
4.5.5	Auswertung der Daten zu den Versuchspersonen	94
4.6	Zusammenfassung	98
5	Literaturverzeichnis	99
6	Anhang A	105
7	Anhang B	109
7.1	Hörversuch zu Surround - Mikrofonierungs - Techniken	110
7.2	Zusatzinformation für die Versuchspersonen	114

7.3	Auflistung aufgezeichneter Signale und Ablauf der Versuchsreihe	115
7.4	Beispiel zur Analyse der statistischen Auswertung	117

1 EINLEITUNG

Der Wunsch des Menschen, Geschehnisse aus der Natur abzubilden, ist nicht neu und zeigt sich nicht nur in der Akustik. Statuen und Bildnisse von vergangenen Hochkulturen sind Hinweise dafür aus einer Zeit, in der von elektroakustischer Übertragung wohl noch nicht die Rede war. Wesentliche Informationen werden aber seit jeher durch Schall übertragen. Unsere akustische Umgebung trägt auch entscheidend zur Qualität von Erlebnissen bei, - seien sie alltäglicher oder besonderer Natur wie z.B. ein gutes Konzert. Die Möglichkeit zu haben, Erinnerungen an solche Erlebnisse zu unterstützen bzw. die Erlebnisse mehreren Menschen zugänglich zu machen, mag auch ein Grund für den Wunsch einer möglichst guten Ab- bzw. Nachbildung sein.

Waren vor der Ära der Schallaufzeichnung die Möglichkeiten der Aufzeichnung und Reproduktion z.B. von Musik nur durch Noten, Papier und weiteren Aufführungen gegeben, so stehen heute Techniken und Equipment zur Verfügung, welche eine Aufzeichnung und Wiedergabe von Schallereignissen mit sehr guter Qualität erlauben. Obwohl die Realisierung dieser beiden Möglichkeiten völlig unterschiedlich ist, ist doch das Prinzip ähnlich: Das Quellsignal kann mit Hilfe eines gemeinsamen Wertevorrates kodiert, gespeichert, dekodiert und schließlich wiedergegeben werden.

Die Entwicklung bis hin zu den heutigen Formen der Mehrkanal – Stereophonie verlief über viele Stationen. Zu Beginn der mechanischen und elektrischen Musikaufzeichnung stand die Monophonie, gefolgt von der Stereophonie in den 30iger Jahren. Damals beschäftigte man sich schon mit der Frage, wie akustische Eigenschaften eines Raumes dem Zuhörer vermittelt werden könnten um somit die Qualität der Ab- bzw. Nachbildung zu verbessern. Die Erweiterung auf mehr als auf zwei Wiedergabekanäle war eine Konsequenz aus dem Bedürfnis, akustische Situationen möglichst ‚naturgetreu‘ wiedergeben zu können. Das ursprüngliche Erlebnis kann zwar durch die Aufnahme und Wiedergabe nicht ersetzt werden, es ist aber jetzt mit der entsprechenden Aufnahmetechnik möglich, Verbesserungen bei der Darstellung von akustischen Situationen in Bezug auf die zwei – Kanal – Stereophonie zu erreichen.

Surround Sound hat sich deshalb in den letzten Jahren nicht nur im Kino, sondern auch im Konsumentenbereich (Fernsehen und reines Audio) immer stärker durchgesetzt und ist am Weg dazu, nach Stereo ein neuer Audiostandard der Zukunft zu werden.

2 RÄUMLICHES HÖREN

2.1 HISTORISCHES - EIN RÜCKBLICK ÜBER DAS RÄUMLICHE HÖREN

Schon seit der Schwelle zum 20. Jahrhundert werden Theorien zur Lokalisation von Schallquellen diskutiert. Das, obwohl es damals noch gar keine geeignete Technologie gab, um Testschalle für empirische Studien zu generieren und zu kontrollieren.

THOMPSON fasste diese Theorien 1882 in seinem Essay „On the Function of the Two Ears in the Perception of Space“ zusammen [1].

Drei Theorien wurden im Wesentlichen zu dieser Zeit vertreten:

Lokalisation aufgrund von:

- interauralen Intensitätsdifferenzen (interaural intensity differences, IID)
- IID sowohl als auch interauralen Zeitdifferenzen (interaural time differences, ITD)
- IID im Zusammenwirken der Pinnae zur Richtungsbestimmung

JOHN STRUTT, besser bekannt als LORD RAYLEIGH, formulierte 1907 die „Duplex-Theorie“, welche die Untersuchungen über die menschliche räumliche Wahrnehmung die nächsten 50 Jahre dominierte. Er stellte fest, dass es bei Schallereignissen, deren Wellenlänge klein im Verhältnis zu den Kopfabmessungen ist, zu einer Abschattung des Schallereignisses durch den Kopf kommt und damit beim ipsilateralen (dem Schallereignis zugewandten) Ohr zu einer höheren Schallintensität kommt als beim kontralateralen (dem Schallereignis abgewandten) Ohr. Das heißt, es tritt eine IID auf. Weiter bewirken verschiedene Abstände von der Schallquelle zu ipsi- bzw. kontralateralem Ohr eine ITD. STRUTT konnte zeigen, dass sich die ITD's vor allem bei tiefen Frequenzen auswirken. Er schloss daraus, dass die Lokalisation bei tiefen Frequenzen von ITD's und bei hohen Frequenzen von IID's bestimmt wird. Als Übergangsbereich gab er 1,5 kHz an.

Ab 1960 wurde verstärktes Augenmerk auf die Pinnae und deren Auswirkungen auf die Richtungswahrnehmungen gelegt. Publikationen von D. W. BATTEAU [2], J. H. HEBRANK

und D. WRIGHT [3] beschäftigten sich u.a. dahingehend und hatten ein Reihe von Untersuchungen zufolge. Es zeigte sich, dass die Filterung, welche durch die Pinna entsteht, besonders für die Lokalisation in der Medianebene und für die Vorne- / Hinten-Unterscheidung bedeutend ist, da hierbei nur minimale interaurale Differenzen auftreten.

2.2 UMGEBUNGSKONTEXT

2.2.1 DIREKTER UND INDIREKTER SCHALL

Breitet sich Schall in natürlicher Umgebung aus, so tun dies die Schallwellen nach allen Richtungen. Dabei treffen sie auf Raumobjekte, durch welche sie reflektiert oder gebeugt werden. Ein spezielles Raumobjekt ist der/die Zuhörende, dessen/deren Position die Schallquellen zu verschiedenen Zeiten aus verschiedenen Richtungen erreichen. Der Direktschall erreicht diese Position im Gegensatz zum indirekten Schall ohne Umwege. Der direkte Schall enthält die Information über die Richtung des Schallereignisses. Im indirekten Schall steckt die Information über die Umgebung und die relative Position des Schallereignisses, wie z.B. den Abstand von dem /der Hörenden.

2.2.2 KOORDINATENSYSTEM

Es ist zweckmäßig ein Koordinatensystem einzuführen, mit welchem eine möglichst anschauliche Darstellung gewährleistet ist. Im Fall des allgemeinen akustischen, räumlichen Hörens ist dies das Kugelkoordinatensystem mit dem Kopf als Mittelpunkt (siehe Abb. 1). Das Koordinatensystem sei invariant bezüglich der Kopfbewegungen.

Als Azimuth wird der Winkel φ in der Horizontalebene bezeichnet. Er kann Werte zwischen 0° und 360° annehmen, wobei 0° dem „Geradeaus - Schauen“ entspricht. Zum Beschreiben von Bewegungsabläufen in der Azimutebene erweist es sich manchmal sinnvoll, Links - Bewegungen als links $0^\circ - 180^\circ$, bzw. Rechts - Bewegungen als rechts $0^\circ - 180^\circ$ zu beschreiben. Die Elevation entspricht dem Winkel ϑ in der Medianebene. Er nimmt Werte von $0^\circ - \pm 180^\circ$ an.

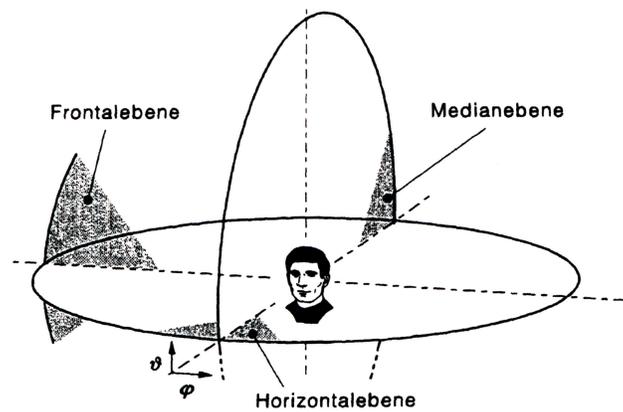


Abb. 1: Das kopfbezogene Koordinatensystem, vgl. [4]

2.3 AUFBAU DES OHRES

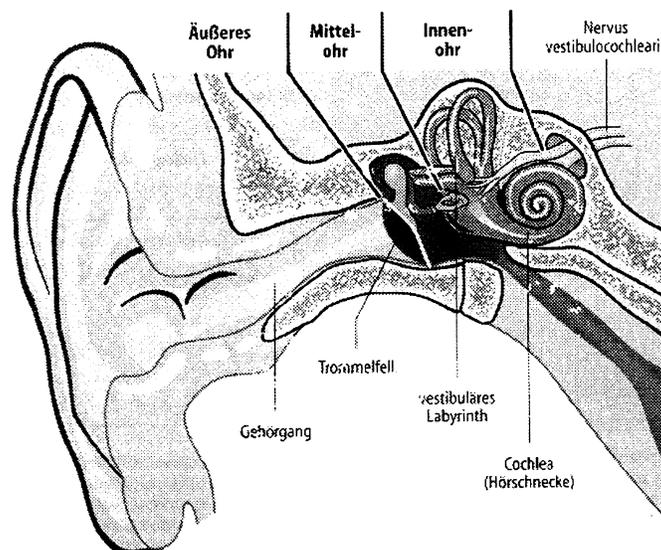


Abb. 2: Darstellung von Außen-, Mittel- und Innenohr vgl. [5]

Die naturgetreue Abbildung von Abb. 2 wird in Abb. 3 abstrahiert und der Weg des Schalls wird anhand einer eintreffenden Schallwelle nachvollzogen.

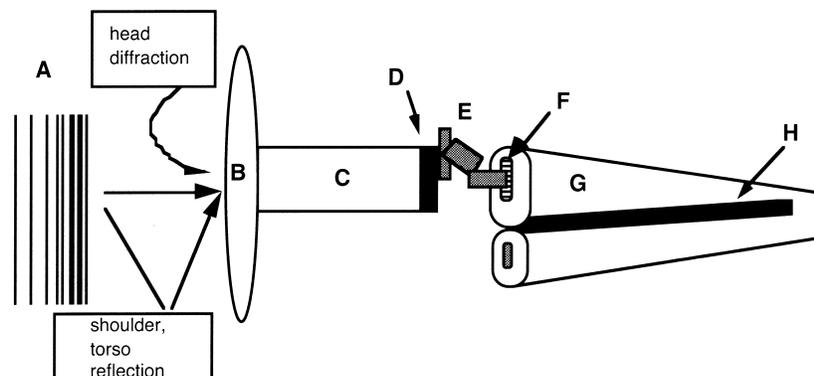


Abb. 3: Aufbau des Ohres, schematisch vgl. [6]

Der Schall (A) wird zuerst durch die Pinna (B), den sichtbaren Teil des Außenohrs, und die nahen Körperteile, wie Schultern und Kopf, beeinflusst. Danach folgen die Effekte des Ohrkanals (oder Meatus, C), der zum Mittelohr führt. Dieses besteht aus dem Trommelfell (D) und den Gehörknöchelchen (E). Die mechanische Energie wird durch Bewegung am ovalen Fenster (F) in Flüssigkeitsdruck im Innenohr (oder Cochlea, G) transformiert. Dieser Druck verursacht frequenzabhängige Erregungsmuster auf der Basilarmembran (H). Dadurch werden Haarzellen angeregt, die wiederum elektrische Aktionspotentiale in den Neuronen des Auditiven Systems auslösen. Diese werden auf höheren Niveaus mit den Informationen des anderen Ohrs kombiniert und verarbeitet.

2.4 WAHRNEHMUNG VON AZIMUTH, ELEVATION UND ENTFERNUNG

2.4.1 LOKALISATION UND LOKALISATIONSUNSCHÄRFE

Hierbei wird der Frage nachgegangen, wie gut das akustische Auflösungsvermögen des Menschen überhaupt ist.

BLAUERT definiert in [7] die Lokalisationsunschärfe als: „ ... die kleinste Änderung eines bestimmten Merkmals oder bestimmter Merkmale des Schallereignisses oder eines anderen, mit dem Hörereignis korrelierten Ereignisses, die gerade zu einer Ortsänderung des Hörereignisses (z.B. bezüglich Richtung und/oder Entfernung) führt.“

Die Lokalisationsunschärfe hängt von vielen Parametern ab. Abhängigkeiten bestehen vor allem von der Schalleinfallrichtung, der Art des Signals (z.B. Sprache, Rauschen), Frequenz und Dauer aber auch vom Alter der Versuchspersonen.

Besser als in der Medianebene wird in der Horizontalebene aufgelöst und hier wiederum zeichnet sich der Bereich um $\varphi = \vartheta = 0^\circ$ aus. Die größten Unschärfen treten über dem Kopf auf.

So liegt z.B. die Lokalisationsunschärfe bei Breitbandrauschen in der „Vorne“- Richtung bei $\pm 3.6^\circ$, vergrößert sich dann mit zunehmenden Winkel bis $\varphi = 90^\circ$ auf ca. $\pm 10^\circ$ und nimmt dann bis $\varphi = 180^\circ$ wieder ab auf etwa $\pm 5 - 6^\circ$.

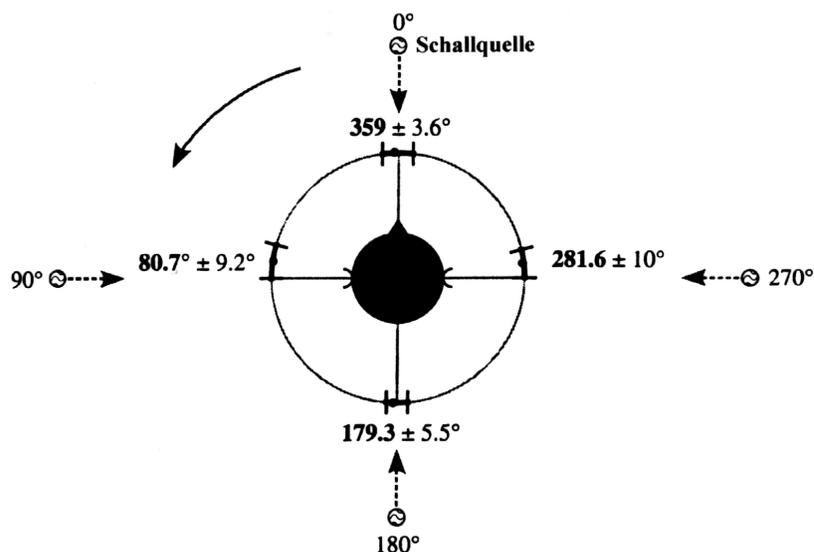


Abb. 4: Lokalisationsunschärfe in der Horizontalebene vgl. [8]

2.4.2 ENTFERNUNGSHÖREN IM FREIEN SCHALLFELD

Abhängig vom Abstand der Schallquelle sind verschiedene Merkmale in den Ohrsignalen dafür verantwortlich, wie das Gehör die Entfernung des Hörereignisortes beurteilt.

- 1) Bei mittleren Schallquellenentfernungen (Punktquelle, zeitlich stationäres Signal, ca. 3 – 15m) hängt nur der Schalldruckpegel¹ der Ohrsignale von der Schallquellenentfernung ab. Dieser sinkt im freien Schallfeld um 6dB pro Entfernungsverdopplung ($1/r$ – Gesetz).

¹ Mit Schalldruckpegel ist hier der breitbandig gemessene Effektivwertpegel gemeint.

- 2) Bei Schallquellenentfernungen größer als 15m muss zusätzlich zum $1/r$ – Gesetz die frequenzabhängige Dämpfung der Luft berücksichtigt werden (Dispersion: Hohe Frequenzen werden stärker gedämpft als niedrige). Außer dem Schalldruckpegel der Ohrsignale ändert sich damit auch die Form des Spektrums (Pegel und Phase).
- 3) Bei kleinen Schallquellenentfernungen ($< 3m$) treten spektrale Einfärbungen aufgrund der Krümmung der eintreffenden Wellenfronten am Kopf und der Pinnae auf. Diese sind anders geartet als jene, die aufgrund der Dispersion auftreten.

Wird Schall über Kopfhörer wiedergegeben, kann es zur ‚Im – Kopf - Lokalisation‘ kommen, da dabei durch das direkte Anbringen der Lautsprecher am Gehörgang die normale Filterwirkung der Ohrmuscheln ausgeschaltet wird. Dem kann aber z.B. durch die sogenannte Kunstkopftechnik entgegengewirkt werden. Das Signal muss dabei bereits bei der Aufnahme durch künstliche Ohren hindurchtreten und erfährt so eine Filterung entsprechend einer (allgemeinen) kopfbezogenen Übertragungsfunktion (HRTF – Head Related Transfer Function).

2.4.3 RICHTUNGSHÖREN IN DER MEDIANEBENE

Auch wenn nicht von exakt symmetrischen Verhältnissen für Kopf und Ohren ausgegangen werden kann, fehlen bei Schallereignissen in der Medianebene doch im Wesentlichen interaurale Merkmale, sodass vom Gehör v.a. monaurale Merkmale zur Bewertung herangezogen werden.

In [7] schildert BLAUERT eine 1967 gemachte, im Folgenden geschilderte Beobachtung: Er saß vor einem Lautsprecher, der einen Ton abstrahlte, dessen Frequenz sich gleitend von 200 Hz bis 16 kHz erhöhte. Dabei nahm er wahr, dass das Hörereignis sich mehrmals auf einer Bahn über seinen Kopf hinweg von vorn über oben nach hinten und zurück bewegte. Siehe dazu Abb. 5.

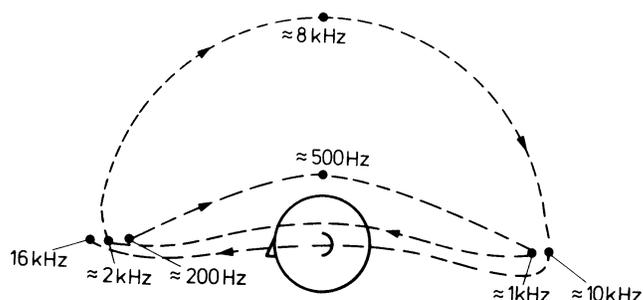


Abb. 5: Bahn der Hörereignisrichtung, vgl. [7]

Bahn der Hörereignisrichtung als Funktion der Bandmittenfrequenz bei Abstrahlung von Schmalbandrauschen aus einer beliebigen Richtung der Medianebene (1 Versuchsperson, Kopf fest, stark schematisiert).

Weitere Versuche mit terzbandbreitem Rauschen ließen „richtungsbestimmende Bänder“ [7] erkennen woraus zu schließen war, dass das Gehör die Hörereignisrichtung aufgrund der Frequenz der Ohrsignale² bildet.

Schallereignisse in der Medianebene werden demnach entsprechend ihres spektralen Anteils geortet. So ordnet man z.B. Schallsignalen mit einer Frequenz unter 500 Hz und um 3 kHz die Schallquellenrichtung „vorne“, jenen um 8 kHz „oben“ und Schallsignalen mit Frequenzen bei 1 kHz die Schallquellenrichtung „hinten“ zu.

² Als Ohrsignale werden die Schallsignale in den Gehörgängen bezeichnet. Sie sind laut BLAUERT [7] die wesentlichsten Eingangssignale einer (Versuchs-) Person bezüglich des räumlichen Hörens.

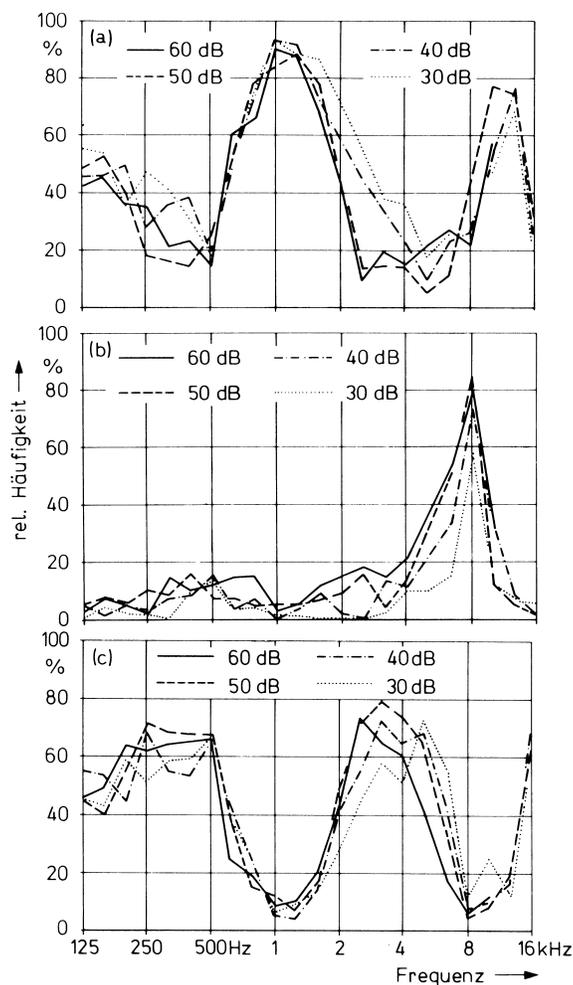


Abb. 6: Richtungsbestimmende Bänder, vgl. [7]

Relative Häufigkeiten der Antwort „h“ (hinten), 20 Versuchspersonen, bei jeder Terzmittenfrequenz jeweils einmal von vorn und einmal von hinten beschallt, b) wie a), jedoch Antwort „o“ (oben), c) wie a), jedoch Antwort „v“ (vorne).

Geht man weg von der Medianebene und über zum allgemeinen Fall, dann werden sich linkes und rechtes Ohrsignal unterscheiden und zwar in einer für Schalleinfallrichtung und Entfernung charakteristischen Weise. Es wird vermutet, dass das Gehör diese Unterschiede nicht in allen Details auswertet, sondern gut erkennbaren Merkmalen dieser Unterschiede Informationen entnimmt.

Dies sind zum Einen Unterschiede der beiden Ohrsignale bezüglich des Zeitpunktes ihres Auftretens (interaurale Zeitdifferenzen) und zum Anderen sind dies Unterschiede der beiden Ohrsignale bezüglich ihres Schalldruckpegels (interaurale Pegeldifferenzen).

2.4.4 INTERAURALE ZEITDIFFERENZEN

Schallwellen, die seitlich (z.B. $\varphi = 90^\circ$) auf den Kopf auftreffen, weisen am kontralateralen Ohr eine Laufzeit in Bezug auf das ipsilaterale Ohr auf.

Die meisten diesbezüglichen Versuche beschränkten sich auf das Herausfinden einer seitlichen Auslenkung (Lateralisation) des Hörereignisses aufgrund zweier, zueinander zeitverzögerter aber ansonsten gleicher, mittels Kopfhörer dargebotener Signale.

Eine typische Lateralisationskurve für kurze Knackimpulse (Klicks), die gleichermaßen für andere impulshaltige Signale, wie z.B. Rauschen oder Sprache, zutrifft, ist in der folgenden Abbildung dargestellt.

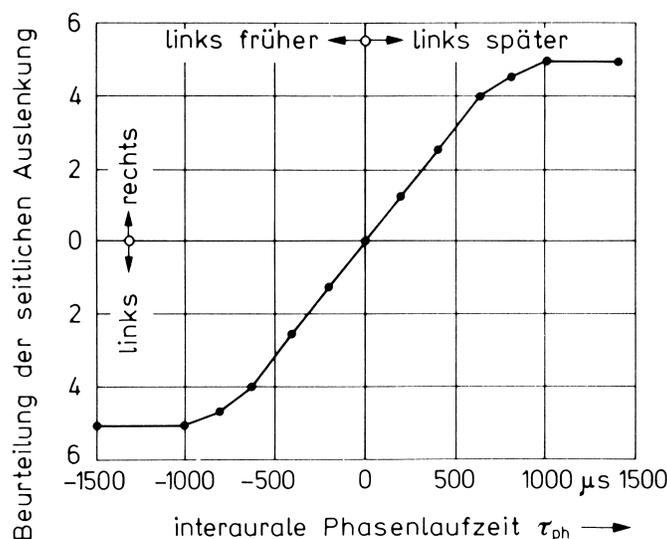


Abb. 7: Seitliche Auslenkung des Hörereignisses, vgl. [7]

Seitliche Auslenkung des Hörereignisses als Funktion einer verzerrungsfreien interauralen Signalverschiebung. Gültig für Impulse und impulshaltige Signale. 5 Versuchspersonen, Lautstärken von 30 bis 80 Phon.

Für Töne gilt eine ähnliche Abhängigkeit. Nimmt man an, dass das Gehör einen bestimmten Pegel - Schwellenwert zur Beurteilung dieser Zeitverzögerung heranzieht, folgt daraus, dass dieser zweimal während einer harmonischen Schwingung erreicht wird. Dominant ist aber jeweils das der Medianebene näher gelegene Hörereignis, also jenes mit der kürzeren interauralen Zeitdifferenz.

Maximale seitliche Auslenkung für Töne wird nur dann entsprechend obiger Abbildung bei einer interauralen Zeitdifferenz von $630\mu s$ erreicht, solange deren halbe Periodendauer $T/2$ $630\mu s$ nicht unterschreitet, d.h. ihre Frequenz etwa 800Hz nicht überschreitet. Bei höheren

Frequenzen wird die maximale Auslenkung immer kleiner und tritt nicht mehr bei $\tau_{ph} = 630\mu s$ sonder bei $\tau_{ph} = T/2$ auf.

Bei Frequenzen oberhalb von 1,5 bis 1,6kHz ist eine seitliche Auslenkung aufgrund interauraler Zeitdifferenzen in der Regel nicht mehr erkennbar.

2.4.5 INTERAURALE PEGELDIFFERENZEN

Ähnlich wie bei den interauralen Zeitdifferenzen werden hier jedoch Pegeldifferenzen zwischen kontra- und ipsilateralen Ohr aufgrund von Beugung und Abschattung behandelt.

Abhängig von der Wellenlänge eines Signals sind dessen Druckmaxima und -minima verschieden weit von einander entfernt.

Trifft eine Schallwelle auf ein ‚Hindernis‘ (z.B. den Kopf) welches klein im Vergleich zu seiner Wellenlänge ist, so werden um dieses Hindernis herum für einen festen Zeitpunkt der sich ausbreitenden Schallwelle kaum Druckunterschiede auftreten (Beugung).

Trifft die Schallwelle andererseits auf ein Hindernis, welches groß im Vergleich zu seiner Wellenlänge ist, so wird an der Vorderseite des Hindernisses Reflexion stattfinden, seitlich wird die Schallwelle ungehindert passieren und an der Hinterseite des Hindernisses wird ein für diese Frequenz schalltoter Bereich entstehen – die Druckmaxima einer ebenen Welle z.B. können sich nun auch seitlich ausbreiten und „laufen aus“, bzw. gleichen sich dem atmosphärischen Druck hinter dem Hindernis an (→ keine weitere Schallausbreitung hinter dem Hindernis – Abschattung).

Die Pegeldifferenzen der im freien Schallfeld auftretenden Ohrsignale sind stark frequenzabhängig. Diese Abhängigkeit wurde jedoch meistens bei den in der Literatur beschriebenen Experimenten nicht nachgebildet, - vielmehr beschränkte man sich auf die Untersuchung der seitlichen Auslenkung des Hörereignisses bei frequenzunabhängiger Pegeldifferenz. Siehe dazu Abb. 8. Dennoch ist diese Abhängigkeit bei der Verallgemeinerung der Ergebnisse zu bedenken.

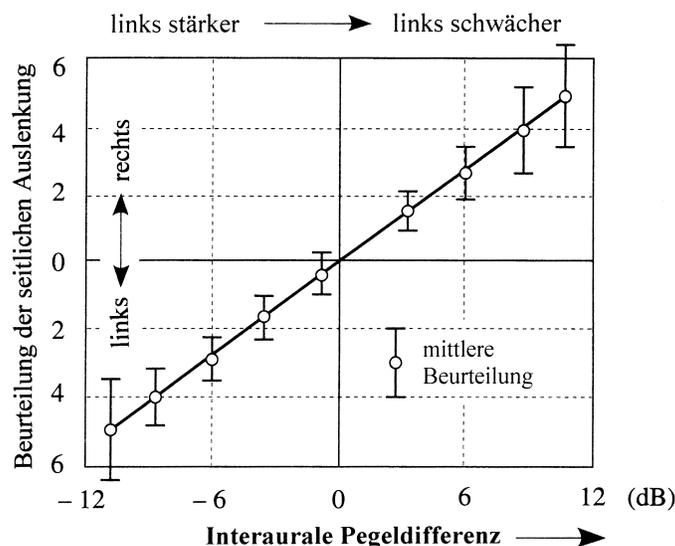


Abb. 8: Seitliche Auslenkung des Hörereignisses, vgl. [8]

Seitliche Auslenkung des Hörereignisses als Funktion der interauralen Pegeldifferenz.

2.4.6 GESETZ DER ERSTEN WELLENFRONT / HAASEFFEKT

Ähnlich den Verhältnissen bei der interauralen Zeitverzögerung wandert das Hörereignis bei der Darbietung zweier gleicher Signale zuerst in Richtung des nichtverzögerten Signals, wenn das zweite Signal verzögert wird. Tritt diese Zeitverzögerung in einem natürlichen Umfeld auf, spricht man vom Haaseffekt.

Dieser besagt nun, dass für die Ortung eines Schallereignisses die erste am Ohr eintreffende Wellenfront ausschlaggebend ist (Gesetz der ersten Wellenfront). Danach eintreffende Wellenfronten können ohne die Ortung zu beeinträchtigen z.B. um bis zu 10 dB lauter sein, wenn nur die Verzögerungszeit groß genug ist. Siehe Abb. 9.

Steigt die Verzögerungszeit t_v weiter an, kommt es zur Wahrnehmung von Echo. Die Verzögerungszeit, bei welcher dieser Übergang stattfindet, ist abhängig von der Signalart und dem Pegel des Primärsignals. Ab $t_v \sim 35\text{ms}$ kann bereits Echo auftreten.

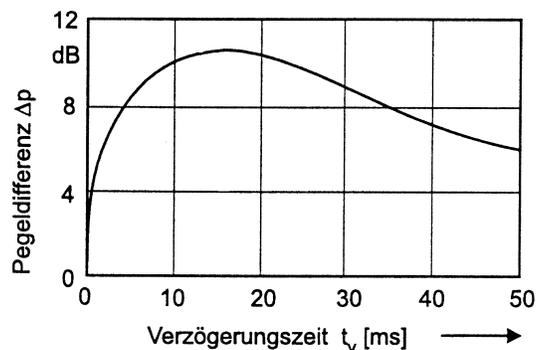


Abb. 9: Gesetz der ersten Wellenfront / Haaseffekt, vgl. [9]. Die erste am Ohr eintreffende Wellenfront trägt die Richtungsinformation der Schallquelle. Später eintreffende Schallwellen können bis zu 10 dB lauter sein, ohne die Richtungswahrnehmung zu ändern.

2.4.7 AUßENOHRÜBERTRAGUNGSFUNKTIONEN

Außenohrübertragungsfunktionen oder HRTF – Kurven (Head – Related – Transfer – Function), wie sie auch genannt werden, sind wesentlich für das räumliche Hören.

Ein Schallereignis welches an das Trommelfell gelangt, wird durch die Einflüsse von Außenohr, Gehörgang, Rumpf und Kopf, abhängig von seiner Einfallrichtung und den jeweils individuellen Außenohren, spektral gefärbt. Die HRTF beschreibt den Schalldruck im Freifeld im Verhältnis zum Druck am Trommelfell und ist vom Einfallswinkel und der Entfernung der Schallquelle abhängig.

Die binaurale HRTF kann man sich als frequenzabhängige Amplituden- und Zeitverzögerungsdifferenzen vorstellen. Kleine Unterschiede zwischen linker und rechter HRTF sind noch zusätzlich zu berücksichtigen, da die beiden Ohren nicht exakt gleich sind. Insgesamt führen also die asymmetrischen, komplex geformten Pinnae zu einem eindeutigen Set von kleinsten Zeitverzögerungen, Filter- und Beugungserscheinungen, das in eine eindeutige HRTF für jede Schallquellenposition übertragen wird.

2.4.8 BEWEGUNGEN VON KOPF UND QUELLE

Bewegungen des Kopfes sind wichtig, um Schalle genauer lokalisieren zu können. Die Drehtheorie besagt, dass dazu meist unbewusste Drehbewegungen, sogenannte Peilbewegungen ausgeführt werden. Ein Grund für diese Notwendigkeit liegt im ‚Cone of Confusion‘ (Abb. 10). Abstrahiert man den Kopf zu einer Kugel ohne Außenohren und betrachtet man nur die IID und ITD für einen Schallquellenort, so kann man diesen nicht eindeutig zuordnen.

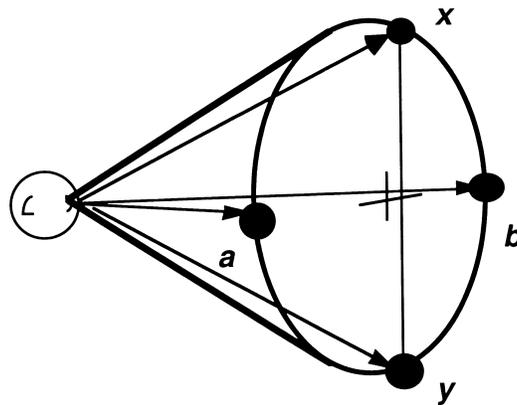


Abb. 10: ‚Cone of Confusion‘, vgl. [6]

Das Schallereignis am Ort a erzeugt dieselbe IID und ITD wie jenes am Ort b . Dies nennt man die Vorne – Hinten - Ambiguität. Ebenso verhält es sich mit den Orten x und y , an denen Schallquellen die Elevationsambiguität hervorrufen können.

Obwohl dieser Ansatz natürlich theoretisch ist, überführen doch die Peilbewegungen die vorher statische Hörsituation in eine dynamische und liefern damit dem Gehör einen größeren Informationsgehalt über das Schallfeld und machen damit das Orten einer Schallquelle einfacher. Zum Teil wird dabei der Kopf gedreht, bis das Schallereignis in den Bereich der größten Lokalisationsunschärfe (vorne) gelangt. Peilbewegungen werden beim natürlichen Hören sehr oft durchgeführt, vor allem auch dann, wenn die Schallquelle zuerst optisch noch nicht gesichtet werden kann.

Einen auch sehr wichtigen Effekt bei bewegten Schallquellen stellt der sogenannte Dopplereffekt dar. Darunter versteht man die Veränderung der Tonhöhe, wenn sich eine Schallquelle an dem/der ruhenden Hörenden vorbeibewegt.

3 VERSCHIEDENE SURROUND – WIEDERGABEKONFIGURATIONEN

3.1 ALLGEMEINES

Es ist bekannt, dass die herkömmliche 2 – Kanal – Stereophonie Grenzen bei der Abbildung von akustischen Ereignissen hat. Zusätzliche Stereokanäle sind wünschenswert um für eine vergrößerte Anzahl von Zuhörern eine verbesserte Darbietung zu erreichen. Es wird eine Ausdehnung des Referenzpunktes (sweet spot) angestrebt.

3.2 WIEDERGABEKONFIGURATION FÜR DAS FORMAT 3/2 BZW. 5.1

3.2.1 ALLGEMEINES

Im Folgenden sei das Hauptaugenmerk auf die Konstellation 3/2³ gerichtet. Die Anordnung 5.1⁴ ist vor allem für Bilddarbietung mit Tonwiedergabe gedacht und der dabei zusätzlich auftretende Tieftonlautsprecher soll wiederum nur für Spezialeffekte eingesetzt werden.

Das 3/2–Stereo Format wird den unter 3.1 genannten Anforderungen insofern gerecht, da durch den zusätzlichen Mittenkanal eine Verbreiterung der vorderen Stereo Basis - bei stabilerer Darstellung der Mitte für eine größere Hörfläche - erreicht wird.

Weitere Dienste leisten die zwei Surround Kanäle, - nicht, um eine Rundherum -, d.h. 360° - Abbildung zu erreichen, sondern um den linken und rechten Stereokanal zu ergänzen.

Zusätzlich will das genormte 3/2 Format dem Anspruch gerecht werden, ein bestmöglicher Kompromiss - für Bildanwendungen als auch für Tondarbietungen ohne Bild - zu sein.

3.2.2 MÖGLICHKEITEN DES FORMATS

Dieses Format gibt dem Tontechniker Verbesserungsmöglichkeiten (siehe auch Tabelle 1) im Vergleich zum herkömmlichen 2-Kanal Stereo. Hierbei ist es wichtig, sich die relevanten psychoakustischen Prinzipien vor Augen zu halten, um die gegebenen Mittel effizient einzusetzen. Das gilt speziell für jene Fälle, in welchen eine möglichst optimale, ‚naturgetreue Abbildung‘ gewünscht wird. Diesem Anspruch kann aber nur mit Einschränkungen Rechnung getragen werden.

Faktoren wie schlechte Aufnahmebedingungen im Konzertsaal, schlechte Hörbedingungen, reduzierte Dynamik, geforderte Abwärtskompatibilität etc. tragen dazu bei, dass selbstverständlich Abweichungen vom Original entstehen. Diese Abweichungen können andererseits auch als bewusst gewählte Gestaltungsmittel eingesetzt werden.

³ „3/2“ bezeichnet die genormte und in Abb. 11 dargestellte Lautsprecheraufstellung und gibt Auskunft über die Anzahl der Front - und Surround – Lautsprecher (3 Frontlautsprecher, 2 Surroundlautsprecher)

⁴ „1“ bedeutet, dass der Tieftonlautsprecher mit eingeschränkter Bandbreite (ca. 1/10 der natürlichen Hörbandbreite) betrieben wird.

Besteht der Wunsch nach der natürlichen stereofonen Abbildung, sollte diese jedoch den zwei folgenden Anforderungen gerecht werden: Sie sollte erstens ästhetisch gefallen und gleichzeitig sollte sie zweitens den tonalen und räumlichen Verhältnissen des Originals entsprechen. Dies wird in manchen Situationen zu Widersprüchlichkeiten führen und der resultierende Kompromiss, optimiert vom Tontechniker, wird umso besser gelingen, je flexibler die Aufnahmetechnik ist.

	<u>2/0-Stereo</u>	<u>3/2-Stereo</u>	<u>Kunstkopf</u>
Azimuth	+30°...-30°	+30°...-30°, Surround Effekte	Surround, vorne instabil
Elevation	nicht möglich	?	Möglich
Tiefe	simuliert	?	Möglich
Entfernung Lautsprecher - Kopf*	nicht möglich	?	Möglich
Raumeindruck	simuliert	möglich	Möglich
Umhüllende Quellen	nicht möglich	?	Möglich

Tabelle 1: Abbildungsmöglichkeiten stereofoner Systeme, vgl. [18]

Das 3/2-Stereo Format erweitert die Möglichkeiten der herkömmlichen zwei-Kanal Lautsprecher Stereophonie. Es ist jedoch ein Kompromiss und Beschränkungen sind durch Richtungs- und Entfernungsdarstellung gegeben, vergleicht man mit der Kunstkopftechnik. Andererseits ist es erwähnenswert, dass z.B. Kunstkopftechnik wenig Platz für kreatives Sound-Design offenlässt. „?“ – Einträge bezeichnen noch ungenügenden Informationsstand über die jeweiligen Parameter.

* Gemeint ist hiermit die Möglichkeit der Abbildung von Schallquellen in den Bereich zwischen Lautsprecher und Kopf.

Die Möglichkeiten, um jene das 3/2-Format erweitert wurde, liegen vor allem in den zusätzlichen Surround-Lautsprechern. Der zusätzliche Mittenkanal dient vor allem, wie bereits erwähnt, der Richtungsstabilität (Verbreiterung der Hörzone) durch Verwendung von zwei 30°-Stereobasen (L-C und C-R) anstelle von einer 60° Stereobasis L-R. Auch wird durch

ihn im Mittenbereich die Klarheit und Klangfarbe des Abbildes verbessert. Dies schlägt sich natürlich in der Mikrofonierungstechnik nieder. Siehe dazu 3.2.6.

Die drei prinzipiellen Anwendungen der Surround-Kanäle LS und RS liegen in der Generierung von

- Raum
- Atmosphäre und
- Effekten.

Dies jedoch immer in einer ergänzenden und/oder unterstützenden Art und Weise für das vordere stereofone Bild.

Das neue 3/2-Format kann also eine verbesserte stereofone Darstellung in Hinsicht auf Lokalisierung (speziell Wahrnehmung von Tiefe), räumlichen Eindruck, Einhüllende Atmosphäre (z.B. Applaus) bieten (siehe dazu auch Tabelle 2). Die entsprechenden psychoakustischen Prinzipien sollen als Phänomene des räumlichen Hörens verstanden werden und erfordern jeweils spezielle Mikrofonierungs- und Mischtechniken. Der bestimmten Handhabung von Delay, der Korrelation zwischen einzelnen Kanälen sowie der Pegelabstimmung von direktem und indirektem Schall kommen dabei eine besondere Bedeutung zu.

DIREKTER SCHALL
Lokalisierte Schallquellen → Richtungs- / Entfernungswahrnehmung
INDIREKTER SCHALL
Reflektionen / Hall → Raumeindruck
UMGEBUNGS- (NICHT REFLEKTIERTER) SCHALL
Ambiente, nicht lokalisierte Schallquellen → Einhüllende Atmosphäre

Tabelle 2: Drei Arten von 3/2-Stereofonie Schall, vgl. [18]

Lokalisation von (Phantom-) Quellen, Wahrnehmung von Raumeindruck und Wahrnehmung Einhüllender Atmosphäre (z.B. Applaus) sollten unterschieden werden und jeweils als Phänomene

des räumlichen Hörens verstanden werden. Diesen liegen spezifische Gesetze zugrunde, welche adäquate Mikrofonierungs- und Mischtechniken erfordern.

3.2.3 WIEDERGABEANORDNUNG UND HÖRBEDINGUNGEN

Die Abhöranlage sollte entsprechend den Empfehlungen des Surround Sound Forums (SSF) – aufgestellt und eingemessen werden. Diese Expertengruppe stellt eine interdisziplinäre und überregionale Arbeitsgemeinschaft dar, welche 1996 anlässlich der 19. Tonmeistertagung vom Verband Deutscher Tonmeister (VDT) gegründet wurde.

Die Empfehlungen sind vor allem für den Bezugsort in Studios gedacht, und sollen dort eine neutrale und kritische Kontrolle des Tonsignals ermöglichen. Die Heim - Abhöranlage sollte jedoch diesen Bedingungen möglichst nahe kommen, um die ursprüngliche Intention des Produzenten gut wiedergeben zu können. Aus [19] seien hier einige Punkte aufgegriffen und dargestellt.

Alle nachfolgenden Angaben sind als Mindestanforderungen zu verstehen, sie sind noch nicht hinreichend, um optimale Anforderungen genau zu beschreiben.

In Bezug auf **Wiedergabeanordnung und Aufstellungshinweise** finden sich recht detaillierte Angaben. Zur Referenzanordnung für das Format 3/2, welches als Weltstandard (auch für Bildwiedergabe) einen optimalen Kompromiss zwischen akustischen Anforderungen und den praktischen Realisierungsbedingungen für Studio und Heim darstellt, siehe Abb. 11.

Bei einer Abweichung von dieser Referenzaufstellung sind gegebenenfalls Laufzeit- und Pegelkorrekturen (für Wohn – Hörräume z.B. mit sogenannten Home - Prozessoren) durchzuführen.

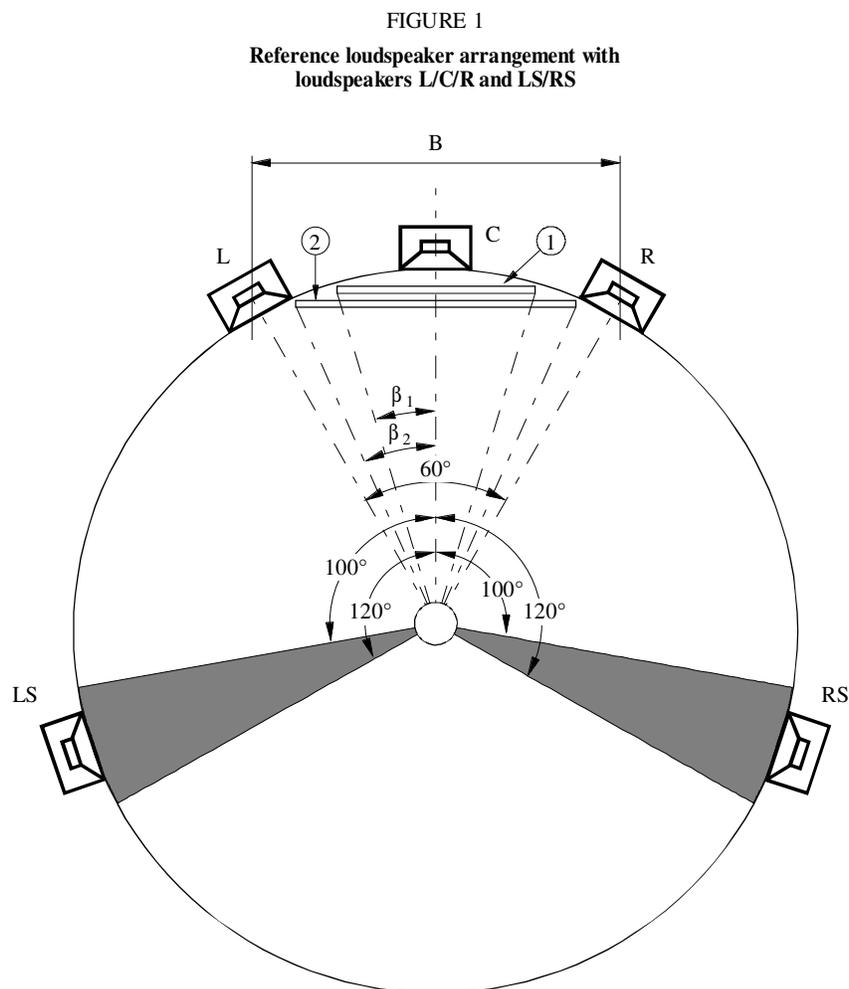
Neben diesem dem ITU-Standard entsprechenden Format 3/2 sind weitere Konstellationen (z.B. zur Vergrößerung der Hörfläche 5/2, 5/4,...) möglich und auch in Verwendung. Die zweite Ziffer bezieht sich jeweils auf die Anzahl der Surround – Kanäle.

Bei der Verwendung eines zusätzlichen Subwoofers (bis max. 120 Hz), welcher optional gemäß Standard zulässig ist, werden diese Produktions - Formate zu 5.1, bzw. 3/2/1 usw.

Bei der Produktion in all diesen Formaten ist jeweils auf Abwärtskompatibilität zu achten, was auch beinhaltet, dass z.B. keine programmrelevanten Signalanteile in den LFE (Low Frequency Enhancement bzw. Extension) – Kanal gelegt werden sollen.

Zur Verringerung der Volumen der 5 Stereo - Lautsprecher ist es bei Heim - Wiedergabeanlagen auch möglich die untere Abstrahl - Grenzfrequenz auf 80Hz – 100Hz hinaufzusetzen und den Bereich unter dieser Grenzfrequenz über (mehrere oder auch nur einen) Tieftonlautsprecher wiederzugeben. Diese Konfiguration wird weiterhin als 3/2 bezeichnet.

Die Tieftonlautsprecher sind nicht mit dem LFE – Kanal zu verwechseln, welcher ja schon vom Produzenten angeboten wird, - sie können aber zur Wiedergabe dieses Kanals (dazu sind verschiedene Einmeßvorgänge nötig) durch Umschalten verwendet werden.



Screen 1 HDTV – Reference distance = $3 H$ ($2\beta_1 = 33^\circ$)

Screen 2 = $2 H$ ($2\beta_2 = 48^\circ$)

H: height of screen

B: loudspeaker base width

Loudspeaker	Horizontal angle from centre (degrees)	Height (m)	Inclination (degrees)
C	0	1.2	0
L, R	30	1.2	0
LS, RS	100 ... 120	≥ 1.2	0 ... 15 down

DOI

Abb. 11: Referenz - Lautsprecheranordnung mit den Lautsprechern L/C/R und LS/RS in Kombination mit Bildwiedergabeeinrichtungen (gemäß ITU-R BS. 775-1 [20])

Die **Anforderungen an das Schallfeld** unterteilen sich in jene an den Direktschall und jene an den Reflektierschall, welcher sich wiederum in Erste Reflektionen (hier der Bereich von 1 kHz bis 8 kHz innerhalb der ersten 15 ms) und Diffusanteil gliedert.

Eine wesentliche Größe für diese Belange stellt die Nachhallzeit T dar. Sie ist definiert als jene Zeitspanne, innerhalb welcher die Schallenergie in einem Raum von ihrem Anfangswert auf ein Millionstel, d.h. auf -60 dB abgenommen hat.

Abb. 12 zeigt ein Toleranzschema, innerhalb welchem T_m^5 zu liegen hat.

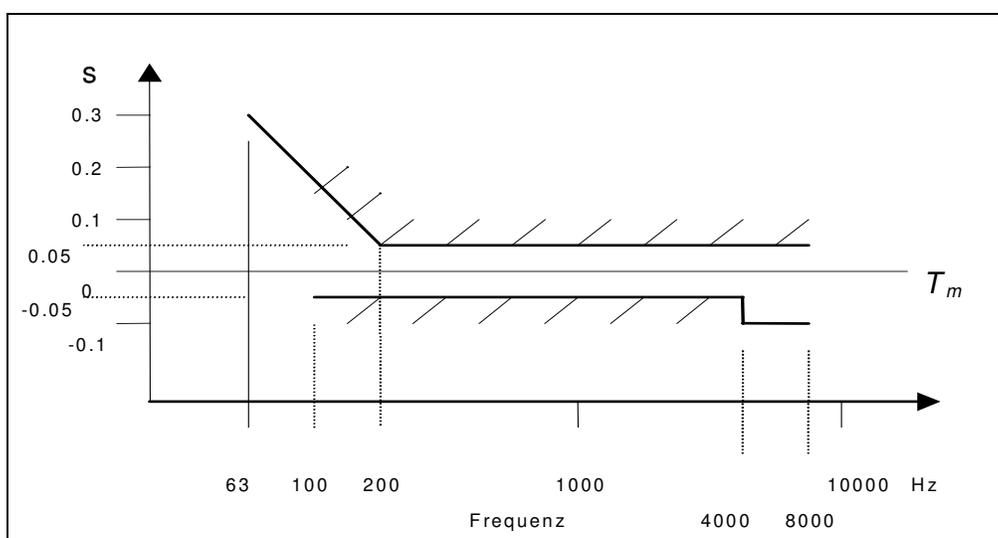


Abb. 12: Toleranzgrenzen für die Nachhallzeit, relativ zum arithmetischen Mittelwert T_m , [20]

Ein weiteres, wichtiges Kriterium ist die Betriebs - Schallpegelkurve. Rosa Rauschen terzbandgefiltert soll sich für jeden einzelnen Lautsprecher innerhalb der Toleranzgrenzen entsprechend Abb. 13 befinden.

⁵ T_m steht hier für den arithmetischen Mittelwert, welcher über die Terzbänder von 200Hz bis 2,5 kHz gebildet wird.

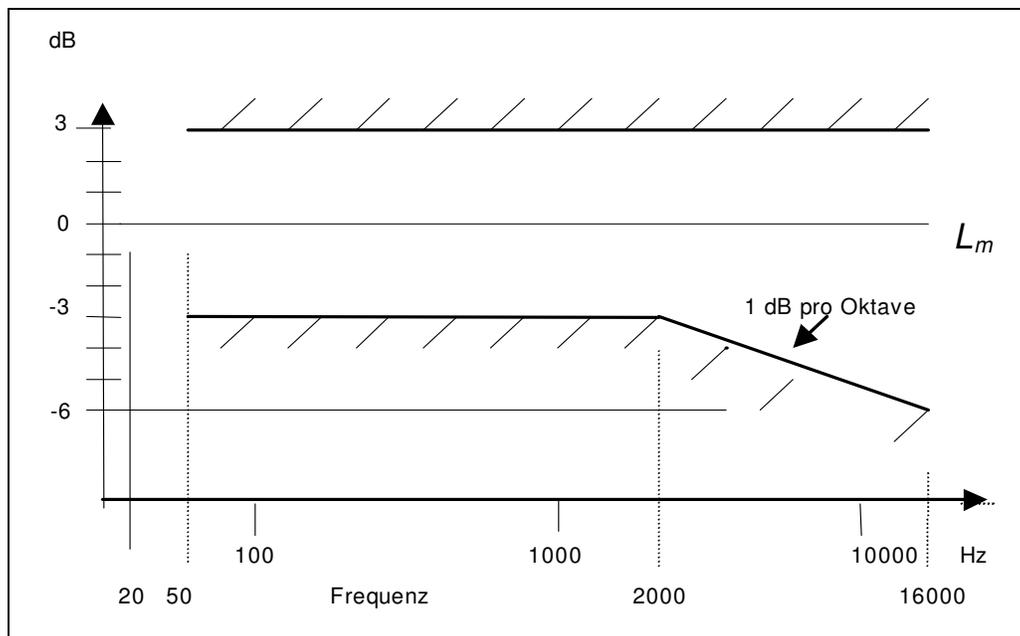


Abb. 13: Betriebs - Schallpegelkurve, relativer Pegel, Toleranzgrenzen [21]⁶

Für den **Referenz - Abhörpegel** soll am Bezugs - Hörpunkt mit allen Lautsprechern ein A - bewerteter Schalldruckpegel von 85dB erreicht werden. Dabei soll die Pegeldifferenz einzelner Lautsprecher kleiner 1 dB sein.

Zu den **Anforderungen für Referenz - Monitor - Lautsprecher** sei hier nur gesagt, dass aus Gründen der Kompatibilität fünf gleiche Lautsprecher empfohlen werden, obwohl, abhängig vom Programmmaterial, für die Surround – Lautsprecher manchmal diffus strahlende Monitore eventuell zu bevorzugen wären, diese aber als Frontlautsprecher wiederum nicht geeignet sind.

Für den **Referenz - Hörraum** wird ein Maximal - Volumen von 300 m³ und Symmetrie in Bezug auf die Hörrichtung gefordert.

⁶ Die Diagramme in Abb. 12 und Abb. 13 wurden der EBU-Empfehlung [21] entnommen. In entsprechenden Dokumenten der ITU [22] und von DIN [23] gibt es für Nachhallzeiten und Betriebsschallpegelkurven demgegenüber Abweichungen bzw. größere Toleranzen (siehe [19], Anhang 1). Für die vorliegende Empfehlung wurde den Angaben der EBU der Vorzug gegeben.

3.2.4 EINSTELLUNG DER WIEDERGABEANLAGE

Für die Einstellung einer Wiedergabeanlage hat das SSF eine Einmeß- und Testaufzeichnung vorgesehen, - die Test - DVD des SSF. Die Bedingungen dafür sind Bestandteile von [24].

Diese Einmeß- und Testaufzeichnung umfasst zwei Teile:

Abschnitt 1: Subjektive Kontrolle der Mehrkanal - Wiedergabe

Abschnitt 2: Meßtechnische Überprüfung der Wiedergabeanlage

Für Empfehlungen bezüglich Spurenbelegung sowie Aufzeichnungspegel siehe 3.2.5, Richtlinien für Aufzeichnung und Programmaustausch.

In [24, Anhang 1] sind Einzelheiten dazu und Hinweise zur Handhabung angegeben. Entnommen sei diesem Anhang der Teil 2, die

Subjektive Kontrolle der Wiedergabeanlage am Referenz – Hörort:

- a) Kanal - Identifikation: Prüfung der Spur- und Kanalzuordnungen durch zugeordnete Ansagen (für alle 8 Kanäle; auf der Test - DVD des SSF lediglich Kanal 1 bis 6)
- b) Lautstärkevergleich zwischen jeweils zwei Wiedergabekanälen durch Rauschimpulse
- c) Lautstärkevergleich zwischen Subwoofer und Stereo - Lautsprechern
- d) Prüfung der Phasenrichtigkeit (Gegenüberstellung phasenrichtig – verpolt – phasenrichtig) mittels Ansagen.

Die Beurteilungen auf Phasenrichtigkeit müssen jeweils vor der Mitte des jeweiligen Lautsprecherpaares (ggf. außerhalb des Referenzplatzes) erfolgen.

Nacheinander jeweils die Paare C + L, C + R, R + RS, L + LS.

- e) Überprüfung der vorgenommenen Einstellungen mit ausgewählten Hörbeispielen
- f) Digital „0“

Die Ansagen und Testsignale sind dabei lautheitsgerecht in Bezug auf Orchesteraufnahme als Referenz zugeordnet.

3.2.5 RICHTLINIEN FÜR AUFZEICHNUNG UND PROGRAMMAUSTAUSCH

3.2.5.1 Spurenbelegung⁷ für die Anwendung eines 8-Kanal-Aufzeichnungsformates

Bei der Festlegung der Aufzeichnungscharakteristika wurden die internationalen Empfehlungen zur Mehrkanal - Stereophonie zugrunde gelegt, insbesondere ITU-R BS.775-1 [20], ITU-R BS.1116-1 [22] sowie ITU-R BR.[DOC.10/10] [25]. Weiter wurden die Empfehlungen [26] bis [32] berücksichtigt. Siehe Tabelle 3, aus [85].

Spur	Signal	Belegung
1	L Links (Left)	
2	R Rechts (Right)	
3	C Mitte (Center)	
4	LFE Zusätzliches Tiefsttonsignal	Low Frequency Enhancement (Subbaß - und Effektsignal) für Subwoofer, optional ⁸
5	LS linkes Surround - Signal	- 3 dB im Fall Mono - Surround (MS = -3 dB)
6	RS rechtes Surround - Signal	- 3 dB im Fall Mono - Surround (MS = -3 dB)
7	Im Programmaustausch frei nutzbar ⁹	Vorzugsweise linkes Signal einer 2/0-Stereo-Abmischung.
8	Im Programmaustausch frei nutzbar	Vorzugsweise rechtes Signal einer 2/0-Stereo-Abmischung

Tabelle 3: Spurenbelegung bei Verwendung eines 8-Kanal-Aufzeichnungsformates

⁷ Der Begriff „Spuren“ wurde aus dem allgemeinen Sprachgebrauch übernommen - entsprechend engl. „tracks“ bei Magnetband -, auch wenn bei anderen Speichermedien keine realen Spuren mehr existieren.

⁸ Vorzugsweise angewendet bei Spielfilmen; für Heimwiedergabe optional, siehe Anmerkung unter 3.2.5.2, sofern kein LFE – Signal vorhanden ist, kann Spur 4 frei verwendet werden, z.B. für Kommentare, gemäß [25]. In einigen Regionen wird hier zusätzlich das monofone Surroundsignal MS = LS + RS vorgesehen, wobei der Pegel um jeweils 3 dB verringert wird.

⁹ Die Spuren 7 und 8 können alternativ, z.B. für Kommentare, für zusätzliche Surround - Signale oder für halblinke / halbrechte Front - Signale (z.B. für spezielle Filmformate) bzw. aus Matrixformaten gewonnene Summensignale L_t / R_t verwendet werden.

Diese Spurenbelegung ist für die Anwendung bei Hörfunk und Fernsehen verbindlich.

Es wird empfohlen, die im Einzelfall gewählte Nutzung der Spuren 4, 7 und 8 auf dem Aufzeichnungsträger anzugeben.

3.2.5.2 Aufzeichnungspegel

Es gelten die Einstellrichtlinien der EBU - Recommendations 64 [27] und 68 [28] sowie die Pegeldefinitionen in ITU-R BS. 645-2 [26], [33].

	Analoge Aufzeichnung	Digitale Aufzeichnung¹⁰ (bei Programmaustausch, live - Betrieb / Rundfunk / TV)	Digitale Aufzeichnung¹¹ (bei Mastering u.ä.)
Zugelassener Maximal - Signalpegel L_{PMS} (Permitted Maximum Signal Level)_bei 1 kHz Sinussignal	0 dB¹¹ relativ zum Studio - Normalpegel + 6dBu	- 9 dB¹¹ relativ zum digitalen Clip - Pegel 0 dB _{FS}	0 dB¹² relativ zum digitalen Clip - Pegel 0 dB _{FS}
Einstell - Signalpegel L_{AS} (Alignment Signal Level) bei 1 kHz Sinussignal	- 9 dB¹¹ relativ zum Studio - Normalpegel + 6dBu	- 18 dB^{11, 13} relativ zum digitalen Clip - Pegel 0 dB _{FS}	- 18 dB^{11, 12, 13} relativ zum digitalen Clip - Pegel 0 dB _{FS}

Tabelle 4: Verschiedene Aufzeichnungspegel, aus [85]

Anmerkung:

Bei der Wiedergabe von Kinofilmen wird der Schallpegel des LFE - Signals üblicherweise ca. 10 dB höher als der Wiedergabe - Schallpegel der übrigen Einzelkanäle eingepegelt (Näheres dazu in [19]). Zu den Messbedingungen siehe [24, Anhang 1]; weitere Erläuterungen siehe [24, Anhang 2].

¹⁰ bei digitaler Studio-Aufzeichnung lineare Auflösung, mindestens 16 bit und Abtastfrequenz 48 kHz, keine Preemphasis.

¹¹ gemessen mit einem IEC-Quasipeak-Programm-Aussteuerungsmesser ($\tau = 10$ ms). Bei Aussteuerungskontrolle mit $\tau < 0.1$ ms und Kurzzeitimpulsen wird dieser Maximal - Signalpegel überschritten.

¹² gemessen mit EBU-Spitzen-Programm-Aussteuerungsmesser ($\tau < 0.1$ ms) zur Ausnutzung der vollen Wortbreite bei CD-/DVD- u.ä. Produktionen.

¹³ in SMPTE - Regionen ist ein Einstellpegel = - 20 dB_{FS}, gemäß SMPTE RP 155, üblich.

3.2.5.3 Kontrollteil

Vor Beginn des Programmteils ist laut [25] und [27] ein Pegel-Kontrollteil mit 2 Kennsignalen für jede der benutzten Spuren mit Einstell-Signalpegel (Alignment Signal Level) L_{AS} aufzuzeichnen:

- a) Tonsignal mit 1-kHz-Sinuston zur Kontrolle des Einstell-Signalpegels;
- b) Rauschsignal, unkorreliert, zur Prüfung des Schalldruckpegels (siehe folgende Anmerkung sowie [24, Anhang 1]).

Anmerkung:

Die Aufzeichnung des Rauschsignals kann entfallen, sobald Einmeß- und Testaufzeichnungen (u.a. Test-DVD des SSF) gemäß [12, Anhang 1] einheitlich verwendet werden.

3.2.6 MIKROFONAUFNAHMETECHNIKEN: HAUPT- STÜTZ- UND RAUMMIKROFONIE

Das Hauptmikrofon und dessen charakteristische Attribute mögen unterschiedlich sein, wenn man herkömmliche Zwei – Kanal – Anwendungen oder Fünf – Kanal – Anwendungen betrachtet. Es sollte jedoch prinzipiell zwei (psychoakustischen) Aspekten Rechnung tragen:

- Richtungsabbildung: Aufnahme des ersten Schalls einer Quelle oder einer Gruppe von Quellen, welche bei der Wiedergabe die „erste Wellenfront“ (siehe 2.4.6 und [34]) bildet (direkter Schall).
- Räumliche Abbildung: Aufnahme der entsprechenden natürlichen Reflektionen und des Halls (indirekter Schall)

Im Falle der herkömmlichen Zwei – Kanal – Stereophonie erscheint die Umsetzung dieser beiden Aspekte mit einem Stereo (Haupt-) Mikrofon vorteilhaft. Vorausgesetzt natürlich, dass günstige Aufnahmebedingungen gegeben sind und die richtige Mikrofonposition gefunden wird, sowohl um eine adäquate Richtungsabbildung als auch ein entsprechendes Verhältnis von direktem und indirektem Schall zu gewährleisten (R/D – Ratio [51]).

Betrachtet man 3/2 – Stereophonie, dann kommen weitere Aspekte zum Tragen.

Erstens muss berücksichtigt werden, dass die Richtungsabbildung des vorderen Bereichs durch drei Lautsprecher, L C, R, realisiert wird, welche wiederum die Sub – Stereo – Basen [40] L-C und C-R bilden und deshalb einer passenden 3 – Kanal – Aufnahme des direkten Schalls bedürfen.

Zweitens muss Bedacht darauf genommen werden, dass, abhängig vom Aufnahmeraum ca. 50 % der Energie des indirekten Schalls aus den Surround - Kanälen LS und RS kommen soll, um annähernd der natürlichen Situation zu entsprechen. Das hat eine entsprechende Mikrofonaufstellung und auch adäquate Richtcharakteristika der Mikrofone zur Folge, um eine ausreichende Trennung von direktem und indirektem Schall zu gewährleisten.

Der Anspruch an das 3/2 - Hauptmikrofon ist jener, dass es alleine schon bei der Abbildung des vorderen Bereichs der herkömmlichen Zwei – Kanal – Stereophonie überlegen sein soll. Diesem Anspruch kann das 3/2 – Format zumindest in zwei Punkten gerecht werden: Zum einen ist das die Richtungsstabilität und die daraus resultierende, vergrößerte Hörfläche [40]. Zum anderen sind das bei der Schall – Wiedergabe – Qualität die Attribute ‚Klarheit‘ und ‚Klangfarbe‘, wie eine Reihe von Untersuchungen [54], [55] belegt.

Ein optimales 3 – Kanal – Hauptmikrofon für den vorderen Bereich muss sich also diese Vorteile zu eigen machen. Dabei gilt es vor allem ein ‚Kanal – Übersprechen‘ und damit das ‚Dreifach - Phantomquellen - Problem‘ (siehe dazu Abb. 14) bei der Aufnahme so gut als nur möglich zu vermeiden, denn dies hat wiederum ein Verwischen der Abbildungsschärfe und Kammfiltereffekte zur Folge.

Findet Übersprechen statt, generieren mehr oder weniger korrelierte Signale drei Phantomquellen deren Richtung und Ausdehnung von den Pegel– und Zeitdifferenzen abhängig sind, welche durch Mikrofonposition und Richtcharakteristik beeinflusst werden können. Idealerweise sollen diese Parameter so gewählt werden, dass eine Schallquelle links von der Mikrofonkapsel rechts nicht aufgenommen wird und umgekehrt. Schallquellen in der Mitte sollten nur vom mittleren Mikrofon aufgenommen werden.

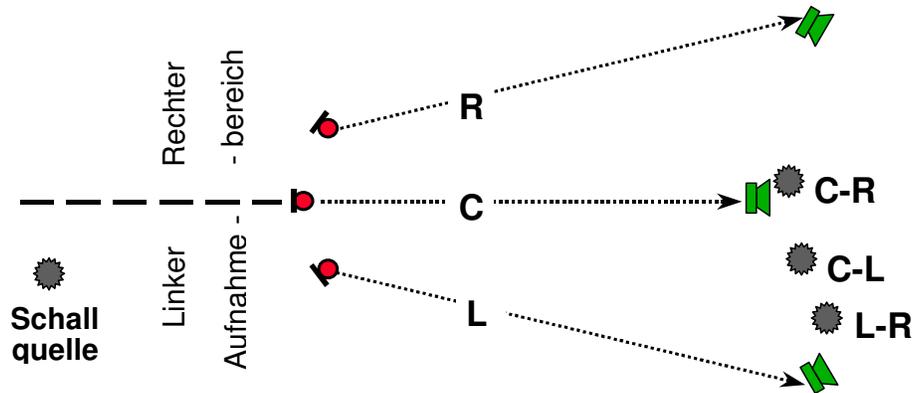


Abb. 14: Das 'Dreifach - Phantomquellen - Problem' beim 3 - Kanal Stereomikrofon, aus [18]

Prinzipiell produziert jede Stereobasis C-L, C-R und L-R seine eigene Phantomquelle und jede dieser Phantomquellen findet sich an verschiedenen Orten wieder. Das resultiert in verminderter Abbildungsschärfe, weniger Klarheit und in Klangfarbeneffekten.

3.2.6.1 L-C-R Hauptmikrofon nach Theile [18]

Diese Anordnung ist prinzipiell eine herkömmliche, auf engen Raum angeordnete, Drei – Kanal - Mikrofon - Konfiguration. Sie basiert nicht auf der Anwendung der Mikrofon – Array – Technologie (siehe dazu 3.2.6.2).

Die prinzipielle Anordnung ist in Abb. 15 dargestellt. Sie wurde bezüglich Kanal - Übersprechen und Aufnahmewinkel optimiert. Um dem Übersprechen beizukommen, wurden für L und R Mikrofone mit Supernieren – Richtcharakteristik, welche jeweils 90° nach außen gedreht sind, verwendet. Bei diesem Einsatzzweck werden die Supernieren natürlich ungünstig bezüglich ihres Frequenzganges angeordnet. Diesen Nachteil kann man aber dadurch kompensieren, dass man Mikrofone bzw. Kapseln mit kleinem Durchmesser verwendet und auch der Freifeld – Frequenzgang kann für 60° anstelle für 0° bezüglich der Mikrofonachse eingestellt werden. Das Mikrofon C weist Nierencharakteristik auf.

Nimmt man frontalen Schalleinfall ($\Omega = 0^\circ$) an, so ergeben sich für die drei Kanäle die folgenden Pegel- und Zeitdifferenzen (vgl. [18]):

Ω	L	C	R	Δt
0°	-9 dB	0 dB	-9 dB	0,24 ms

Die Zeitdifferenz Δt ergibt sich aus den geometrischen Abmessungen und errechnet sich aus (vgl. [18]):

$$\Delta t = \left[\sqrt{h^2 + (1/2b)^2} \cdot \cos \left(90^\circ - \Omega + \arctan \frac{2h}{b} \right) \right] \cdot 0,03 \frac{\text{ms}}{\text{cm}}$$

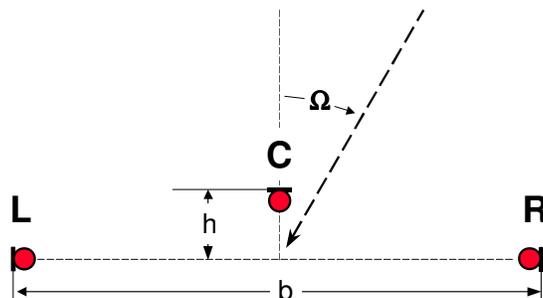


Abb. 15: Optimierte Dreieckskonfiguration / OCT (optimised cardioid triangle) nach Theile

L und R: Supernierencharakteristik, C: Nierencharakteristik. Der Freifeld – Frequenzgang sollte vorzugshalber für $\Omega = 30^\circ$ eingestellt werden. Der Abstand b ist abhängig vom gewünschten Aufnahmewinkel φ :

$\varphi = 90^\circ$	$b = 80 \text{ cm}$
$\varphi = 100^\circ$	$b = 70 \text{ cm}$
$\varphi = 110^\circ$	$b = 60 \text{ cm}$

Der Abstand h beträgt 8 cm, - werden Nieren anstelle von Supernieren verwendet, dann soll $h = 12 \text{ cm}$ betragen (vgl. [18]).

Die Größe b ist abhängig vom Aufnahmewinkel und beträgt (für diesen Fall für $\varphi = 100^\circ$) 70 cm. Δt wird durch den Abstand h bestimmt und resultiert aufgrund der hier verwendeten 8 cm in 0,24 ms. Für die Pegeldifferenz ΔL zwischen C und L bzw. R ergibt sich aufgrund der Richtcharakteristika ein Wert von 9 dB. Beide Werte lassen die jeweils entstehende Phantomschallquelle in die Mitte rücken, - ΔL um ungefähr 10° , Δt um ungefähr 5° .

In [40] und [53] wird gezeigt, dass eine in der Mitte von zwei Lautsprechern wahrgenommene Phantomschallquelle in gewissen Grenzen entsprechend den Gesetzmäßigkeiten von

Tabelle 5 verschoben werden kann. Wichtig zu bemerken ist, dass die Verschiebung der Phantomquelle für bestimmte Intensitäts- und Zeitdifferenzen einen konstanten Bezug zur Lautsprecherbasis hat. Tritt z.B. eine Intensitätsdifferenz von 6 dB zwischen zwei Lautsprecherkanälen auf, wird die Quelle um 45% verschoben. Dem entsprechen $\Delta\vartheta_i = 13.2^\circ$ in einer herkömmlichen 60° - Lautsprecheranordnung. Eine Verringerung des Winkels der Stereobasis von $\Psi = 60^\circ$ auf $\Psi = 30^\circ$ wird die Verschiebung der Phantomschallquelle entsprechend auf $\Delta\vartheta_i = 6,6^\circ$ halbieren. Dieses Phänomen widerspricht zwar den Theorien der Summenlokalisation, es wird jedoch durch das Assoziationsmodell nach Theile (vgl. [40] und [53]) erklärt.

Verschiebungsfaktoren aufgrund:	2-Kanal-Stereo Lautsprecher L-R Basiswinkel: $\Psi = 60^\circ$	3-Kanal-Stereo Lautsprecher L-C-R Basiswinkel jeweils: $\Psi = 30^\circ$
Intensität: $Z_i = 7,5\% / \text{dB}$	$\Delta \vartheta_i = 2,2^\circ / \text{dB}$	$\Delta \vartheta_i = 1,1^\circ / \text{dB}$
Zeit: $Z_t = 15\% / 0,1 \text{ ms}$	$\Delta \vartheta_t = 4,4^\circ / 0,1 \text{ ms}$	$\Delta \vartheta_t = 2,2^\circ / 0,1 \text{ ms}$
	$\Delta \vartheta_\Sigma = \Delta \vartheta_i + \Delta \vartheta_t$	

Definition:

$\Delta \vartheta_i, \Delta \vartheta_t$: Verschiebungswinkel der Phantomquelle aufgrund von Intensitäts und Zeitdifferenzen

Verschiebungsfaktor Z

Links Mitte Rechts

$Z_i = 7,5\% / \text{dB}$

$Z_t = 15\% / 0,1 \text{ ms}$

Tabelle 5: Verschiebungsfaktoren für Intensitäts- und Zeitdifferenzen, vgl [18]

Wird also eine Stereo – Wiedergabebasis in zwei Hälften geteilt, so ist das auch für den Aufnahmewinkel erforderlich und der Abstand zweier Mikrophonkapseln in einer engen, d.h. auf engem Raum angeordneten, Hauptmikrofonkonstellation muss entsprechend vergrößert werden.

Anders ausgedrückt heißt das: Verglichen mit einem engen Zwei – Kanal Hauptmikrofon muss der Abstand zwischen den Kapseln L und R ungefähr viermal größer sein wenn ein C – Mikrofon zwischen den beiden Mikrofonen L und R eingeführt wird (vorausgesetzt der gleiche Aufnahmewinkel und die gleiche Wahrnehmung der räumlichen Information wird gewünscht. Die vergrößerte Laufzeit zwischen den Kapseln L bzw. R und C wird den Aufnahmewinkel entsprechend Tabelle 5 verkleinern).

Unter diesen Voraussetzungen wurden in [18] die wahrgenommenen Winkel der Quellen sowie das Kanal – Übersprechen ermittelt und hier in Tabelle 6 dargestellt. Dazu wurden die Werte der Richtcharakteristika aus Tabelle 7 verwendet. Es wurde angenommen, dass sich die Quelle im rechten Sektor, entsprechend Abb. 15 befindet. In diesem Fall sollte vom linken Mikrofon kein Schall aufgenommen werden. Die Verhältnisse bezüglich Übersprechen sind für die Supernieren – Konfiguration akzeptabel. Werden jedoch für L und R Nieren anstelle von Supernieren verwendet, tritt das Übersprechen als Nachteil zu Tage, welcher nicht mehr vernachlässigbar ist. Betrachtet man die Translationskurve, so weisen beide Konfigurationen ähnliche Eigenschaften, vergleichbar mit jenen einer Zwei – Kanal Stereokonfiguration, auf.

Der Aufnahmewinkel beträgt ungefähr $\varphi = 100^\circ$.

Quellwinkel Ω	SUPERNIERE h = 8 cm / b = 70 cm		NIERE h = 12 cm / b = 70 cm	
	Übersprechen →L	Winkel ϑ wahrgenommen	Übersprechen →L	Winkel ϑ wahrgenommen
+ 90°	-11 dB	(30°)	<-18 dB	(30°)
+ 60°	<-18 dB	(30°)	<-18 dB	(30°)
+ 45°	<-18 dB	27°	-15 dB	26°
+ 30°	<-18 dB	19°	-11 dB	17°
+ 15°	-14 dB	9°	-8 dB	8°
0°	-9 dB	0°	-5 dB	0°

Tabelle 6: Abbildungscharakteristika von zwei optimierten Dreieck - Konfigurationen (Quelle im rechten Sektor), vgl. [18]

E.R	0°	15°	30°	45°	60°	75°	90°	105°	120°	135°	150°	180°
N	0	-0,3	-1	-2	-3	-4	-6	-8	-11	-15	<-18	<-18
S.N.	0	-0,3	-1	-2	-3,5	-5	-9	-14	<-18	<-18	<-18	-11
H.N.	0	-0,3	-1	-2	-4	-7	-12	<-18	<-18	-12	-9	-6

E.R...Einfallrichtung, **N...**Niere [dB], **S.N...**Superniere [dB], **H.N...**Hyperniere [dB]

Tabelle 7: Zur Ermittlung von Tabelle 6 herangezogene Richtcharakteristika und deren Werte unter verschiedenen Schalleinfallswinkeln, vgl. [18]

Eine weitere Optimierung bezieht sich auf das Bass – Verhalten der Hauptmikrofon – Anordnung. Es ist ein bekanntes Faktum, dass Mikrofone mit Nierencharakteristik (speziell Super – und Hypernierencharakteristik) eine Schwäche bei der Übertragung von tiefen Frequenzen haben. Ein Hauptmikrofon, welches auf Druckempfängern aufgebaut ist, ist in dieser Beziehung überlegen. Deshalb erscheint es sinnvoll, die Vorteile dieser beiden Charakteristika zu verbinden. Die zwei Mikrofone L und R werden hier im Bassbereich mit zwei Tiefpass - gefilterten Mikrofonen mit Kugelcharakteristik ergänzt. In dieser Anwendung (siehe Abb. 16) resultiert somit über der Trennfrequenz die Richtungsabbildung wie oben

beschrieben. Unter der Trennfrequenz liefert das C – Mikrofon keinen Beitrag und das hat zur Folge, dass eine gewöhnliche AB – Konfiguration mit 70 cm Abstand wirksam wird.

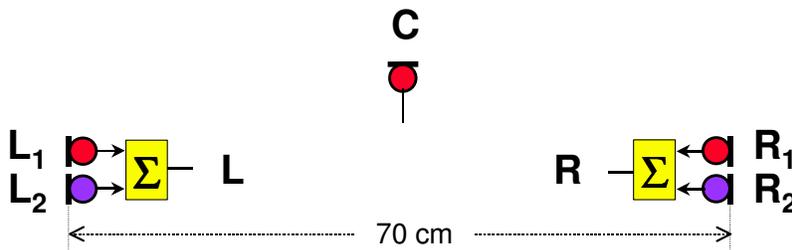


Abb. 16: Supernieren - Konfiguration mit verbessertem Baßverhalten, [18]

Die Kapseln L_1 und L_2 , sowie die Kapseln R_1 und R_2 bilden das ‚Zwei – Weg – Mikrofon‘ L und R.

Kapseln L_1 und R_1 : Supernieren, kleiner Durchmesser, 100 Hz Hochpass – gefiltert

Kapseln L_2 und R_2 : Kugeln, 100 Hz Tiefpass – gefiltert

Kapsel C: Niere, 100 Hz Hochpass – gefiltert

Erweiterung mittels Raummikrofon für fünf Wiedergabekanäle

Dazu schlägt Theile [18] vor, das L-C-R Hauptmikrofon und das Raummikrofon separat zu betrachten. Das Hauptmikrofon wurde in erster Linie in Hinblick auf stereofone Abbildungseigenschaften und nicht auf räumliche Abbildung optimiert. Für die optimale räumliche Abbildung wäre ein passendes Verhältnis von direktem / indirektem Schall (R/D – Ratio) in den vorderen Kanälen vonnöten und dafür wiederum spezielle Mikrofon – Charakteristika, Anordnung und Platzierung des Hauptmikrofons.

Somit sind entsprechende Pegelanpassungen des indirekten Schalls in erster Linie in den Surroundkanälen möglich.

Es kann angenommen werden, dass das Vorne / Hinten – Verhältnis des halligen Anteils des Schalles ungefähr der natürlichen Situation (für eine bestimmte räumliche Perspektive) gleichen soll. Dazu sollten für die vier Kanäle L, R, LS, RS in etwa gleiche Pegelverhältnisse hergestellt werden und das wiederum ist durch die adäquate Mikrofon – Platzierung möglich.

Pegelanpassung des indirekten Schalls ausschließlich in den Surroundkanälen ist nicht das optimale Werkzeug zur natürlichen Darstellung einer räumlichen Perspektive. In [18] wird zur Aufnahme von umgebendem, indirektem und einhüllendem Schall die Verwendung von (mindestens) vier äquivalenten Kanälen vorgeschlagen. Davon werden zwei Kanäle für vorne und zwei für hinten verwendet, wobei jedes Paar L-R, R-RS, RS-LS, LS-L eine stereofone Darstellung der frühen Reflektionen und des Nachhalls liefert. Eine entsprechende quadratische Anordnung ist in Abb. 17 dargestellt.

Die vier Mikrofone sollten unter Berücksichtigung der Prinzipien der Wahrnehmung von einhüllendem Schall angeordnet werden. Im Fall von Nachhall kann der reflektierte Schall als von Spiegelquellen generierter Schall [34] angesehen werden. Das resultierende Schallfeld kann als ‚subjektiv diffus‘ bezeichnet werden, wenn eine gleichmäßige Verteilung der Hörereignisse um den Hörer herum wahrgenommen werden kann. In [57] wurde die Wahrnehmung von subjektiv diffusen Schallfeldern untersucht, - speziell die Auswirkungen vom Vorhandensein von Kohärenz¹⁴ zwischen den einzelnen Kanälen in einer quadratischen Lautsprecheranordnung um den Hörer herum. Die Ergebnisse zeigen, dass ein niedriger Kohärenzgrad ($k < 0,2$) zwischen den einzelnen Kanälen eine Wahrnehmung von Sammlung der Hörereignisrichtungen in den Bereichen um die Lautsprecher zufolge hat. Je inkohärenter die Lautsprecher signale, welche Beiträge liefern, sind, umso eher ist das Gehör in der Lage die Schallquellen (Lautsprecher) zu unterscheiden. Anders gesagt: Zunehmende Kohärenz zwischen den einzelnen Kanälen verursacht Verschmelzungseffekte. Siehe dazu auch Abb. 18.

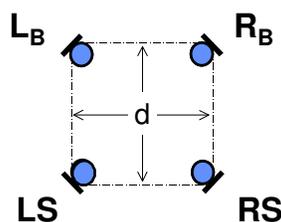


Abb. 17: Quadratische Anordnung für räumliche Abbildung / Atmo-Kreuz, aus [18]

Es sind prinzipiell (mindestens) vier äquivalente stereofone Kanäle wünschenswert, um einen realistischen räumlichen Eindruck und eine einhüllende Atmosphäre zu erlangen. Das Quadrat sollte in Hinblick auf benötigte Korrelation (s.u.) zwischen den Kanälen angeordnet werden. Der Aufstellungsort des Quadrates sollte normalerweise weit außerhalb des Hallradius liegen um eine hohe R/D – Ratio zu erlangen. Weiter werden der Aufstellungsort und die Wahl der Mikrofon – Charakteristika immer von der beabsichtigten stereofonen Perspektive und der jeweiligen Aufnahmesituation abhängen (wird eine Aufnahme mit wenig indirektem Schall gewünscht, so können dafür z.B. Nierencharakteristika und ein Aufstellungsort nahe dem Hallradius gewählt werden).

¹⁴ Der Kohärenzgrad ist gleich dem maximalen Absolutwert der normalisierten Kreuzkorrelationsfunktion (siehe dazu [34] und [58]) „Kohärenz zwischen den Kanälen“ ist ein relevantes Maß für interaurale Beziehungen und Beziehungen zwischen den Kanälen beim Räumlichen Hören.

Der Grad der subjektiven Diffusität oder der subjektiven Einhüllung hängt vom Wert d des Atmo-Kreuzes ab. Ist d zu groß, folgt daraus eine Verminderung der subjektiven Einhüllung. Die ausgewogene Verteilung der einhüllenden Quellen bricht ein und es bilden sich „Wolken von Hörereignissen“ in den Bereichen um die Lautsprecher. Geht d gegen Null (d.h. es liegt bereits eine koinzidente Anordnung vor), so wird eine Phantomquelle über dem Kopf des Hörers wahrnehmbar und die subjektive Einhüllung verschwindet.

Optimale Werte für d sollten im Bereich von $d = 25$ cm (bei der Verwendung von Nierencharakteristika, siehe auch [59]) oder $d = 40$ cm bei der Verwendung von Kugelcharakteristika) liegen.

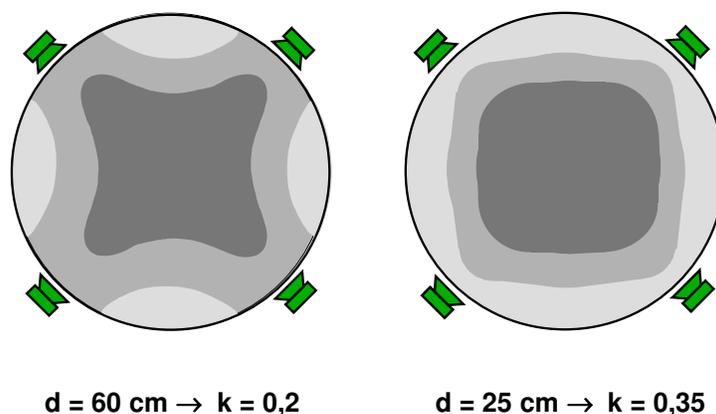


Abb. 18: Effekte von Kohärenz auf die subjektive Diffusität, [57]

Wahrgenommene Richtungen in einem diffusen Schallfeld, welches durch das abgestrahlte Rauschen (0,25 – 2,5 kHz) von vier Lautsprechern mit Kohärenzfaktoren zwischen den Kanälen von $k = 0,2$ und $k = 0,35$ erzeugt wurde. Das Rauschen wurde in einem Hallraum mit dem Atmo-Kreuz unter Verwendung von Nierencharakteristika und Abständen von $d = 60$ cm und $d = 25$ aufgenommen. Dunklere Bereiche markieren wahrscheinlicheres Auftreten von richtungsindizierenden Frequenzen von Hörereignissen.

Es ist klar, dass eine Reihe von praktischen Aspekten (z.B. Gegebenheiten am Aufnahmeort) einen Einfluss haben werden, und dass Kompromisse gemacht werden müssen. Wenn z.B. der Sweet Spot so groß wie möglich werden soll, so muss die Korrelation zwischen den einzelnen Kanälen minimiert und demnach der Abstand d maximiert werden. Wenn andererseits die natürliche Abbildung des Raumes und der Einhüllung die vordergründigen Anforderungen sind, dann ist die Wahrung der Schallankunftszeiten und die Beibehaltung der Diffusfeld-Effekte wichtig und der Wert d muss dementsprechend bestimmt werden.

Die Platzierung der in 3.2.6.1 eingangs beschriebenen OCT – Anordnung orientiert sich vor allem am Aufnahmewinkel, während der Aufstellungsort des Vier – Kanal – Raummikrofons anhand der Eigenheiten des indirekten Schalls in einem bestimmen Raum gewählt werden soll.

Durch die Kombination dieser beiden Konstellationen soll nun ausreichende Flexibilität gewährleistet sein, da die vorrangige Abbildung von Richtung und Raum durch zwei unterschiedliche ‚Werkzeuge‘ realisiert wird.

Für die optimale Abbildung der Richtung und des Spektrums kann nun eben die OCT – Anordnung (oder eine beliebige andere Hauptmikrofon – Anordnung) verwendet werden, welche gleichzeitig mit einer niedrigen R/D - Ratio noch genügend ‚Platz‘ für die Raumabbildung mittels Vier – Kanal – Raummikrofon lässt. Somit kann nun die quadratische Mikrofonanordnung angewandt werden, um ein realistisches Abbild der räumlichen Situation und eine einhüllende Atmosphäre zu generieren.

Aufgrund der unterschiedlichen Aufstellungsorte der beiden Anordnungen entstehende Laufzeitdifferenzen zwischen den Signalen in den Kanälen L_B , R_B , RS , LS und den Signalen in den Kanälen L_F , C , R_F . Um die Vorteile der einzelnen Konfigurationen optimal zu nutzen, muss diese Laufzeit berücksichtigt und entsprechend korrigiert werden (siehe 3.2.6.4 und 3.2.6.5).

3.2.6.2 Mehrkanal – Mikrofon – Arrays nach Williams (vgl. [11], [12])

Allgemeines

Ursprünglich war diese Aufnahmetechnik für Wiedergabekonfigurationen ersonnen, bei welchen die Lautsprecher auf einem Kreis in gleich großen Abständen zueinander standen. Weiterentwickelt wird sie für das Format 3/2 verwendet.

Dabei wird das Schallfeld segmentiert und ein von einem Mikrofonpaar (Nierencharakteristik) aufgenommenes Segment wird über ein jeweils korrespondierendes Lautsprecherpaar wiedergegeben. Es ist hierbei möglich Mehrkanal – Mikrofon – Arrays zu gestalten, bei welchen dem Toningenieur sehr viele Freiheiten bezüglich der Verteilung der Haupt - Schallquelle über einen beliebigen Winkel offen stehen. Weiter ist es möglich die natürliche Struktur der ersten Reflektionen und des Nachhalles beizubehalten und zu integrieren.

Um das Array entsprechend zu gestalten, bedarf es der Kenntnis der fundamentalen Charakteristika eines jeden Segmentes und einer entsprechenden Handhabung der jeweiligen Parameter. Obwohl die Betrachtung und grafische Darstellung in zweidimensionalen Segmenten erfolgt, muss beachtet werden, dass wie bei Stereoaufnahmen das aufgenommene (3 – dimensionale) Schallfeld nur in einer horizontalen Ebene der Lautsprecher wiedergegeben wird. Insofern unterliegt man trotz Mehrkanal – Surround – Sound mit diesen Systemen immer noch Beschränkungen bezüglich der Abbildung des gesamten Schallfeldes.

Segmentierung des Schallfeldes

Der Aufnahmebereich / Abdeckwinkel (s.u.) eines Mikrofonpaares ist relativ zum Winkel der beiden Hauptachsen der Mikrofone beweglich. Damit wird es möglich, die einzelnen Aufnahmebereiche des Mikrofonarrays so aneinander zufügen, dass eine glatt verlaufende Aufnahme des Schallfeldes erreicht wird.

Um eine einfache und flexible Handhabung zu erreichen, wurde bereits der Prototyp einer Software vorgestellt [12]. Sie arbeitet basierend auf den Williams - Kurven und erlaubt den PC - unterstützten Entwurf von jenen Mikrofon - Arrays. Damit wird den verschiedenen Anforderungen für die spezifische Schallfeld – Segment - Aufnahme und Wiedergabe Rechnung getragen.

Bei der Zweikanal - Stereophonie wird von IID's und ITD's Gebrauch gemacht, um für verschiedene Aufnahmesituationen möglichst optimale Mikrofonsysteme für Aufnahme und Wiedergabe zu entwickeln (XY, ORTF, OSS, KFM, Kunstkopf, AB(C), etc. ...). Hier werden diese Prinzipien und die Möglichkeiten, welche in [13], [14], [15], [16] beschrieben wurden auf Mikrofon – Arrays für Mehrkanal – Systeme angewandt.

Per Definitionen wird in obiger Literatur [13], [14] der Aufnahmewinkel (Stereophonic Recording Angel – SRA) immer mit $+ / - 50^\circ$ angegeben. Das war sinnvoll, um Verwirrung mit dem Winkel zwischen den Mikrofonen zu vermeiden. Im Fall der Mehrkanal – Mikrofon – Arrays ist die Aufnahme/Wiedergabe Symmetrie bezogen auf die jeweiligen Array – Segmente nicht mehr gegeben (die Mittenachse eines segmentierten Schallfeldes ist im Allgemeinen ungleich der Mittenachse des entsprechenden Lautsprecherpaares). Deshalb wurde der Begriff ‚Abdeckwinkel‘ (Coverage Angel) eingeführt. Dieser bezeichnet den gesamten Bereich, welcher durch den SRA eines einzelnen Mikrofonpaares in einem Mikrofon – Array abgedeckt wird. In obigem Beispiel wird also aus dem Aufnahmewinkel $+ / - 50^\circ$ ein Abdeckwinkel von 100° .

Grundsätzliches

Betrachtet sei ein einfaches Mikrofonpaar, dessen aufgenommenes Schallfeld von zwei Lautsprechern wiedergegeben wird. Bei stereofoner Wiedergabe ist links / rechts – Symmetrie des wiedergegebenen Schallfeldes normalerweise wünschenswert. Der Aufnahmewinkel von zwei Mikrofonen kann entweder kleiner oder größer als der physikalische Winkel zwischen den Hauptachsen der Mikrofone sein, und ist normalerweise symmetrisch in Bezug auf die Achse zwischen den Mikrofonen.

Beim Entwurf von Mehrkanal – Mikrofon – Arrays muss man nun den Winkel – Offset des wiedergegebenen Schallfeld – Segmentes in Relation zur Symmetrieachse zwischen den Mikrofonen kontrollieren (s.u.). Betrachtet man die vordere Dreiergruppe von Mikrofonen, so muss der Offset der zwei wiederzugebenden Schallfeld - Segmente bei der Aufnahme so eingestellt werden, dass deren jeweils innere Grenze mit der Hauptachse des nach vorne gerichteten Center – Mikrofons übereinstimmen. Für die seitlichen Segmente muss gelten, dass der entsprechende Abdeckwinkel frei und ohne jeglicher Koinzidenz zwischen den physikalischen Achsen der Mikrofone und den Grenzen der entsprechenden Segmente eingestellt werden kann. Für das hintere Segment wird wieder zur symmetrischen Stereophonie übergegangen.

Offset und Kopplung der Segmente

- In Abb. 19a ist der herkömmliche stereofone Aufnahmewinkel (SRA) dargestellt, er ist jedoch größer gewählt, als der physikalische Winkel zwischen den Mikrofonen.
- Abb. 19b zeigt nun, wie dieser Abdeckwinkel im Uhrzeigersinn mittels positivem Winkel – Offset gedreht wird.

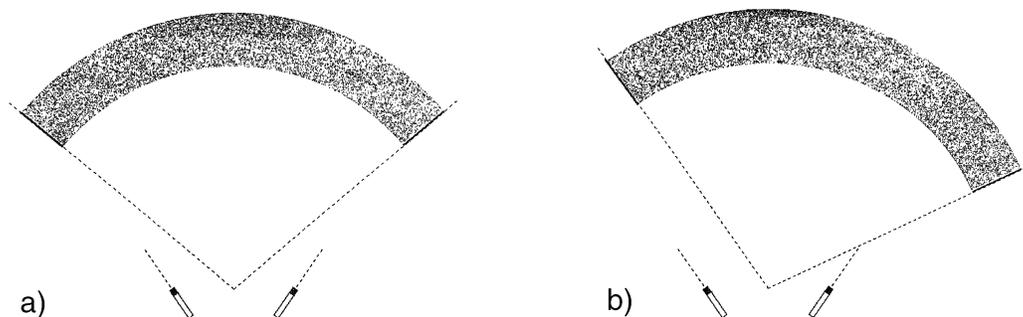


Abb. 19, aus [11]: a): Stereofoner Aufnahmewinkel, größer als Winkel zwischen Mikrofonen. b): Positiver Offset von 15° (Linke SRA – Grenze mit Mikrofon – Achse ausgerichtet)

Obige zwei Punkte sind natürlich auch mit einem SRA, welcher kleiner als der physikalische Winkel zwischen den Mikrofonen ist, und auch gegen den Uhrzeigersinn möglich.

In Abb. 19a) und b) ist der SRA in seinem tatsächlichen Ursprung zwischen den Mikrofonkapseln dargestellt. In Abb. 19b) ist die linke Grenze des Abdeckwinkels parallel zur Hauptachse des linken Mikrofons ausgerichtet.

Der Vorgang des Ankoppelns wird in Abb. 20 und Abb. 21 dargestellt. Da die Schallquelle in Bezug auf den Abstand der Mikrofone zueinander relativ weit entfernt ist, ist es eine vertretbare Annäherung den grafischen Ursprung der SRA - Abdeckung beim Schnittpunkt der Mikrofon - Achsen anzusetzen. Das ermöglicht es im Folgenden den Vorgang des Ankoppelns besser darzustellen. Dieser Vorgang ist wesentlich für eine glatte, verlaufende Gesamt - Abdeckung beim Entwurf von Mehrkanal - Mikrofon - Arrays.

Abb. 20 zeigt die vordere Dreiergruppe eines Mehrkanal – Mikrofon - Arrays mit den entsprechenden Abdeckwinkeln. Die Mikrofone bilden vorderes linkes und rechtes Segment indem sie sich das Center – Mikrofon teilen. Die Ankopplung der beiden Abdeckwinkeln, welche jeweils größer als die physikalischen Winkeln zwischen den Mikrofonen sind, wurde durch Winkel – Offset erreicht: Negativer Winkel – Offset für FLS (Front Left Segment) und positiver Winkel – Offset für FRS (Front Right Segment).

Analog zu diesem Entwurf ist es möglich, eine vordere Dreiergruppe mit jedem beliebigen Abdeckwinkel zu kreieren, womit der Toningenieur flexibler bei der Gestaltung der Hauptschallquelle agieren kann.

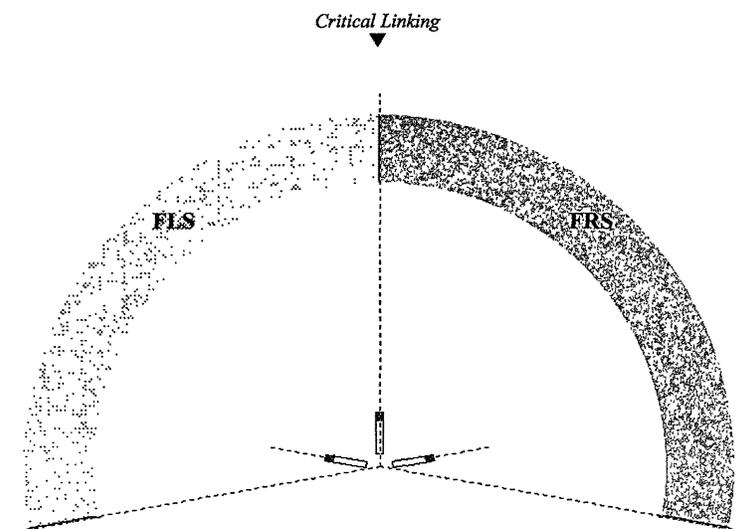


Abb. 20, aus [11]: Ankopplung des ‚Front Left Segments‘ (FLS) und des ‚Front Right Segments‘ (FRS) bei der vorderen Dreiergruppe eines Mehrkanal – Mikrofon – Arrays

In Abb. 21 ist dargestellt, wie der Vorgang des Ankoppelns auf das gesamte Mehrkanal – Mikrofon – Array angewandt wurde. Nur das vordere Center - Mikrofon ist mit der Kopplungsgrenze zweier Abdeckwinkel (FLS und FRS) ausgerichtet. Das hintere Segment ist, wie hier, normalerweise symmetrisch bezüglich der hinteren Achsen des Systems.

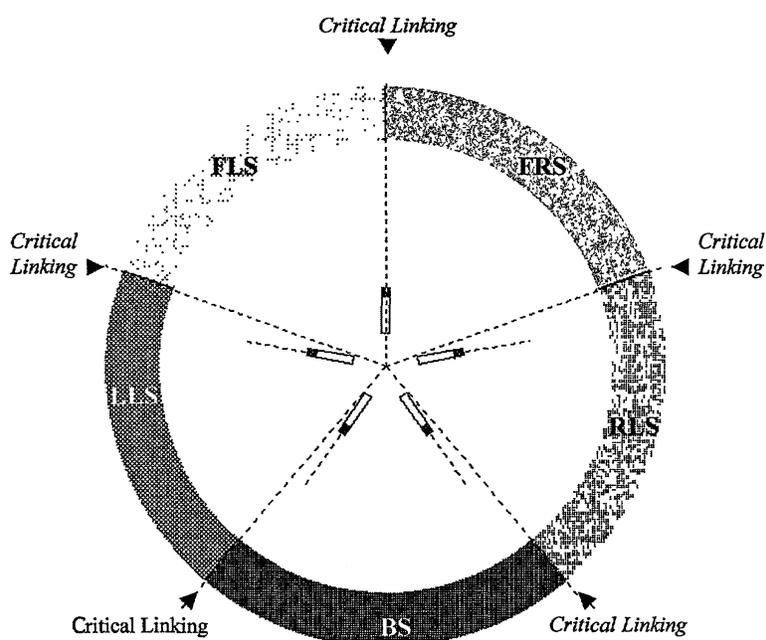


Abb. 21: Ankopplung aller Segmente eines Mehrkanal – Mikrofon – Arrays, vgl. [11]

Erzeugung von Intensitäts – und Zeitoffset

Die einzelnen Methoden der jeweiligen Offset – Technik sind bemerkenswert einfach. Die Schwierigkeit liegt vielmehr darin, die richtige Technik auszuwählen, sodass durch die Ankopplung ein glatt verlaufendes Schallfeld erreicht wird.

Grundsätzlich gibt es vier verschiedene Offset – Typen, die angewandt werden können:

- Elektronischer Intensitäts – Offset: Addition einer konstanten Intensitätsdifferenz zur Intensitäts/Zeit – Differenz – Funktion zwischen zwei Mikrofonen
- Elektronischer Zeit – Offset: Addition einer konstanten Zeitdifferenz zur Intensitäts/Zeit – Differenz – Funktion zwischen zwei Mikrofonen
- Intensitäts – Offset durch Mikrofonposition
- Zeit – Offset durch Mikrofonposition

Die beiden letzten Punkte bedingen einander: Nimmt man einen Intensitäts – Offset durch Mikrofonposition in die eine Richtung vor, entspricht das gleichzeitig einem Zeit – Offset durch Mikrofonposition in die entgegengesetzte Richtung.

In Abb. 22 sind die Mikrofon – Hauptachsen und die Intensitäts- bzw. Zeitachse eines Stereopaars ohne Offset dargestellt. Abb. 23 a) und b) illustrieren nun den Offset durch Mikrofonposition, und betrachten ihn aus zwei verschiedenen Sichtweisen.

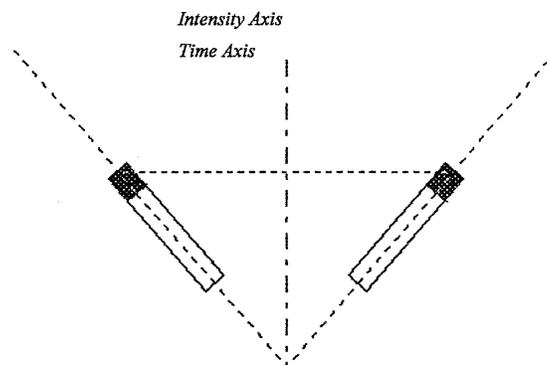


Abb. 22: Mikrofon – Hauptachsen, Intensitäts-, bzw. Zeitachse, kein Offset, [11]

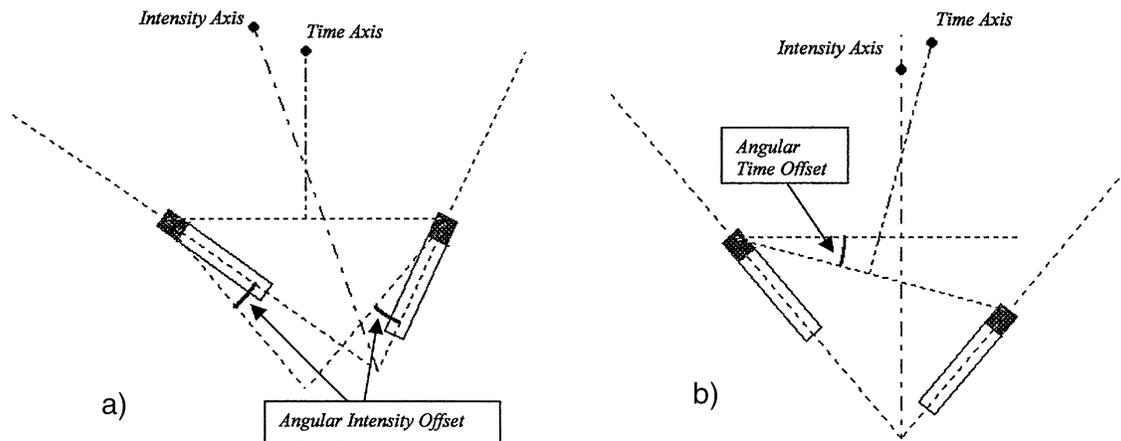


Abb. 23, aus [11]: Offset durch Mikrofonposition. a) Negativer Intensitätsoffset, b) Positiver Zeitoffset

Verwendet man einerseits die Mikrofonkapseln der Mikrofone als Rotationspunkt und dreht beide Mikrofone z.B. 15° gegen den Uhrzeigersinn, so dreht man damit die Intensitätsachse ($\Delta I = 0$) ebenfalls im selben Ausmaß: Negativer Intensitäts – Offset durch Mikrofonposition.

Bewegt man andererseits das rechte Mikrofon nach hinten (konstanter Abstand zwischen den Mikrofonkapseln, Orientierung der Mikrofonachsen bleibt gleich), so dreht man damit die Zeitachse im selben Ausmaß im Uhrzeigersinn: Positiver Zeit – Offset durch Mikrofonposition von 15° .

Durch beide Sichtweisen entsteht schlussendlich das gleiche Resultat, nämlich die relative Orientierung der Intensitäts- und Zeitachsen zueinander.

Abgesehen davon existiert aber natürlich auch ein Unterschied zwischen elektronischem Offset und Offset durch Mikrofonposition.

- Elektronischer Offset entsteht einfach durch das Hinzufügen eines konstanten Intensitäts- oder Laufzeitwertes zur Intensitäts/Zeitfunktion eines Mikrofonpaares, welches ein bestimmtes Segment abdeckt.
- Offset durch Mikrofonposition entsteht durch Veränderung der physikalischen Position der Mikrofone, welche ein Paar bilden. Dadurch entsteht ein Winkel zwischen der Intensitäts- und der Zeitachse.

Per Definition wird positiver Offset als jener Offset definiert, mit welchem eine Drehung des Abdeckwinkels im Uhrzeigersinn verursacht wird (abgebildetes Panorama wandert nach links, was wiederum einer Drehung des Aufnahme / Abdeckwinkels nach rechts entspricht).

So wird z.B. positiver Offset erreicht durch:

- Elektronischen Zeitoffset: Rechtes Mikrofon wird in Relation zum linken verzögert
- Elektronischen Intensitätsoffset: Rechtes Mikrofon wird in Relation zum linken abgeschwächt
- Offset durch Mikrofonposition: Rechtes Mikrofon wird im Uhrzeigersinn gedreht (in diesem Fall zurück), wobei die linke Mikrofonkapsel als Drehmittelpunkt verwendet wird, zwischen den Mikrofonen gleicher Abstand und zwischen deren Achsen derselbe Winkel beibehalten wird.

Jeder dieser unterschiedlichen Offsets hat eine Verschiebung des Graphen in der jeweiligen Intensitäts / Zeit Differenzfunktionen zufolge. Entweder nach links für positiven Zeitoffset, oder nach unten für positiven Intensitätsoffset. Zur Veranschaulichung wurden die Schallquellenkoordinaten, welche die relative Position des Abdeckwinkels angeben, in Ellipsen auf den Graphen eingetragen.

Als Beispiele seien in Abb. 24 bis Abb. 27 die verschiedenen Offsetmöglichkeiten für einen Abdeckwinkel, welcher größer als der physikalische Winkel zwischen den Mikrofonen ist, dargestellt.

Elektronischer Offset

Abb. 24 zeigt die Intensitäts / Zeit Differenzfunktionen für ein Mikrofonpaar mit Abstand 25cm und einem Winkel von 70° zueinander ohne jeglichen Offset. In Abb. 25 ist ein positiver elektronischer Zeitoffset von -0.28 ms (das rechte Mikrofon wurde in Relation zum linken verzögert) dargestellt. Die gesamte Intensitäts / Zeit Differenzfunktion wurde nach links geschoben und auch der Ursprung der Schallquelle (0°) befindet sich nun bei -0.28 ms.

In Abb. 26 wurde ein positiver elektronischer Intensitätsoffset von -2.5 dB zugewiesen (das rechte Mikrofon wurde in Relation zum linken abgeschwächt). In Vergleich zu Abb. 24 wurde die Intensitäts / Zeit Differenzfunktionen nun um 2.5 dB nach unten verschoben. Der Ursprung der Schallquelle (0°) folgte dieser Verschiebung entsprechend und ist nun auf der Intensitätsachse bei -2.5 dB zu finden.

In diesen Beispielen wird der Einfluss der verschiedenen Arten von Offset auf den stereofonen Aufnahmewinkel (SRA) deutlich. In Abb. 24 beträgt der SRA $\pm 50^\circ$ (entsprechend einem Abdeckwinkel von 100°). In Abb. 25 verursachte der Offset eine

gesamte Abdeckung von 103° , - nun aber eine unsymmetrische Abdeckung von -35° (links) bis $+68^\circ$ (rechts). In Abb. 26 hat der Intensitätsoffset von $-2,5$ dB einen Abdeckwinkel von 105° zufolge, wiederum unsymmetrisch von -35° (links) bis $+70^\circ$ (rechts).

Es ist also festzuhalten, dass die Verwendung von Offset den Abdeckwinkel leicht verändern kann. Dies ist bei der endgültigen Gestaltung des Mikrofonarrays zu berücksichtigen.

Es kann hiermit gesehen werden, wie mit positivem und negativem Intensitäts- und Zeitoffset beinahe jede beliebige (Sinn machende) Position des Abdeckwinkels erreicht werden kann. Exzessive Offsetwerte können allerdings unerwartete Artefakte hervorbringen.

Dazu folgende Anmerkungen:

- Es muss einen Schnittpunkt zwischen der Intensitäts / Zeit Differenzfunktion und den psychoakustischen Kurven für $\pm 30^\circ$ geben, ansonsten kann es keine korrekte Ankopplung aller Segmente geben.
- Die Winkelverzerrung ist abhängig vom Schnittpunkt der Intensitäts / Zeit Differenzfunktion mit den psychoakustischen Kurven (von 10° , 20° und 30°). Diese repräsentieren die Positionen der einzelnen, wiedergegebenen Schallquellen als Phantomschallquellen zwischen den Lautsprechern. Wie aus Abb. 25 und Abb. 26 hervorgeht, kann es durch zu große Offsetwerte aufgrund der Nichtlinearität der Intensitäts / Zeit Differenzfunktion zu einer wahrnehmbaren Vergrößerung der Winkelverzerrung kommen.
- Vorsichtig muss mit dem elektronischem Intensitätsoffset umgegangen werden, da dieser die Balance der allgemeinen Energieverteilung verändert.

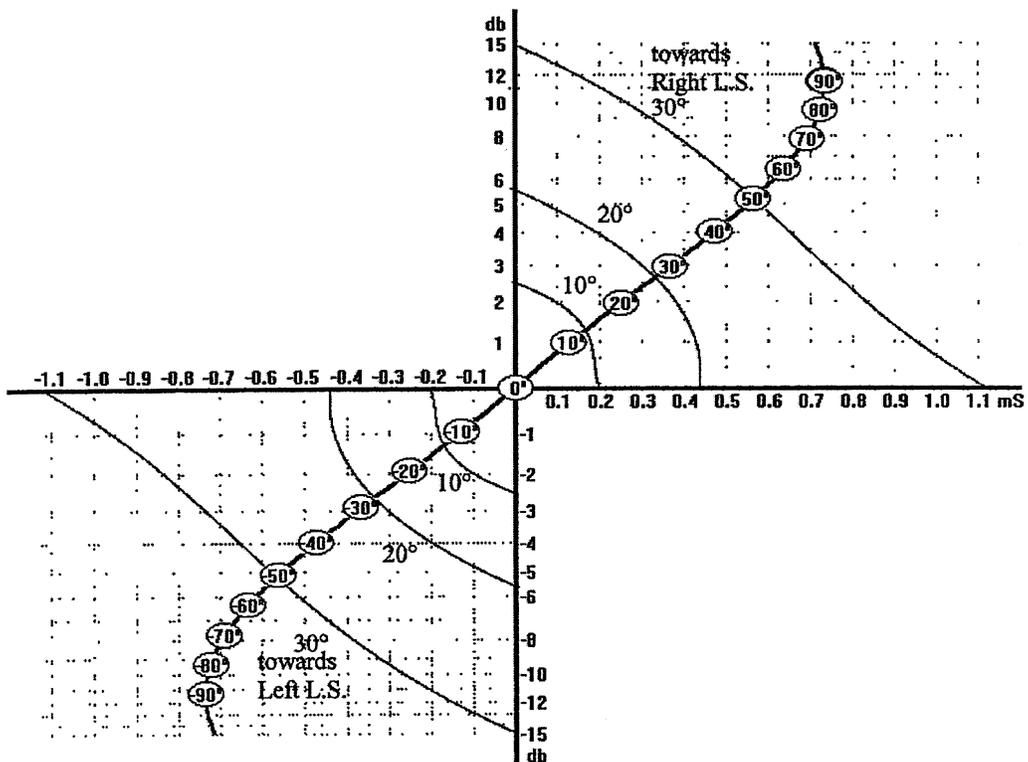


Abb. 24: Intensitäts / Zeit Differenzfunktion ohne Offset, [11]

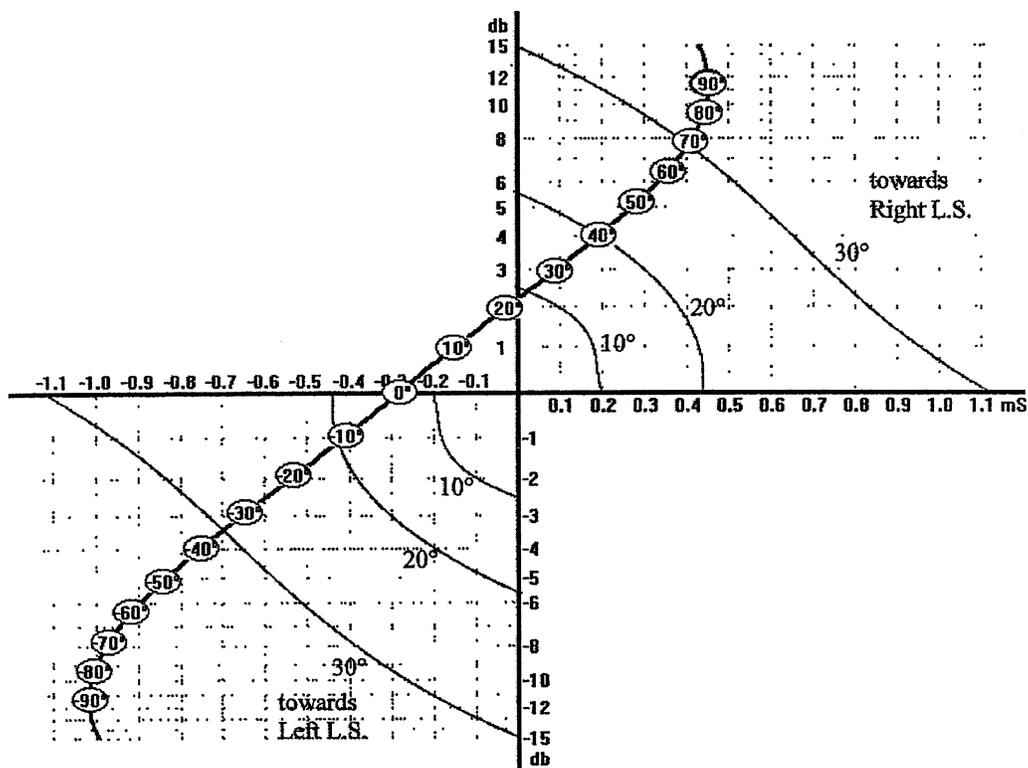


Abb. 25: Intensitäts / Zeit Differenzfunktion mit pos. Elektron. Zeitoffset: -0.28ms, aus [11]

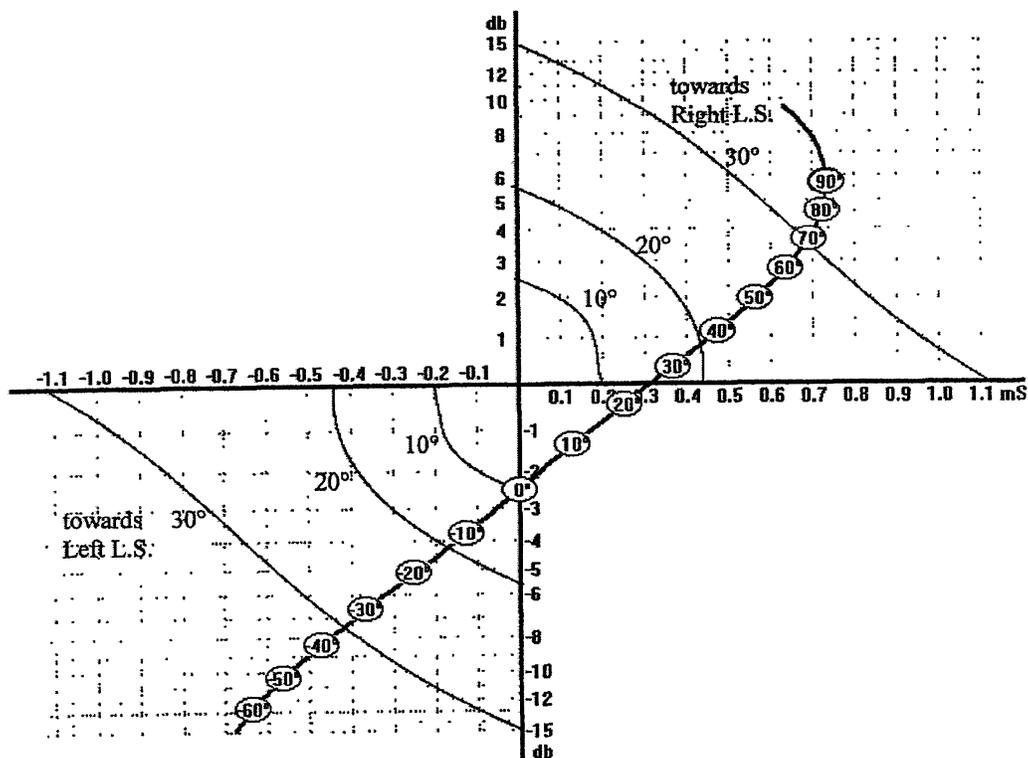


Abb. 26: Intensitäts / Zeit Differenzfunktion mit pos. elektron. Intensitätsoffset: -2.5dB, [11]

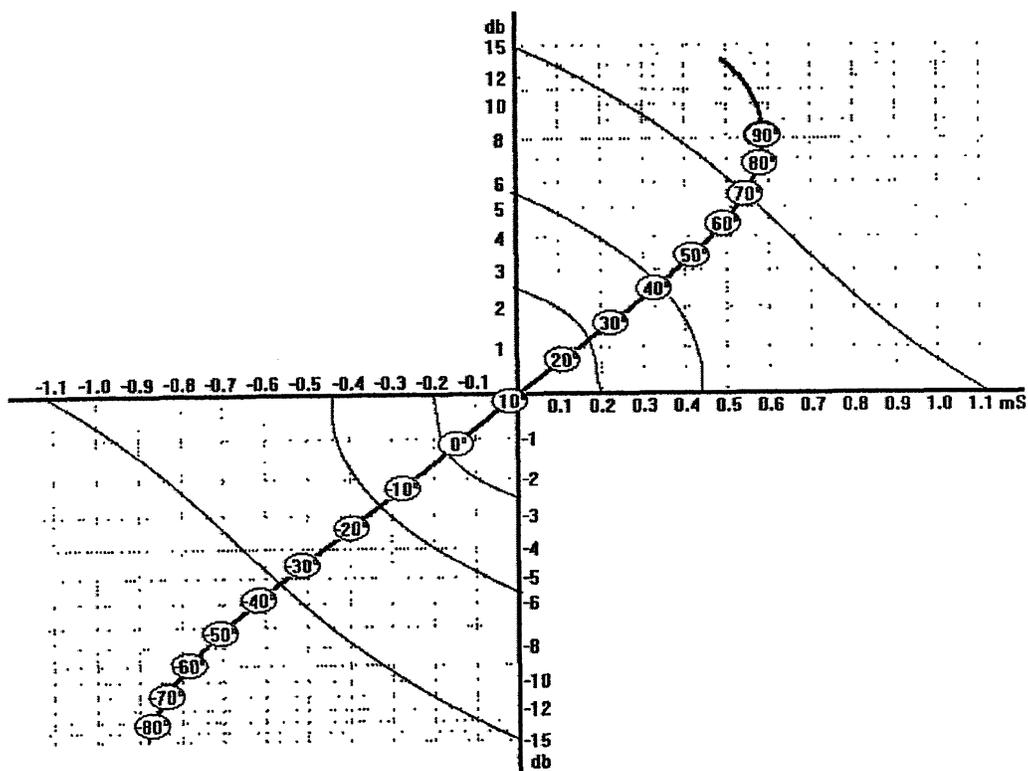


Abb. 27: Intensitäts / Zeit Differenzfunktion mit positivem elektronischen Intensitätsoffset von -1.15dB und pos. elektronischen Zeitoffset von -0.145ms, vgl. [11]

In Abb. 27 ist nun dargestellt wie es möglich ist, durch eine Kombination von elektronischem Intensitäts- und Zeitoffset den Graphen wieder durch den Intensitäts- und Zeiter sprung (0dB und 0ms) wandern zu lassen.

Offset durch Mikrofonposition

In Abb. 28 ist der Effekt dargestellt, den ein Zeitoffset durch Mikrofonposition in der Intensitäts / Zeit Differenzfunktion bewirkt. Angewandt wurde eine Winkelrotation von $+37^\circ$ (Drehung des rechten Mikrofons von 37° im Uhrzeigersinn, Mikrofonkapsel des linken Mikrofons als Drehmittelpunkt, konstanter Winkel zwischen den Mikrofonachsen).

Kombinierte Offsets

In Abb. 29 wird der gleiche Abdeckwinkel – Offset durch eine Offset – Kombination erreicht: Offset von 13° durch Mikrofonposition und -1.3 dB elektronischen Intensitätsoffset. Das hat im Vergleich zu Abb. 24, wo kein Offset angewandt wurde einen Abdeckwinkel – Offset von 15° zufolge.

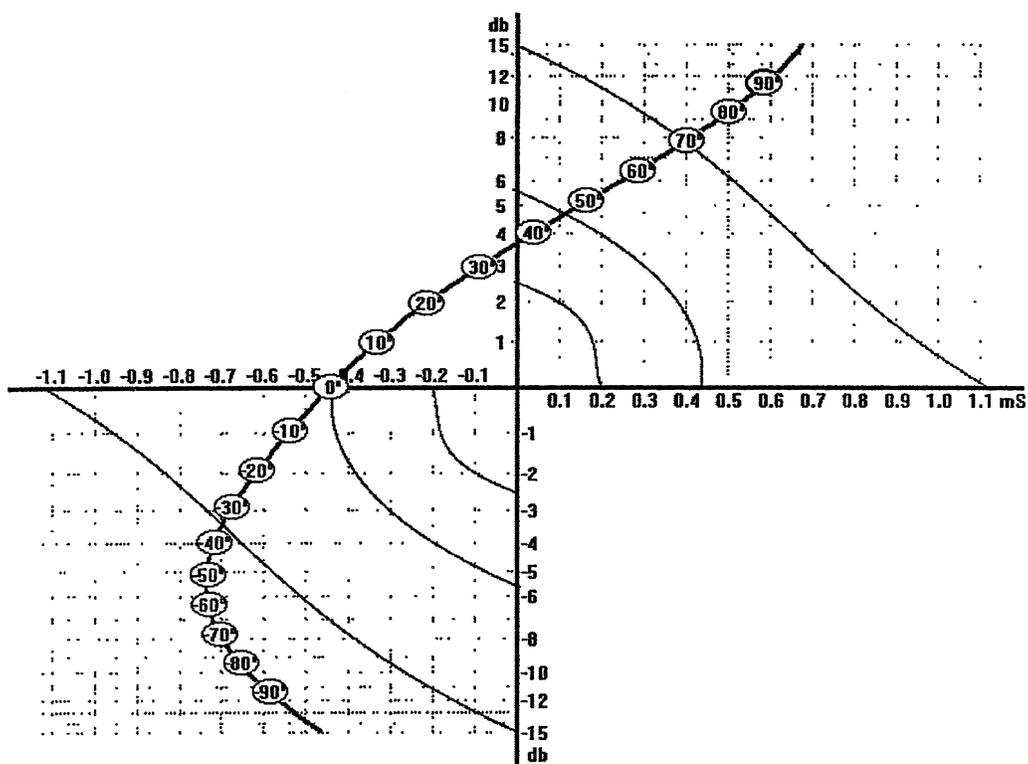


Abb. 28: Intensitäts / Zeit Differenzfunktion mit pos. Offset durch Mikrofonposition von 37°

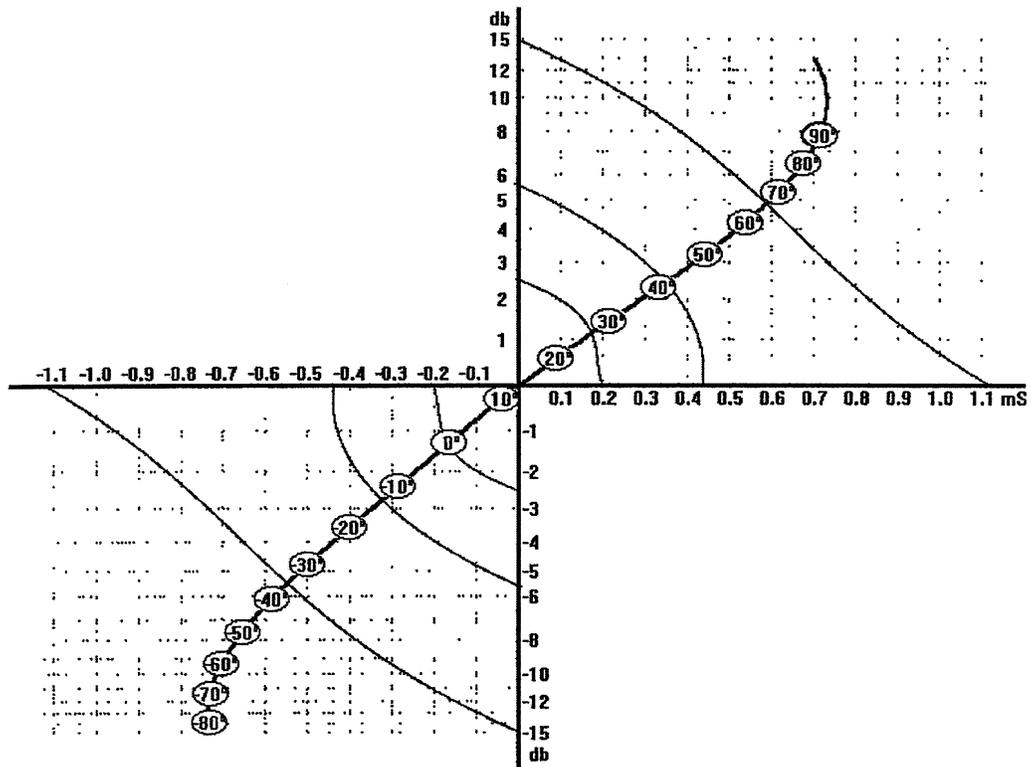


Abb. 29: Intensitäts / Zeit Differenzfunktion mit positivem Offset durch Mikrofonposition von $+13^\circ$ und positivem elektronischem Intensitätsoffset von -1.3dB , aus [11]

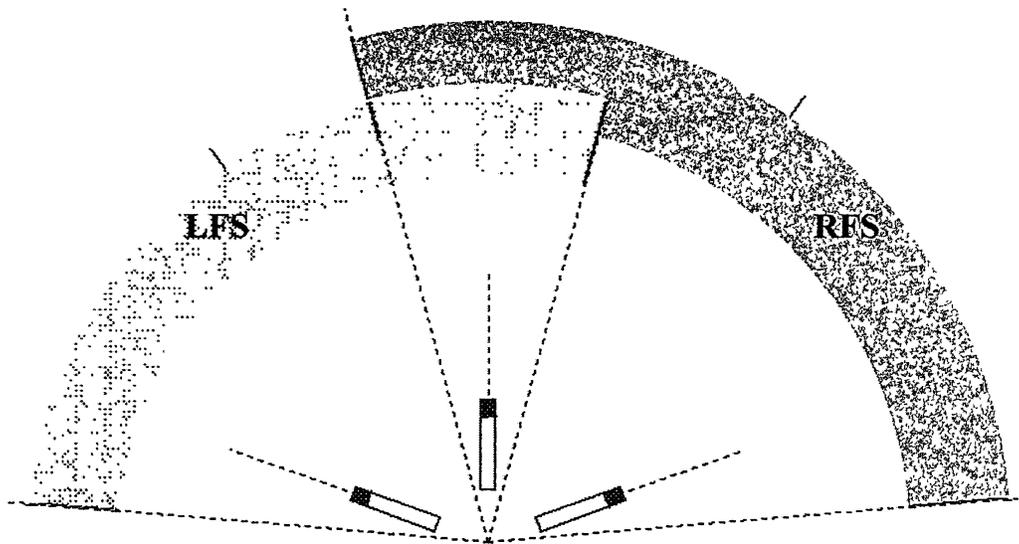


Abb. 30: Vordere Dreiergruppe 25cm / 70° - Abdeckwinkel 100° pro Segment - kein Offset, aus [11]

Kopplung der vorderen zwei Abdeckwinkel

In Abb. 30 wird die vordere Dreiergruppe durch Mikrofone mit jeweils 25cm / 70° zwischen den Mikrofonen repräsentiert. Der SRA beträgt +/-50°, der Abdeckwinkel demnach 100°. Demnach wird ein Winkel - Offset von 15° benötigt, um die Grenzen der Abdeckwinkel mit der Achse des mittleren Mikrofons auszurichten

Anstelle einer symmetrischen Abdeckung von +/-50° wird hier bei Abb. 30 versucht eine Abdeckung von -65° bis +35° bzw. von +65° bis -35° zu erreichen (negativer Winkel – Offset für das LFS bzw. positiver Winkel – Offset für das RFS). Dadurch wird schließlich eine verlaufende Ankopplung, wie sie in Abb. 20 dargestellt, ist erreicht.

Die Beispiele in Abb. 25 bis Abb. 29 illustrieren die Anwendung von Offset bei der Gestaltung der vorderen Dreiergruppe eines Mikrofon – Arrays (jedes Mikrofon – Paar mit einem Abstand von 25 cm zwischen den Mikrofonkapseln und einem Winkel von jeweils 70° zwischen den Mikrofonachsen). Aus diesen Methoden kann nun gewählt werden um den gewünschten Offset zu erreichen und die jeweils inneren Grenzen der Abdeckwinkeln mit der Achse des mittleren Mikrofons auszurichten.

Zusammenfassung der angewandten Methoden der vorangegangenen Abbildungen:

- Abb. 25: Positiver elektronischer Zeitoffset von -0.28ms (das rechte Mikrofon ist in Bezug auf das linke verzögert)
- Abb. 26: Positiver elektronischer Intensitätsoffset von -2.5dB (das rechte Mikrofon ist in Bezug auf das linke abgeschwächt)
- Abb. 27: Kombination von positivem elektronischen Intensitätsoffset von -1.15dB und positivem elektronischen Zeitoffset von -0.145ms
- Abb. 28: Positiver Offset durch Mikrofonposition von 37°
- Abb. 29: Kombination von positivem Offset durch Mikrofonposition von +13° und positivem elektronischem Intensitätsoffset von -1.3dB

Gestaltung der Segmente eines Mehrkanal – Mikrofon – Arrays (MMA)

Es gibt drei verschiedene Stadien bei der Gestaltung eines kompletten MMA:

1. Gestaltung der vorderen Dreiergruppe
2. Wahl der hinteren Segment – Abdeckung
3. Ankopplung der seitlichen Segmente

Gestaltung des Mehrkanal – Mikrofon – Arrays (MMA)

Es ist ratsam solche Mikrofon Abstand/Winkel – Kombinationen zu wählen, sodass der Abdeckwinkel der jeweiligen Mikrofonpaare annähernd gleich dem Winkel zwischen den Mikrofonen ist. Das hat zur Folge, dass nur kleine Offsetwerte für eine glatt verlaufende Ankopplung der einzelnen Segmente notwendig sein werden. Es ist dagegen nicht notwendig Abdeckwinkel zu wählen, welche den Winkeln zwischen den Lautsprechern ähnlich sind.

Vordere Dreiergruppe (Front Facing Triplet - FFT)

Der Abdeckwinkel (**Coverage Angel – CA**) der vorderen Dreiergruppe (**Front Facing Triplet – FFT**) ist offensichtlich der wichtigste Parameter, welcher gleich am Anfang des Gestaltungsprozesses eines MMA bestimmt werden muss. Die Wahl des Abstands vom MMA zur Schallquelle und die Wahl des Abdeckwinkels bestimmen die wiedergegebene ‚Winkelgröße‘ der Schallquelle. Ein CA größer als die entsprechende Lautsprecherbasis wird die Winkelgröße des wiedergegebenen Schallbildes reduzieren, ein CA kleiner als die entsprechende Lautsprecherbasis wird die Winkelgröße des Schallbildes aufweiten – dies ist auch übergehend bis in die linken und rechten seitlichen Segmente möglich.

Diese Punkte sind Teile der persönlichen Entscheidungen des Toningenieurs und der beabsichtigten Perspektive. Die vorderen drei Lautsprechern sind verantwortlich für die Wiedergabe des FFT. Eine Überlappung, übergehend in die seitlichen Segmente, wird den Zuhörer mit einer breiteren Schallquelle einhüllen. Präzise Lokalisation über die Bereiche der seitlichen Segmente kann jedoch schon aufgrund der Prinzipien des räumlichen Hörens nicht erwartet werden. Das ist aber kein Grund, die Möglichkeit des Überlappens außer Acht zu lassen. Der Zuhörer hat die Möglichkeit seinen Kopf während des Hörens zu drehen, was wiederum die Wahrnehmung der einzelnen Segmente verändert.

Mit der Wiedergabe von breiten, ausgedehnten Schallquellen, wie z.B. Symphonie Orchestern kann die Technik des Überlappens definitiv von Vorteil sein. Die stereofone Aufnahme von Schallquellen dieser Art war immer ein Kompromiss zwischen der Aufdehnung der Mitte des

Orchesters und dem Abschneiden des Klangkörpers auf den Seiten (z.B. auf der einen Seite bei den Violinen und auf der anderen Seite bei den Cellos und den Kontrabässen). Mit dem MMA stehen nun die seitlichen Segmente zur Verfügung und können zur Vermeidung des oben beschriebenen Effektes ausgetestet werden.

Hinteres Segment

Die gewählte Kombination für das hintere Segment, welches normalerweise den Raumhall aufzunehmen hat, ist relativ beliebig. Ein zu starkes hinteres Schallfeld kann störend wirken. Da beim 5.1 Standard zwischen den hinteren Lautsprechern ein Winkel von 140° besteht, ist damit eine bei diesem Segment zu berücksichtigende Winkelverzerrung gegeben. Dies resultiert in einer Aufweitung der Mitte des hinteren Segments. Ein zu klein gewählter hinterer SRA würde diesen Effekt noch verstärken.

Seitliche Segmente

Nachdem der Abdeckbereich für die vordere Dreiergruppe und das seitliche Segment bestimmt wurde, muss der Abstand zwischen den entsprechenden Mikrofonen so gewählt werden, dass der Abdeckwinkel der seitlichen Mikrofonpaare dem benötigten Abdeckwinkel für die seitlichen Segmente entspricht. Die Orientierung der Mikrofone, welche die seitlichen Paare bilden, ist, allgemein betrachtet, nicht korrekt. Deswegen nicht korrekt, da ein Zeitoffset zwischen dem zur Seite gerichteten Mikrofon vom FFT und dem nach hinten gerichteten Mikrofon vom hinteren Segment besteht. Dieser Zeitoffset muss einfach durch Intensitäts – Differenz – Offset so kompensiert werden, dass die Segment – Orientierung im Sinne einer verlaufenden Ankopplung mit den benachbarten Segmenten stimmt.

Es stehen nun mit oben beschriebenen Techniken einige Möglichkeiten an Kombinationen der einzelnen Segmentierungen zur Verfügung. Detailliertere Beschreibungen werden in [17] gegeben.

3.2.6.3 Weitere 5 – Kanal – Mikrofonkonstellationen

Hier seien noch weitere Möglichkeiten zur Mikrofonierung im Format 3/2 angeführt. Sie sollen allerdings nicht im Detail besprochen werden, da sie zum Teil ‚Vorgängermodelle‘ sind, bzw. an die Theorie der beiden vorangestellten Kapitel 3.2.6.1 und 3.2.6.2 angelehnt sind.

INA 5

Das ist die 5 – Kanal – Version der **I**dealen **N**ieren **A**nordnung. Sie wurde konzipiert auf der Basis der Williams – Kurven um den ganzen Surround – Aufnahmebereich (360°) abzudecken.

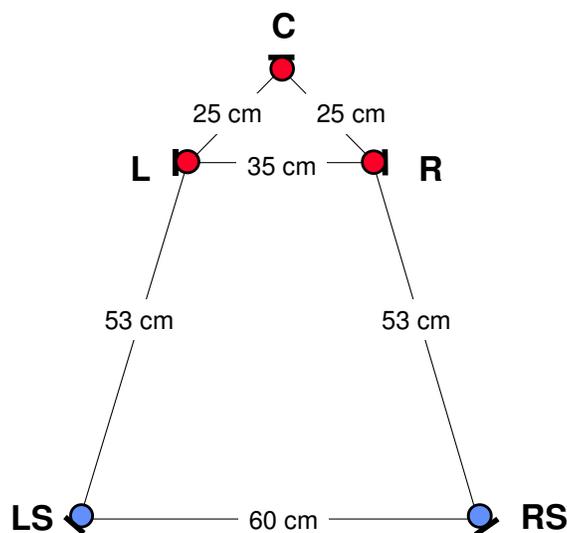


Abb. 31: Anordnung "INA 5", vgl [36]

Die Dreiecksanordnung L-C-R (INA 3, Aufnahmewinkel $\varphi = 180^\circ$) wird durch die Surround – Mikrofone LS und RS ergänzt. Die jeweiligen Entfernungen der Mikrofone zueinander wurden aufgrund der Williams – Kurven konzipiert. Es soll eine optimale Anknüpfung der drei zusätzlichen Aufnahmebereiche links, rechts und hinten (mit jeweils $\varphi = 60^\circ$) erreicht werden.

Sub - Basis	L - C	C - R	R - RS	RS - LS	LS - L
Aufnahmebereich	$-90^\circ - 0^\circ$	$0^\circ - +90^\circ$	$+90^\circ - +150^\circ$	$+150^\circ - -150^\circ$	$-150^\circ - -90^\circ$

Tabelle 8: Aufnahmebereiche von "INA 5"

Bei dieser Anordnung sind die fixierten Dimensionen zwischen den Mikrofonen für die allgemeine Verwendung problematisch, da z.B. für den vorderen Bereich ein

Aufnahmewinkel von $\varphi = 180^\circ$ gegeben ist. Dadurch muss man, wenn man den Aufstellungsort entsprechend dem Aufnahmewinkel wählt, sehr nahe an das Ensemble herangehen. Die R/D – Ratio muss an diesem Ort im Raum passen oder eben akzeptiert werden. Umgekehrt nimmt man über die Maßen Einbußen bei der Translation der abzubildenden Schallquellen in Kauf, wenn bei der Wahl des Aufstellungsortes die R/D – Ratio der bestimmende Parameter ist.

ASM 5 (vgl. [46])

Das ASM 5 (Adjustable Surround Mikrophone) besteht, ähnlich der Anordnung INA 5, aus fünf Mikrofonen, welche an fünf langen Auslegern so befestigt sind, wie das bei INA 5 vorgesehen ist. Im Gegensatz zu INA 5 lassen sich bei ASM 5 aber die Mikrofone in ihrer Richtcharakteristik einstellen. Auch die Winkel zwischen den Mikrofonen sind veränderbar.

„Fukuda-Tree“

Beim „Fukuda-Tree“ ist das den Bühnenbereich abbildende Dreieck L-C-R eine abgewandelte Form vom „Decca-Tree“. Beim „Fukuda-Tree“ werden jedoch die beim „Decca-Tree“ verwendeten Mikrofone mit Kugelcharakteristik durch Mikrofone mit Nierencharakteristik verwendet, um die indirekte Schallenergie in den vorderen Kanälen zu verringern. Aufgrund der großen Abstände zwischen den Mikrofonen gibt es keinen Aufnahmewinkel (vgl. A – B), da das Gesetz der ersten Wellenfront wirksam wird und z.B. Schallquellen von der linken Seite vor allem links, und Schallquellen von der rechten Seite vor allem rechts abgebildet werden. Siehe dazu auch [18]. Schallquellen im Bereich $\pm 30^\circ$ werden also im Center- Lautsprecher oder sehr nahe dazu wiedergegeben, und jene außerhalb von $\pm 50^\circ$ werden in, oder sehr nahe den Lautsprechern L bzw. R wiedergegeben.

Eine solche Anordnung mit großen Mikrofonabständen ist demnach nicht geeignet, um ein genaues stereofones Abbild zu generieren, sondern sie erzeugt ein offenes, räumliches Klangbild mit stabiler Mitte aufgrund des Center – Mikrofons (vgl. [37]). Diese Art von Konstellationen werden normalerweise ein paar Meter hinter und über dem Dirigenten aufgestellt und an die jeweiligen Verhältnisse angepasst.(vgl. [38]).

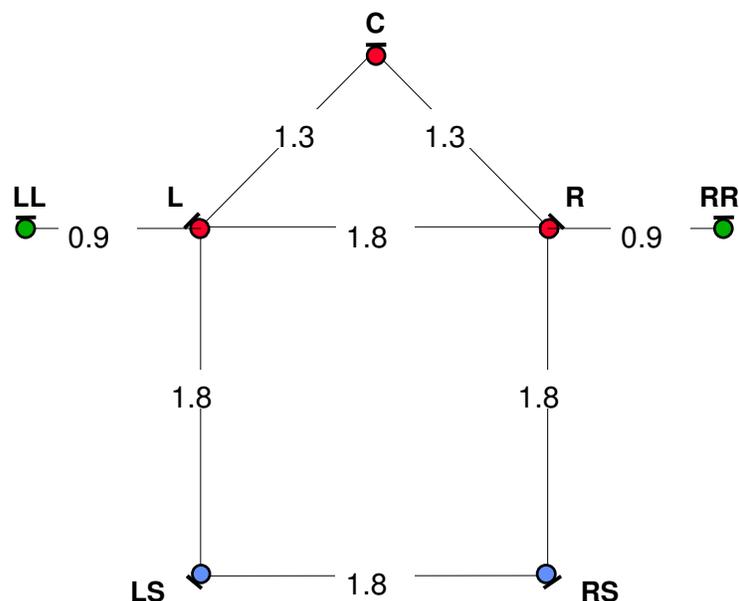


Abb. 32: "Fukuda-Tree" [50]

Die vordere Dreiecksanordnung ist eine abgewandelte Form vom „Decca-Tree“. Die Mikrofone L, C, R, RS, LS sind Mikrofone mit Nierencharakteristik. Zusätzlich seitlich flankierende Mikrofone mit Kugelcharakteristik werden verwendet, um einen Eindruck von der Breite des Orchesters zu geben und um eine Verbindung der Schallcharakteristika von vorne und hinten zu realisieren („...to present a sense of the orchestra width and to smooth the sound connection between the front and the rear...“ [50]). Die räumliche Darstellung wird mit Hilfe des Quadrates L, R, RS, LS realisiert. Weiter wird in [50] erwähnt, dass die jeweiligen Abstandswerte durchaus als variabel zu betrachten sind und an die aktuelle Situation angepasst werden können (Zitat: „The configuration of the tree can vary depending on the hall’s acoustic characteristics, while the microphone intervals may be changed conforming to the orchestra’s size and formation.“).

KFM 360 (vgl. [46])

Das KFM (Kugelflächenmikrofon) 360 nach Bruck (vgl. [47] und [48]) basiert auf einem Kugelflächenmikrofon, bei dem jeweils dicht unterhalb der in die Kugel eingebauten Druckempfänger ein Mikrofon mit Achtercharakteristik montiert ist. Diese Mikrofone sind mit ihrer Hauptachse nach vorne gerichtet. Durch MS – Matrizierung ($L = M + S$, $R = M - S$) bildet sich auf jeder Seite durch die Summe zwei virtuelle Mikrofone, welche Schalle, die von vorne einfallen, bevorzugt aufnehmen. Durch Differenzbildung ergibt sich die gleiche Richtwirkung nach hinten. Die Polardiagramme der beiden virtuellen Mikrofone hängen deshalb vom Pegelverhältnis der beiden Signale ab. Werden die vier Signale unmatriziert aufgenommen, so können die Pegelverhältnisse der Signale, welche mit Druckempfänger und Druckgradientenempfänger (mit Achtercharakteristik) aufgenommen wurden, bei der Nachbearbeitung eingestellt werden.

Im Folgenden werden prinzipielle Überlegungen präsentiert um die fünf Kanäle optimal zu nützen. Das Prinzip kann auf die klassischen Konzepte von „Hauptmikrofon / Stützmikrofon / Raummikrofon“ und „Polymikrofonie“ angewandt werden.

3.2.6.4 Dreidimensionale Raumabbildung / Räumlicher Höreindruck

Der spontan wahrgenommene räumliche Höreindruck, hervorgerufen durch tatsächlichen oder reproduzierten indirekten Schall eines Raumes, enthält zwei Attribute des Hörereignisses [34]. Das ist zum Einen der „Hall“, beschrieben durch ein zeitliches Verwischen des Hörereignisses verursacht durch späte Reflektionen, zum Anderen ist das „Räumlichkeit“, die durch eine charakteristische Verteilung der Hörereignisse, verursacht durch frühe (10...80 ms) seitliche Reflektionen, gekennzeichnet ist. Die optimale Verzögerung für diese seitlichen Reflektionen liegt im Bereich von 15 bis 25 ms.

Der frühe seitliche Schall verursacht eine interaurale Dekorrelation wie sie für einen speziellen Raum (und demnach eine spezielle Räumlichkeit) typisch ist. Die Abhängigkeit der Räumlichkeit von Verzögerungszeit, Pegel, Einfallswinkel und Spektrum der frühen Reflektionen war Thema diverser Untersuchungen [35]. Als zentralen Punkt erkannte man auch den Gesamt - Pegel von Direktschall und Reflektionen. In [39] wird davon ausgegangen, dass es das Muster der Raumreflektionen selbst ist, und nicht der resultierende niedere Pegel der interauralen Korrelation, der zur Beurteilung herangezogen wird.

Ergebnisse von Experimenten (vgl.[35]) zeigen, dass eine klare Abhängigkeit des Ausmaßes von Räumlichkeit vom Schalleinfallswinkel der seitlichen Reflektionen besteht. Diese Ergebnisse sind von praktischer Bedeutung, da sie zeigen, dass seitliche Reflektionen der effektivste Weg sind, um Räumlichkeit zu generieren. Demgegenüber sind frühe Reflektionen in der Medianebene eher nachteilig.

Es ist ein interessantes Phänomen, dass der natürliche Raumeindruck durch Lautsprecher, welche ausschließlich in der horizontalen Ebene angeordnet sind, erreicht werden kann. Die obere Hemisphäre ist subjektiv eingebunden aufgrund unserer Hörerfahrung („auditory association“ [53], Haas- oder Precedence-Effekt [34]). Der Ausdruck 3-dimensionale Raumabbildung reflektiert dieses Faktum.

Stereofone Wiedergabe

Bei der herkömmlichen Zwei-Kanal-Stereofonie muss der Raumeindruck notwendigerweise ausschließlich von den vorderen zwei Lautsprechern als zweidimensionale Perspektive geschaffen werden. Optimiert kann dies durch die Anwendung der Phänomene des räumlichen Hörens werden (z.B. durch die Einbindung der natürlichen Reflektionsmuster in

das stereofone Signal). Das prinzipielle Ergebnis wird aber ein „perspektives Bild in der Simulationsebene“ zwischen den zwei Lautsprechern sein.

Im Falle vom 3/2-Format kann die akustische Perspektive für den Zuhörer mit Hilfe der zusätzlichen Surround Lautsprecher besser dargestellt werden. Ein natürlicher räumlicher Eindruck kann durch die Abstrahlung der seitlichen Reflektionen durch die ebenfalls seitlich positionierten Lautsprecher erreicht werden.

Die psychoakustischen Grundlagen, welche der Wahrnehmung von zu Grunde liegen, können als Basis für die Erzeugung von Surround Sound hergenommen werden. Um einen natürlichen Schalleindruck zu erreichen, muss eine Ortung der Surround Lautsprecher vermieden werden. Das kann am Besten dadurch gewährleistet werden, indem die Surround Signale im Bezug zu den Front Signalen verzögert werden. Die verzögerten Surround Signale erfüllen dann die Funktion von echten Saal-Reflektionen und das Gesetz der Ersten Wellenfront (Richtungswahrnehmung und (Nach-) Verdeckung (!)) ist wirksam. Gleichzeitig sollen die im Raum verteilten Lautsprecher ein natürlich diffuses Schallfeld im Hörbereich erzeugen. Dadurch entsteht eine Imitation von frühen Reflektionen und Hall, und der Zuhörer nimmt die akustische Perspektive von den Frontlautsprechern wahr, und soll sich vom Hörereignis umhüllt fühlen.

Darüber hinaus können die seitlichen Reflektionen zu einer Wahrnehmung von räumlicher Tiefe führen. Schon früher wurde im Zusammenhang mit der sogenannten „Room-Related-Balancing-Technique“ ([41], [42]) herausgefunden, dass die Darstellung von Entfernung in der Simulationsebene erfolgreich sein kann, wenn nur jeweils eine linke und rechte seitliche Reflektion imitiert wird. Die Darstellung von räumlicher Tiefe kann also besser gelingen, wenn die seitlichen Reflektionen von den Surround Lautsprechern abgestrahlt werden und tatsächlich von seitlichen Richtungen beim Zuhörer eintreffen.

3.2.6.5 Räumliche Gestaltung mittels Delay

Die Darstellung von räumlichem Eindruck und Tiefe verlangt Berücksichtigung der Signalverzögerungen, - z.B. entsprechend den Prinzipien der „Room-Related-Balancing-Technique“ (siehe Abb. 33, Abb. 35, [41] und [42]). In Abb. 33 wird das Grundkonzept gezeigt, ausgehend von Fall: Orchester mit Haupt- und Stützmikrofonen. Das Prinzip ist natürlich für jede Art der Polymikrofonie gültig.

Im Fall von nichtverzögerten Mikrofonensignalen (Abb. 33, obere Situation) wird das Signal der Stützmikrofone früher wiedergegeben als das entsprechende Hauptmikrofon - Signal. Als Folge davon interpretiert man das Stützmikrofon - Signal als Direktschall und vorteilhafte Abbildungscharakteristika des Hauptmikrofons gehen verloren. Solche Aufnahmen klingen

unnatürlich, flach und ohne räumliche Tiefe. Der Grund dafür liegt im Haas - Effekt (Abb. 34). Die Lokalisation eines Hörereignisses wird im Prinzip durch den Schall bestimmt, welcher als erster beim Ohr eintrifft. Deswegen ist in obiger Situation auch die Charakteristik der Stützmikrone maßgebend für die Abbildungsqualität der Aufnahme.

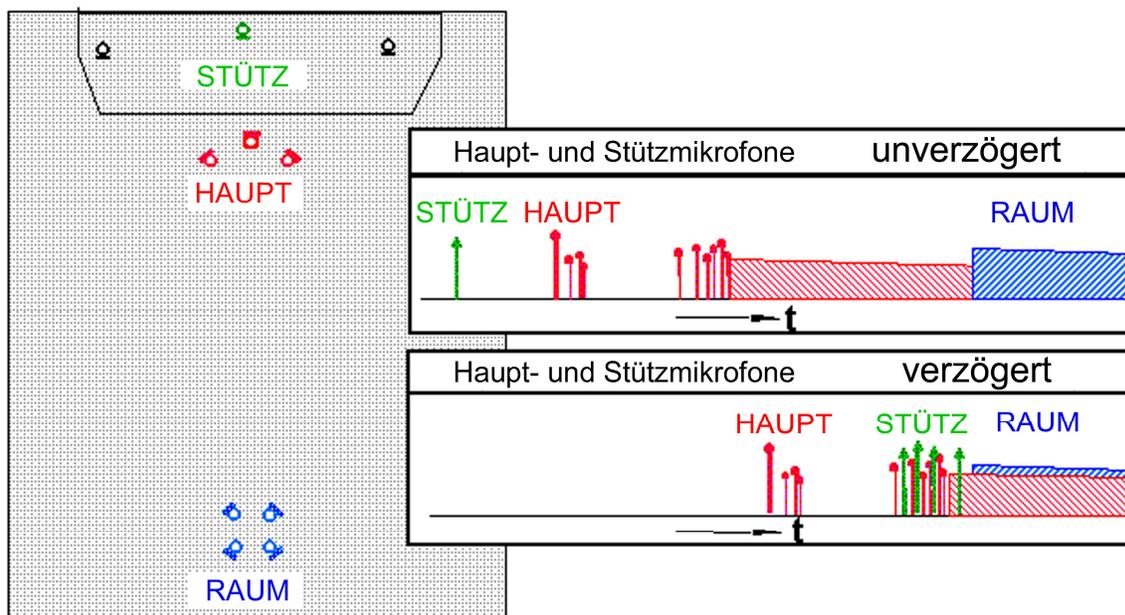


Abb. 33: Delay-Design entsprechend der natürlichen Raumantwort, vgl. [18]

Das Konzept von „Room-Related Balancing“ unterstützt die Natur des Raumeindrucks. Es wird empfohlen die Delay Situation in einer Art und Weise zu gestalten, wie sie der Originalsituation in einem Konzertsaal entspricht. Dieses Konzept ist mit einer beliebigen Mikrofonierung realisierbar.

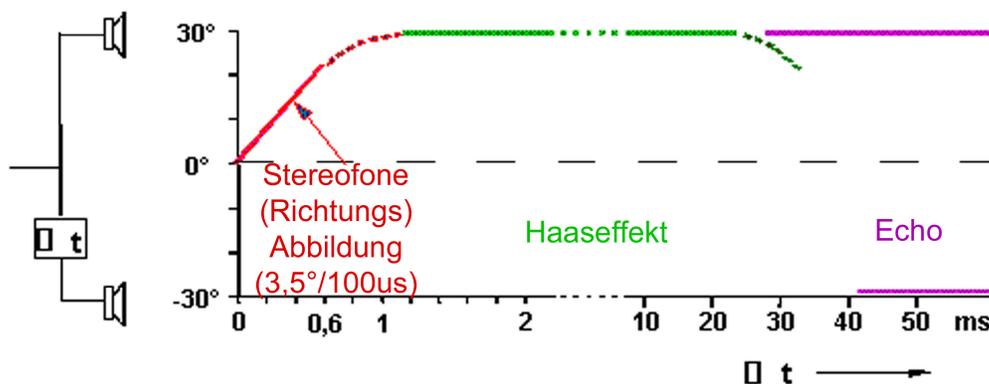


Abb. 34: Phantomquelle, Haas- bzw. Precedence-Effekt oder Echo, verursacht durch Kanalverzögerung, aus [18]

Es ist in der Praxis üblich, diesen den Raum störenden Effekt durch künstlichen Hall zu kompensieren oder die dem Hauptmikrofon entsprechende Verzögerung auszugleichen. Diese Technik ist jedoch nicht befriedigend, da reine Verzögerungskompensation zu Kammfiltereffekten führt, welche besonders störend auftreten, wenn sich die Musiker in der Nähe der Stützmikrofone bewegen. Um dies zu verhindern und die Wahrnehmung der räumlichen Tiefe durch das Hauptmikrofon zu erhalten, sollten die Stützen mehr verzögert werden, als es für die reine Kompensation nötig wäre, - so dass sie in den Bereich der frühen Reflektionen (15...25 ms) fallen.

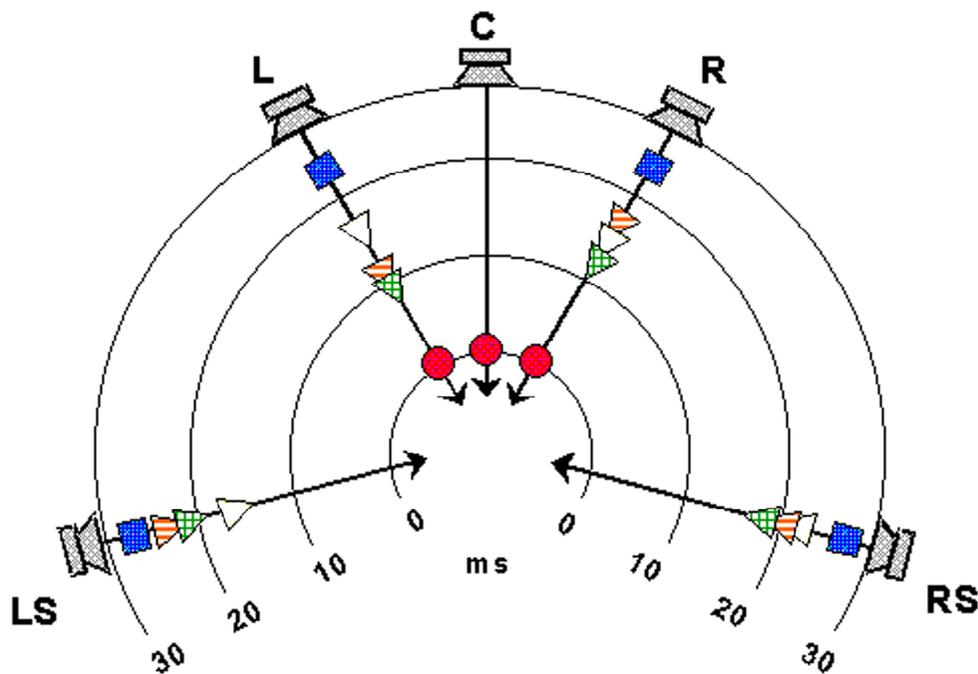
In einer umfassenden Studie von Konzertsälen [56] fand man heraus, dass in den überlegenen Sälen das Einsetzen des indirekten (impulsiven) Schalls ungefähr 20 ms nach dem Direktschall erfolgt. Diese „Arrival – Time - Gap“ [34], [55] gibt dem Zuhörer einen Eindruck der Räumlichkeit, ein Gefühl für die Größe des Raumes.

Die „Room-Related-Balancing-Technique“ bewirkt, dass das zeitliche Reflektionsmuster von Direktschall und ersten Reflektionen, welches ja im Hauptmikrofon enthalten ist, unverändert bleibt. Die Signale der Stützmikrofone tragen dennoch zur Klangregulierung bei.

Werden Raummikrofone verwendet, müssen Haupt- und Stützmikrofone in Bezug auf diese verzögert werden. Dies ist notwendig, um Echo - Effekte zu vermeiden, wenn der Abstand zwischen Haupt- und Raummikrofon größer als 10m (dem entsprechen ungefähr 30ms) ist.

Erfahrungsgemäß ist ein Aufnahmeplan inklusive Delay - Design für jedes verwendete Mikrofon von großer Bedeutung.

Die Grundlage eines Delay - Plans ist ein exakter Entwurf des darzustellenden Reflektionsmusters für den Referenzpunkt (sweet spot) der Hörzone, siehe Abb. 11. Für jeden Teil des Orchesters muss entschieden werden, welches Mikrofon für den Direktschall und die Richtungsabbildung verantwortlich ist. Dies legt wiederum den „Nullpunkt“ auf der Zeitachse fest, z.B. Nullpunkt beim Hauptmikrofon.



Mikrofonsignale:

● Haupt ▼ Stütze A ▽ Stütze B ▽ Stütze C ■ Raum

Abb. 35: Delay-Design entsprechend der natürlichen Raumantwort, aus [18]

Ein mögliches Reflektionsmuster ist hier dargestellt, um das beabsichtigte Misch – Resultat zu zeigen (die Zeitachse (ms) definiert die Schallankunftszeit für den Zuhörer). Die Ausbreitungsgeschwindigkeiten der wiedergegebenen Kanäle sind von jedem Lautsprecher zum Referenzpunkt (sweet spot) gleich. Abweichungen aufgrund extremer Hörpositionen oder asymmetrischer Lautsprecheraufstellung können in einem bestimmten Toleranzbereich (bis ± 3 ms oder ± 1 m) ignoriert werden. Das Zeitmuster des indirekten Schalls ist derart gestaltet, um den gewünschten räumlichen Eindruck und Wahrnehmung von Tiefe entsprechend (siehe Absatz Stereofone Darstellung) zu erreichen. In [18] wird vorgeschlagen den indirekten Schall der Stützen und Raummikrofone in die äußeren Kanäle (L, LS, RS, R) zu legen. Vorteilhaft ist es weiter wenigstens vier Reflektionen (je mehr desto besser, z.B. 8, 12) von jedem Stützmikrofon mittels Delay abzuleiten und ein 4-Kanal-Raummikrofon zu verwenden. Reflektionen im Mittenkanal (Medianebene des Zuhörers) sind unvorteilhaft [39], [41].

Prinzipiell sind mit 3/2-Stereo-Musikaufnahmen bessere Raumeindrücke als mit 2 – Kanal – Aufnahmen reproduzierbar. Optimale Ergebnisse sind sogar in schwierigen Situationen erreichbar, - vorausgesetzt dass jedes Mikrofon individuell verzögert werden kann. Um bei groß angelegter Stützmikrofonierung Verzögerungsglieder zu sparen ist es weiter möglich, Stützmikrofone welche annähernd gleichen Abstand zum Hauptmikrofon haben zu Gruppen A, B, C zusammenzufassen. Die Ableitung von mehr als vier Reflektionen von einem

Stützmikrofonsignal erweist sich als sinnvoll. Weiter kann die Qualität des Raumeindrucks durch adäquaten Nachhall für die Stützmikrofone verbessert werden.

Prinzipiell kann der „Room – Related - Balancing Algorithmus“ auch in Mischpulten implementiert werden und eine entsprechende Konsole wurde bereits in [8] eingeführt. Digitale Prozessoren erlauben weitere Optimierungen z.B. Entfernungs - Equalizer (Miteinbeziehung von spektralen Veränderungen aufgrund von Absorption während der Ausbreitung), zusätzliche Reflektionen pro Stützmikrofonsignal (realistischer Weise 12 ... 24), zusätzlich künstlicher Nachhall (von den Stützmikrofonsignalen in einer Linie mit den zusätzlichen Reflektionen erzeugt), „natural panning“ (Panoramisierungsgesetz entsprechend den interauralen Transferfunktionen des Kugelflächenmikrofons) [7], [8].

Es ist weiter möglich, die frühen Reflektionen mittels in Mischkonsolen integrierter Signalprozessoren zu realisieren. Das Hauptmikrofon würde dann überflüssig werden. Auch deshalb kann das Room – Related - Transfer Konzept ebenfalls auf Polymikrofonie angewandt werden.

3.2.6.6 Ästhetische Abwärtskompatibilität

Ein wichtiger Aspekt ist die ästhetische Abwärtskompatibilität von 3/2-Stereo-Musikaufnahmen. Die Downmix - Gleichungen entsprechend der Empfehlungen von ITU-R BS 775-1 sind in Tabelle 9 dargestellt. Es ist zwar möglich einen Downmix - Koeffizienten zu bestimmen, es ist damit aber nicht gleichzeitig gewährleistet, dass das resultierende zwei - Kanal - Ergebnis auch ästhetisch so gefällt (z.B. bezüglich Hall und Lautheits - Balance oder Tiefenwahrnehmung) , wie es die Original 3/2-Version tut. Idealerweise sollte der Downmix gewährleisten, dass die Aufnahme wie eine herkömmliche 2/0-Mischung klingt, z.B. aufgenommen mit dem gleichen Mikrofonset.

Was die Surround Informationen und das zeitliche Muster von Direktschall, frühen Reflektionen und Nachhall betrifft, ist das Downmix - Resultat korrekt (siehe auch Abb. 36). Das resultierende Reflektionsmuster entspricht dem gewünschten Muster für natürliche 2 - Kanal Aufnahmen (vgl. [41]). Von diesem Standpunkt aus gesehen ermöglicht der Downmix prinzipiell eine perspektivische Darstellung eines dreidimensionalen Raumes in der zweidimensionalen Simulationsebene zwischen den zwei Lautsprechern.

$L_0 = L + 0,7C + k LS$
$R_0 = L + 0,7C + k RS$
Zur Auswahl: $k = 1, 0.71, 0.50, 0.36, 0$

Tabelle 9: Kompatibilitätsmatrix 3/2 -> 2/0, aus [18]

Der Standard Downmixkoeffizient für die Surround-Kanäle ist $k = 0.71$, entsprechend der ITU-R BS 775-1. Natürlich ist es möglich auf der Produktionsseite alternative Koeffizienten zu definieren, welche dann via „Ancillary Data“ [44] oder „Meta Data“ [45] transportiert werden können um den entsprechenden Downmix für das jeweilige individuelle Programmmaterial zu gewährleisten. Der Koeffizient für den Center - Kanal ist mit 0.7 fixiert, da dieser Wert für alle Arten von Material, betrachtet man Richtungsabbildung und Lautheitsbalance, ziemlich gut funktioniert.

Es kann festgehalten werden, dass die Berücksichtigung der Zeitverhältnisse in beiden Fällen, 2 - Kanal Stereo- und Mehrkanal - Aufnahme, keinen Unterschied darstellt. Die entscheidende Änderung passiert nur bei der räumlichen Verteilung des indirekten Schalls.

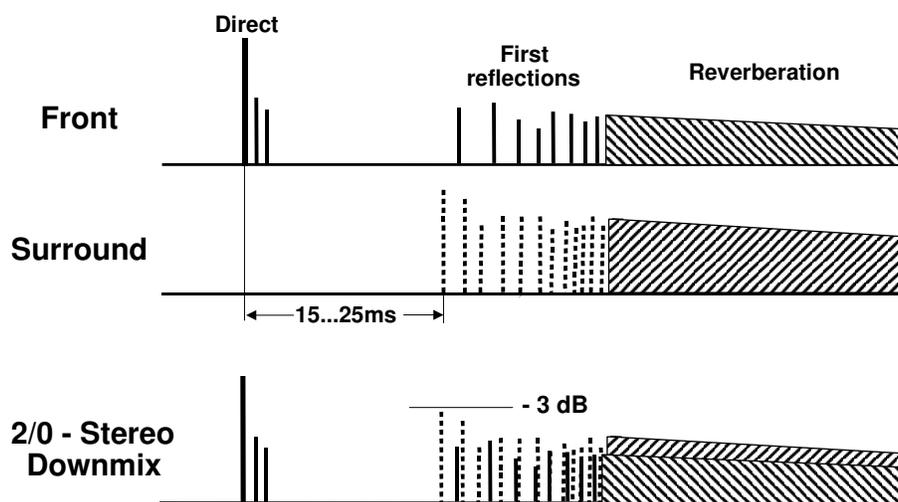


Abb. 36: Reflektionsmuster in einem 3/2- und in einem 2/0-Downmix, aus [18]

Im 2/0-Downmix ist die Rauminformation des 3/2-Stereo-Downmix zur Gänze erhalten geblieben. Zwei - Kanal Reproduktion erlaubt einen entsprechenden räumlichen Eindruck in der Simulationsebene. Dennoch ist nicht in allen Fällen eine optimale stereofone Qualität sichergestellt.

Es kann nicht allgemein vorausgesetzt werden, dass die ursprünglich beabsichtigte Hall – Balance bewahrt werden soll. Es ist ein allseits bekannter psychoakustischer Effekt beim Binauralen Hören, dass ein Raum weniger hallig klingt als beim Monauralen Hören [34], und ein ähnliches Phänomen tritt auf bei der Praxis von Musikaufnahmen wenn man von zwei-Kanal Stereo auf Mono oder von Mehrkanal auf zwei-Kanal Stereo umschaltet. Das könnte z.B. heißen, dass die Gesamt – Nachhall - Energie beim Downmix kleiner als in der Mehrkanaldarstellung sein soll. Das könnte man z.B. mit dem Surround Downmix - Koeffizienten $k = 0.7$ oder $k = 0.5$, abhängig vom Programmmaterial, erreichen.

Andere Erfahrungen werden in [49] und [50] berichtet. Dort wird erwähnt, dass jene Energie, welche Nachhall, bzw. räumliche Information enthält in Stereo und in Surround Produktionen gleich sein muss, wenn sie vom Zuhörer als natürlich wahrgenommen werden soll (“For reverberation component energy containing spatial information to be perceived as natural by the audience, the total sound levels in both stereo and surround – recording procedures must be equal.”). Das würde einen Surround - Koeffizienten von $k = 1$ bedeuten. Der Grund dafür könnte in den verschiedenen Dichte - Charakteren des halligen Schalls liegen, abhängig vom jeweiligen Surround – Mikrofon - Konzept.

Die Pegel - Angleichung des indirekten Schalls scheint eher eine Frage des ästhetischen Gefühles denn eine Frage einer allgemeinen Empfehlung zu sein. Gerade bei Überlegungen zur Abwärtskompatibilität trifft dies zu, da die Wahrnehmung einer akustischen Perspektive und eines räumlichen Eindrucks von einer Reihe von Parametern beeinflusst wird wie z.B. Dichte, zeitliche und richtungsabhängige Verteilung, Reflektionsenergie (R/D - Verhältnis [51]) usw.

Bei Betrachtung von Kopfhörer Wiedergabe muss festgehalten werden, dass ein einfacher Downmix entsprechend den Empfehlungen von ITU-R BS 775-1 (Tabelle 9) keine optimale Lösung darstellt. Die Im - Kopf - Lokalisation [34] stellt eine zu berücksichtigende Beeinträchtigung in Bezug auf Wahrnehmung von Raum und Tiefe, sogar, wenn man mit zwei-Kanal Lautsprecherwiedergabe vergleicht. Vergleicht man mit dem realen räumlichen Eindruck den man mit einer 3/2-Aufnahme erreichen kann, scheint der Mangel an ästhetischer Kompatibilität nicht akzeptabel. Es bedarf einer speziellen „Downmix“ - Methode für Mehrkanal - Kopfhörer Wiedergabe welche es ermöglicht, den räumlichen Eindruck zu bewahren welcher in einem Mehrkanal Hörraum möglich ist.

Ein dahingehender Versuch ist die Anwendung von Auralisationskonzepten um eine virtuelle Lautsprecherwiedergabe zu erreichen. Ein entsprechendes System wird in [14] vorgestellt. Es basiert auf binauralen Daten, welche in einem Mehrkanal Regieraum gemessen wurden.

Generiert wird ein binaurales Signal für Kopfhörerwiedergabe um das Hören in einem virtuellen Regieraum am Referenzpunkt zu ermöglichen. So werden Beeinträchtigungen des räumlichen Eindrucks, wie er mit Lautsprechern möglich ist, weitgehend vermieden.

3.2.7 WIEDERGABE- UND ÜBERTRAGUNGSFORMATE NACH DOLBY¹⁵

Bei der Aufstellung der Wiedergabeanlage schließt sich Dolby den Empfehlungen ITU – R [BS 775-1] an (siehe 3.2.3, Abb. 11). Zusätzlich gibt es eigens von Dolby entwickelte Standard – Raum – Layouts (siehe dazu Abb. 37).

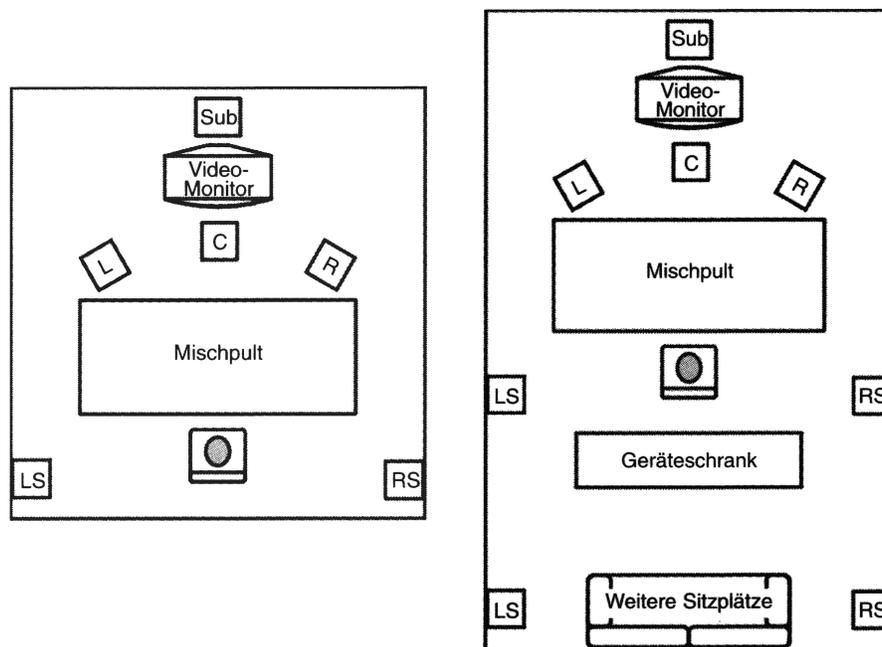


Abb. 37: Dolby Raum - Layouts für 5.1 bzw. 7.1 Wiedergabe, aus [60]

Bei der Einstellung der Lautsprecher empfiehlt Dolby dies mit rosa Rauschen, wiedergegeben mit -20dBfs pro Wiedergabekanal, durchzuführen. Mit einem Schalldruck – Messgerät mit langsamer Ansprechzeit und einer C – Bewertungskurve sollen die Verstärker so eingestellt werden, dass jeweils 85dB(C) gemessen werden. Für kleinere Wiedergaberäume wird für die Surround – Lautsprecher ein Pegel von 83dB(C) empfohlen. Der Subwoofer soll mit einem Terzbandanalyzer so eingestellt werden, dass der vom Subwoofer wiedergegebene Frequenzbereich 10dB über dem entsprechenden Bereich des Center - Lautsprechers liegt (wird hier nicht mit einem Terzbandanalyzer gearbeitet sondern mit einem Schalldruck – Messgerät über den gesamten Audio – Frequenzbereich, so sollte der Schalldruck des bandbegrenzten rosa Rauschen nach dieser Anzeige nicht um 10 , sondern um 6dB angehoben werden, also auf 91dB(C)) (siehe Abb. 38, vgl. [60]).

¹⁵ Dolby steht im Folgenden für Dolby Laboratories Inc.

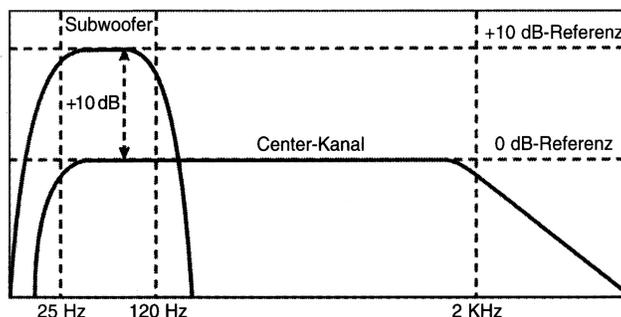


Abb. 38: Frequenzgang der Lautsprecher für 5.1 - kanaliges Abhören (vgl. [60])

Kodierformate

Dolby Surround ist eine Technik, bei der die Information der vier Eingangskanäle (L, R, C, S) am Dolby Surround Encoder durch einen Matrizierungsvorgang in zwei Kanäle (L_t , R_t) enkodiert wird (siehe Abb. 39). Bei dieser Matrizierung werden die Signale L und R ohne Veränderung den Signalen L_t und R_t zugewiesen. Das C – Signal erfährt eine Absenkung um 3dB und wird zwischen L_t und R_t aufgeteilt. Das S – Signal wird ebenfalls zwischen den beiden Kanälen aufgeteilt, wird aber davor noch drei Bearbeitungsschritten unterzogen:

- Bandbegrenzung von 100 Hz bis 7 kHz.
- Enkodierung mit einer modifizierten Form der Dolby (Typ – B) Rauschunterdrückung.
- Eine Phasendrehung von plus und minus 90° wird implementiert um eine 180° - Phasendrehung des aufgeteilten Signals zu erreichen, bevor es den Signalen L_t und R_t hinzugefügt wird.

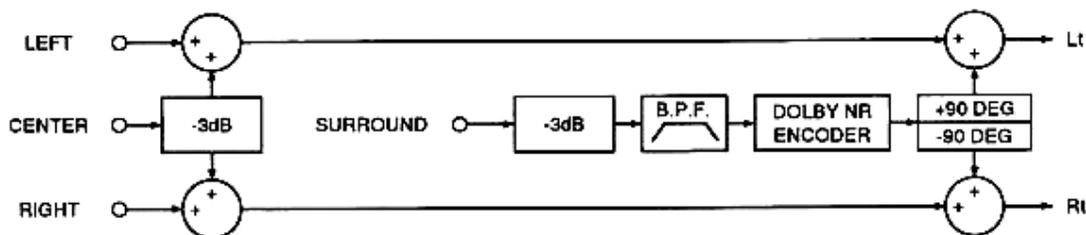


Abb. 39: Dolby Surround Encodierung (vgl. [61])

Die so generierten Signale L_t und R_t können mit jedem zwei – Kanal – Stereo Medium transportiert werden. Die Signale L_t und R_t sind kompatibel mit jeder Mono und zwei – Kanal – Stereo Wiedergabe. Bei Verwendung eines *Dolby Surround Pro Logic* – Systems, bei welchem aus den zwei Kanälen L_t und R_t eine vier – Kanal – Information L, R, C, S mittels aktivem Decoder (siehe [61]) wiedergewonnen wird, ist vier – Kanal – Surround

Sound Wiedergabe möglich. Dabei werden die Surround – Lautsprecher (vgl. Abb. 37) mit dem gleichen S - Signal versorgt.

Dolby Digital ist eine Technik, bei der das Audio Signal in den AC – 3 Code (Audio Coding, dritte Generation) übertragen wird. Die dabei stattfindende Datenreduktion basiert auf dem Modell der Vor- und Nachverdeckung und der Technik des ‚Noise – Shaping‘. Für die Enkodierung müssen die Eingangssignale im linearen PCM – Format vorliegen, verschiedene Abtastraten (empfohlen sind 48kHz, weiter sind möglich: 44.1kHz und 32kHz) und Wortbreiten (16,18 oder 20 Bit) sind möglich. Der AC – 3 Code ist für Mehrkanal – Anwendungen (von Mono bis zum Format 5.1) konzipiert, wobei die kodierten Kanalinformationen diskret verarbeitet werden (vgl. [60]). Das AC – 3 codierte Audiosignal kann mit jedem digitalen zwei – Kanal – Medium transportiert werden.

Der AC – 3 Datenstrom eignet sich für verschieden Wiedergabebedingungen (siehe Abb. 40).

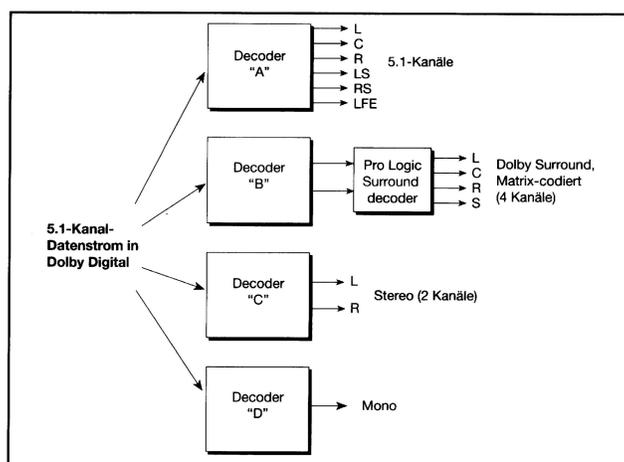


Abb. 40: Möglichkeiten der Wiedergabe von AC - 3 codiertem Audio – Signal, vgl. [60]

Der *Dolby Digital* (AC – 3) Code beinhaltet zusätzlich sogenannte ‚Metadaten‘, - Zusatzinformation, mit welcher vom Decoder das Downmix – Verhalten in Bezug auf Wiedergabe - Dynamik und Kompatibilität mit dem Wiedergabesystem gesteuert werden kann.

Dolby E wurde neben den oben genannten Surround - Formaten für den Konsumenten als professionelles Kodierformat, vor allem für die Produktion und den Austausch von DTV (Digital Television) Mehrkanal Audio, entwickelt. Acht Audio – Kanäle inkl. ‚Metadaten‘ (siehe *Dolby Digital*) können via einem AES/EBU - Paar transportiert werden. Als Datenrate wird eine Standard – Rate von 1,92Mbits/sec (20 – bit Audio bei 48kHz) verwendet. Video – Kompatibilität wird durch die Unterstützung einer ‚Frame – Rate‘ von 29.97fps gewährleistet. Künftig sollen Raten von 25fps und 24fps sowie Audio – Wortbreiten von 16bit und 18bit unterstützt werden (vgl.

[63]). Die acht Kanäle können beliebig verwendet werden, so z.B. für einen 5.1 – Mix und einen herkömmlichen zwei – Kanal – Mix. Im Gegensatz zum Format *Dolby Digital*, in welchem mit dem gleichen Audiomaterial nur ein Kodierungs- Dekodierungsdurchlauf möglich ist, gestattet es *Dolby E* diesen Zyklus (,ohne hörbarer Verschlechterung‘ (vgl. [62])) zehn mal zu durchlaufen.

3.3 WIEDERGABEKRITERIEN NACH THX¹⁶ (VGL. [77])

THX bezeichnet eine Zusammenstellung von technischen Standards für Filmaufführung, entwickelt vor allem für Kinos. Diese Standards beinhalten Spezifikationen in Bezug auf Equipment, Raumakustik, Hintergrund – Geräusch – Pegel und Projektions- und Sichtwinkel.

Entwickelt wurde THX 1982 von George Lucas und Tomlinson Holman um ein einheitliches Wiedergabekriterium zu schaffen. Die Entwicklungen von Dolby Laboratories in Bezug auf die Dekodierung von Audiomaterial, (vgl. 3.2.7), wurden eingebunden und ergänzt durch die Wiedergabekriterien für Filmtheater von THX.

THX gliedert die Signalkette in A - und B – Kette (siehe Abb. 41).

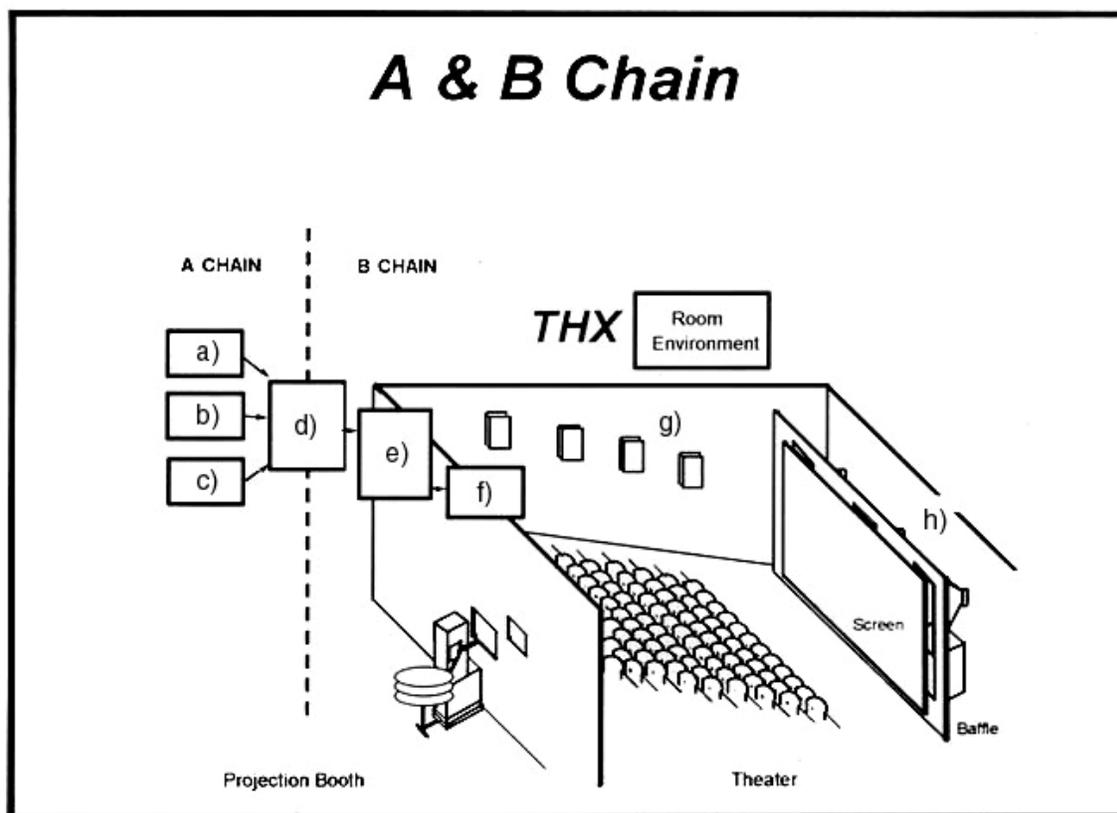


Abb. 41: A - und B - Signalkette nach THX, vgl. [77]. Den A – Teil bilden verschiedene Tonspuraufzeichnungsarten: Magnetisch, Optisch und Digital. Der B - Teil setzt sich aus individueller Raumakustik, spezieller Installationsmethode für Front - Lautsprecher, eigenem elektronischen Verteilernetzwerk und genauen Spezifikationen für Equipment und Wiedergabestandard zusammen. a)...Magnetic Soundtrack, b)...Optical Soundtrack, c)...Digital Soundtrack, d)...Cinema Processor, e)...THX Monitor, f)...Power Amplifiers, g)...Surround Speakers, h)...Screen Channels

¹⁶ THX steht im Folgenden für THX Lucasfilm Ltd.

Wird ein THX – Sound – System installiert, muss das System durch einen autorisierten Ingenieur auf diese Standards hin geprüft werden. THX legt Wert darauf, in die Planungsphase eines individuellen Systems eingebunden zu sein und gibt dann z.B. auch detailliert Auskunft über raumakustische Grenzwerte. Wird den Standards entsprochen, kann eine Zertifizierung erfolgen. Laufenden Prüfungen muss für eine jährliche Rezertifizierung entsprochen werden.

3.3.1 DIGITALE TONSPURFORMATE BEI THX

Zur Zeit existieren drei unterschiedliche digitale Tonspurformate für Film. Jedes dieser Formate wird durch entsprechende Hardware vom Hersteller unterstützt. Weiter stellt jedes dieser Tonspurformate mindestens sechs diskrete Kanäle (L, R, C, LFE (low frequency enhancement), LS (left surround) und RS (right surround)) (basierend auf den Wiedergabe - Spezifikationen von 3.2.3) zur Verfügung.

Dolby Digital (vgl. 3.2.7) encodierte digitale Datenblöcke werden zwischen den Transportlöcher von herkömmlichen 35 mm Film aufgezeichnet (siehe Abb. 42).

DTS (Digital Theatre Systems) encodierter digitaler Soundtrack wird auf einer (bis zu drei) CD – ROM's gespeichert und beim Abspielen mit dem Timecode des Films, welcher angrenzend an den Bildbereich von 35 mm Film aufgetragen ist (siehe Abb. 42), synchronisiert.

SDDS (Sony Dynamic Digital Sound) encodiertes digitales Audiomaterial (Datenreduktion basierend auf dem psychoakustischem Modell der Maskierung mittels ATRAC (Adaptive Transform Acoustic Coding, vgl.

[79])) wird an den zwei äußeren Rändern von 35 mm Film aufgezeichnet (siehe Abb. 42). Dieses System kann abhängig von der jeweiligen Mischung und der Lautsprecherkonfiguration hinter der Filmprojektionsfläche sechs oder acht Kanäle zur Verfügung stellen (vgl. [78]).

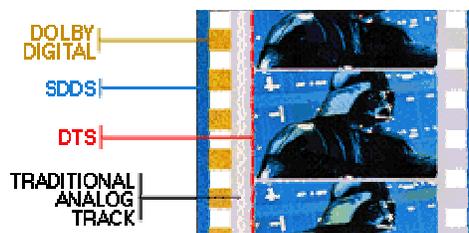


Abb. 42: Aufzeichnung verschiedener Tonspurformate auf 35 mm - Film bei THX, vgl. [77]

3.4 WIEDERGABEKONFIGURATION FÜR AMBISONIC

ALLGEMEINES

Ambisonic ist ein mehrkanaliges Aufnahme – und Wiedergabesystem und geht zurück auf GERZON M. und dessen Publikationen, beginnend mit den 70iger Jahren. Es bietet die Möglichkeit ein Schallfeld dreidimensional über Lautsprecher zu reproduzieren. Bei der Aufnahme wird das Schallfeld in einem Punkt mit Mikrofonen unterschiedlicher Richtcharakteristika aufgenommen. Die Lautsprecher signale werden bei der Wiedergabe in Abhängigkeit von den Lautsprecherpositionen entsprechend aus diesen Aufnahmesignalen zusammengesetzt [64]. Betrachtet man ein elementares System für die Wiedergabe in der horizontalen Ebene, so besteht dieses aus drei Kanälen. Jede ungerade Anzahl von Kanälen kann für dieses System verwendet werden, - für optimalen Einsatz des Systems ist die gleiche Anzahl an Lautsprechern erforderlich wie Kanäle verwendet werden. Im Folgenden sollen zweidimensionale Systeme im Speziellen erörtert werden.

3.4.1 THEORIE VON AMBISONIC

Ambisonic gestattet eine dreidimensionale Schallwiedergabe. Im Folgenden wird der zweidimensionale Fall betrachtet um die mathematische Komplexität in Grenzen zu halten. Die Anzahl der Kanäle zur Realisierung eines zweidimensionalen Ambisonic - Systems ergibt sich durch $2m+1$, wobei m den Grad der Ordnung¹⁷ darstellt. Würde man den dreidimensionalen Fall betrachten, ergibt sich hierbei die Anzahl der Kanäle durch $(m+1)^2$. Dadurch steigen die benötigten Ressourcen und auch die mathematische Komplexität steigt, weshalb hier zur Veranschaulichung der Prinzipien beim zweidimensionalen Fall geblieben wird.

Es ist anschaulich zur Darstellung und Analyse von Ambisonic ein Modell zu verwenden, welches auf dem Vergleich einer in das System hineingehenden Referenzwelle und einer vom System generierten herausgehenden Welle basiert. Es wird die Annahme getroffen, dass der Referenzschall eine ebene Welle ist, was für die Praxis solange gültig ist, solange die Schallquelle weit genug von der Hörposition entfernt ist. Weiter wird die Annahme getroffen, dass die von den Lautsprechern generierte ausgehende Schallquelle ebenfalls eine ebene Welle ist, was wiederum für eine entsprechende Entfernung des Zuhörers von den Lautsprechern erfüllt ist ($r > \lambda_{\min}/2$, vgl. [69], Kapitel 5.1).

¹⁷ siehe auch weiter unten

Es soll hier festgehalten werden, dass, obwohl die Bezeichnungen *hineingehend* und *herausgehend* für die Beschreibung der ursprünglichen ebenen Welle und der reproduzierten Welle verwendet werden, beide im Referenzpunkt konvergieren. Diese Bezeichnungen werden deshalb verwendet um darauf hinzuweisen, dass die ursprüngliche ebene Referenzwelle im Referenzpunkt ‚hineingeht‘ bzw. aufgenommen wird, jedoch der Versuch der Reproduktion der Welle von den Lautsprechern ausgeht.

Bei Ambisonic ist es üblich das herkömmliche kartesische Koordinatensystem (positive x – Achse vom Mittelpunkt nach rechts gerichtet) um 90° gegen den Uhrzeigersinn zu drehen. Demnach wird die x-Achse vom Mittelpunkt nach oben zeigen, die y-Achse vom Mittelpunkt nach links und für den dreidimensionalen Fall wird die z-Achse aus der Zeichenebene heraus zeigen. Der Ursprung des Koordinatensystems ist das Zentrum des Hörbereichs, der Referenzpunkt.

Angenommen die ebene Welle fällt mit einem Winkel ψ in Bezug zur x-Achse ein. Der hier betrachtete Aufpunkt befinde sich in einem Abstand r vom Ursprung entfernt mit dem Winkel ϕ , - wiederum in Bezug zur x-Achse (siehe dazu Abb. 43). Ist \vec{k} der Wellenvektor und \vec{r} der Vektor zum Aufpunkt, dann lässt sich die ebene Welle in dieser Situation als $P_\psi e^{i\vec{k}\cdot\vec{r}}$ schreiben. P_ψ ist die konstante Amplitude der ebenen Welle und k ist die Wellenzahl ($k = \frac{2\pi}{\lambda}$, wobei λ die Wellenlänge darstellt). \vec{k} beinhaltet zusätzlich die Richtungsinformation der sich ausbreitenden Welle, - in diesem Fall jene Richtung, welche mit der x-Achse den Winkel ψ bildet und zum Mittelpunkt orientiert ist. Umgeformt kann dieser Ausdruck als folgende Gleichung angeschrieben werden:

$$S_\psi = P_\psi e^{ikr \cos(\phi-\psi)} \quad (1)$$

wobei S_ψ die ursprüngliche ebene Welle kennzeichnet.

Der Anspruch des Ambisonic – Systems ist es, diese ebene Welle bis zu einer bestimmten Ordnung wiederherzustellen. Wie das versucht wird zu erreichen, wird im Folgenden dargestellt.

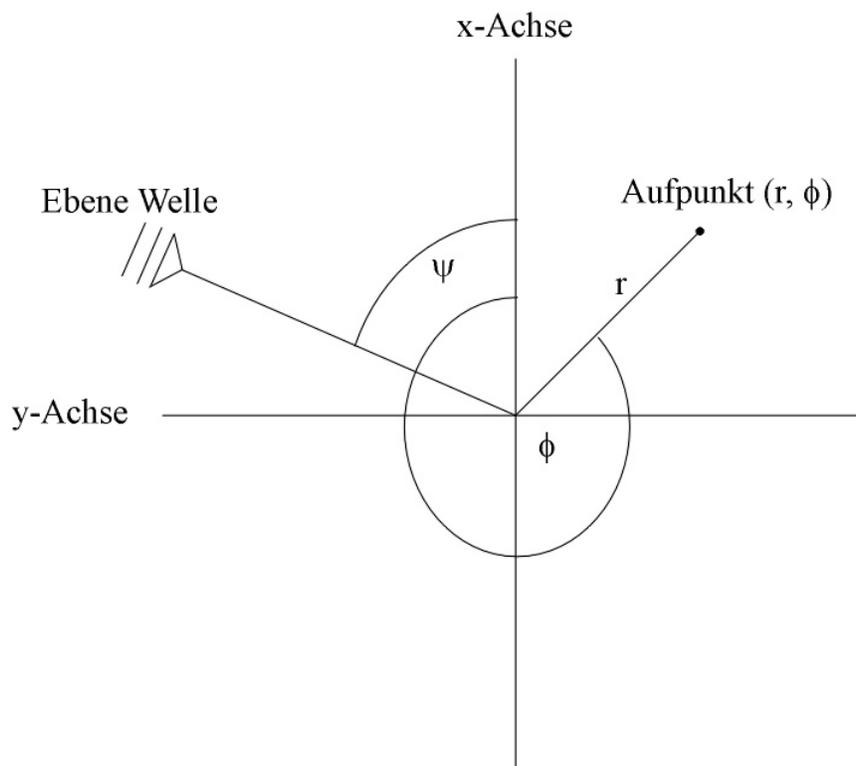


Abb. 43.: Geometrie der einfallenden ebenen Welle

Der Phasenterm der einfallenden Welle lässt sich in eine Reihe entwickeln, d.h. die einfallende Welle ist als Reihe darstellbar. Gelingt es, die Beschaffenheit der Lautsprechersignale derart zu gestalten, dass deren Summe den Gliedern der entwickelten Reihe entspricht (Koeffizientenvergleich, siehe Gleichungen (6) bis (8)), so ist die Reproduktion möglich.

In [65] gibt Gleichung 1.2.9 nun die Beziehung zwischen einer ebenen Welle und einer Reihe mit zylindrischen Bessel Funktionen an. Wenn die ebene Welle als Reihe ausgedrückt werden kann, dann besteht die Möglichkeit aus den niedrigen Termen der Reihe ein Signal generieren, welches als Quellsignal für ein Lautsprecher – System verwendet werden kann. Mit dieser Annahme kann folgendes festgehalten werden [65], [66]:

$$S_{\psi} = P_{\psi} J_0(kr) + 2P_{\psi} \sum_{m=1}^{\infty} i^m J_m(kr) \cos m(\phi - \psi) \quad (2)$$

Der Kosinus – Ausdruck kann erweitert werden, und es ergibt sich folgende Gleichung:

$$S_{\psi} = P_{\psi} \left(J_0(kr) + \sum_{m=1}^{\infty} 2i^m J_m(kr) [\cos(m\phi) \cos(m\psi) + \sin(m\phi) \sin(m\psi)] \right) \quad (3)$$

Aus obiger Gleichung ist ersichtlich, dass die ursprüngliche ebene Welle exakt reproduziert werden könnte, wenn eine unendliche Anzahl von Signalen und Kanälen vorhanden wäre.

Natürlich ist das praktisch nicht durchführbar, da nur eine endliche Anzahl von beidem zur Verfügung steht. Dennoch sei dies hier allgemein festgehalten.

Ein Ambisonic Lautsprecher – Array besteht aus N Lautsprechern, welche entlang einer Kreislinie symmetrisch angeordnet sind, wie z.B. in Abb. 44 wo N = 8 und der Winkel zwischen den Lautsprechern jeweils 45° beträgt.

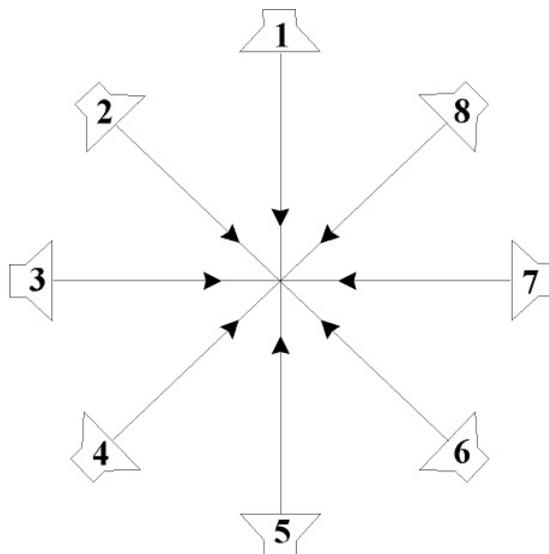


Abb. 44: Typische Ambisonic Lautsprecher – Aufstellung vgl [69]

Wenn nun angenommen werden darf, dass jeder Lautsprecher (von einem genügend großen Abstand aus betrachtet) eine ebene Welle produziert, dann kann der Ausgang des n - ten Lautsprecher folgendermaßen geschrieben werden:

$$S_n = P_n e^{i\vec{k}_n \cdot \vec{r}} = P_n e^{ikr \cos(\phi - \phi_n)} =$$

$$S_n = P_n \left(J_0(kr) + \sum_{m=1}^{\infty} 2i^m J_m(kr) [\cos(m\phi) \cos(m\phi_n) + \sin(m\phi) \sin(m\phi_n)] \right) \quad (4)$$

wobei ϕ_n für den Winkel des n - ten Lautsprecher steht. Jeder Lautsprecher trägt einen Teil zum Gesamtschallereignis bei¹⁸, und so ergibt die Superponierung der Ausgänge aller N Lautsprecher die wiederhergestellte gesamte Welle. Diese vom System erzeugte Welle sollte

¹⁸Wenn jeder Lautsprecher eine ebene Welle erzeugt, ist die Summe darüber nicht notwendigerweise ebenfalls wieder eine ebene Welle darstellt. Diese Summe ist ein kompliziertes Gemisch aus N ebenen Wellen, welches der Einfachheit halber hier als *gesamte Welle* oder als *wiederhergestellte Welle* bezeichnet werden soll.

gleich der einfallenden Referenzwelle sein. Hierbei müssen die Lautsprecher regelmäßig angeordnet sein¹⁹.

Führt man die Summenbildung durch ergibt sich folgender Ausdruck:

$$S = \sum_{n=1}^N P_n J_o(kr) + \sum_{m=1}^{\infty} 2i^m J_m(kr) \left(\sum_{n=1}^N P_n \cos(m\phi) \cos(m\phi_n) + \sum_{n=1}^N P_n \sin(m\phi) \sin(m\phi_n) \right) \quad (5)$$

Vergleicht man Gleichung (3) mit Gleichung (5) fällt auf, dass bestimmte Terme übereinstimmen. Das gewünschte Ergebnis ist die Übereinstimmung der wiederhergestellten und der ursprünglichen ebenen Welle. Dies ist der Fall, wenn folgende Bedingungen, welche sich durch Koeffizientenvergleich ergeben, erfüllt sind (matching conditions):

$$P_\psi = \sum_{n=1}^N P_n \quad (6)$$

$$P_\psi \cos(m\psi) = \sum_{n=1}^N P_n \cos(m\phi_n) \quad (7)$$

$$P_\psi \sin(m\psi) = \sum_{n=1}^N P_n \sin(m\phi_n) \quad (8)$$

Die Gleichungen (6) bis (8) repräsentieren also die Bedingungen welche erfüllt sein müssen, damit die wiederhergestellte Welle der ursprünglichen ebenen Welle entspricht. Jede Ordnung m entspricht der m -ten kugelförmigen Harmonischen im dreidimensionalen Fall. Stünde eine unendliche Anzahl an Kanälen zur Verfügung, könnte die wiederhergestellte Welle der ursprünglichen ebenen Welle zur Gänze entsprechen. Dazu sollte auch eine unendliche Anzahl an Lautsprechern vorhanden sein, welche den Kreis bildet. Die Reproduktion eines Schalles wäre vollkommen²⁰ indem vorzugsweise ein Lautsprecher (abhängig von ψ) für die Wiedergabe des Direktschalls verantwortlich ist, während die Reflektionen und Echo-Effekte der aufgenommenen Umgebung (z.B. des Aufnahmeraumes) von den anderen Lautsprechern hinzugefügt werden.

Natürlich ist weder eine unendliche Anzahl an Kanälen noch eine unendliche Lautsprecheranzahl realisierbar. Daraus folgt, dass nicht alle kugelförmigen Harmonischen erzeugt werden können und die Reihe abgebrochen werden muss.

¹⁹ D.h. gleicher Winkel zwischen den Lautsprechern und der Abstand aller Lautsprecher vom Mittelpunkt ist gleich.

²⁰ Alle vorangegangenen Annahmen und freies Schallfeld vorausgesetzt

Zur prinzipiellen Veranschaulichung genügt es Systeme erster und zweiter Ordnung zu betrachten. Ein System erster Ordnung zeichnet sich dadurch aus, dass die Terme bis zu $m = 1$ übereinstimmen, - bei einem System zweiter Ordnung ist dies bis $m = 2$ der Fall usw.

Die Bedingungen (6) bis (8) beschreiben einen Satz von Gleichungen welcher verwendet wird, um die entsprechenden Ambisonic Signale zu generieren. Um also eine Übereinstimmung nullter Ordnung für die ebene Welle zu erhalten, müssen einander nur die Glieder niedrigster Ordnung gleichen. Dies ist der Fall, wenn die Amplitude der beiden Wellen übereinstimmt, - ursprüngliche und wiederhergestellte Welle hätten also gleiche Amplitude. Gleichung (6) gibt die dahingehende Bedingung an, - das Signal wird in der für Ambisonic typischen Schreibweise durch W gekennzeichnet und beinhaltet keine Richtungsinformation. Die Übereinstimmung der nächsten, also der ersten Ordnung bedeutet, dass ursprüngliche und wiederhergestellte Welle die gleiche Orientierung aufweisen, - sie bewegen sich also auch in die gleiche Richtung²¹. Diese Bedingung wird in Gleichungen (7) und (8) für $m = 1$ erhalten. Es gilt $X = P_\psi \cos\psi$ und $Y = P_\psi \sin\psi$. Diese drei Signale bilden ein Ambisonic – Signal erster Ordnung²² und werden in der vorliegenden Form als B-Format Signale bezeichnet.

Ein horizontales Ambisonic – Signal zweiter Ordnung beinhaltet neben den oben angeführten Informationen noch zusätzlich Information bezüglich der Form der Welle (ähnlich der Krümmung). Diese zwei zusätzlichen Gleichungen werden aus Gleichungen (7) und (8) für $m = 2$ erhalten, $U = P_\psi \cos 2\psi$ und $V = P_\psi \sin 2\psi$. Ein System zweiter Ordnung besteht aus den folgenden fünf Signalen: W , X , Y , U und V . Beinhaltet ein System diese fünf Signale und verfügt es weiter über eine Lautsprecheraufstellung entsprechend Abb. 44, dann kann es eine ebene Welle im Referenzpunkt korrekt bis zur zweiten Ordnung wiedergeben.

Die Lautsprechersignale für einen beliebigen Lautsprecher in einem Ambisonic – System müssen näher betrachtet werden. Für die Wiedergabe einer ebenen Welle bis zu einer bestimmten Ordnung mittels Ambisonic – System müssen die Gleichungen (6) bis (8) erfüllt sein. Dennoch bleibt die Frage, wie das Versorgungssignal P_n für den n -ten Lautsprecher beschaffen sein soll. GERZON gab eine Beziehung für die Lautsprechersignale für ein Ambisonic – System erster Ordnung in [67] und [68] an und VANDERKOOY und LIPSHITZ verwendeten eine ähnlich in [66]. Der einzige Unterschied zwischen den beiden ist, dass VANDERKOOY und LIPSHITZ einen Faktor $\sqrt{2}$ nicht mit verwenden, welcher normalerweise bei *B – Format Ambisonic* für die Signale X und Y mit verwendet wird (vgl. [71]). Um also

²¹ Das heißt, die wiederhergestellte Welle wird so wahrgenommen, dass sie sich prinzipiell in die gleiche Richtung wie die ursprüngliche Welle bewegt, z.B. von in einem Raum vorne links über die Mitte nach hinten rechts.

²² Wenn von einem „Ambisonic – Signal erster Ordnung“ gesprochen wird, so ist gemeint, dass das Signal Informationen bis hin zur ersten Ordnung beinhaltet.

die folgenden Gleichungen für *B – Format Ambisonic* zu verwenden, muss der Faktor $\sqrt{2}$ bei den X - und Y – Signalen entfernt werden. Entsprechend VANDERKOOY und LIPSHITZ lautet das Versorgungssignal P_n für ein Ambisonic – System erster Ordnung mit N Lautsprechern folgendermaßen [67]:

$$P_n = \frac{1}{N}(W + 2X \cos \phi_n + 2Y \sin \phi_n) \quad (9)$$

Es kann gezeigt werden²³, dass Gleichung (9) die Bedingungen aus den Gleichungen (6) bis (8) erfüllt. Für P_n in einem System zweiter Ordnung werden untenstehend zwei Terme hinzugefügt, sodass die Erfüllung dieser Bedingungen weiterhin gewährleistet ist.

$$P_n = \frac{1}{N}(W + 2X \cos \phi_n + 2Y \sin \phi_n + 2U \cos 2\phi_n + 2V \sin 2\phi_n) \quad (10)$$

Mit diesen zwei Gleichungen (9) und (10) ist es nun möglich, Ambisonic – Systeme erster bzw. zweiter Ordnung zu realisieren.

Um die Qualität der reproduzierten ebenen Welle zu quantifizieren existieren Methoden, von denen zwei in [69] ausführlich erläutert werden.

3.4.2 KOMPATIBILITÄT DES SYSTEMS

Für die Etablierung eines Systems spielt die Kompatibilität mit bestehenden Komponenten im Hinblick auf Akzeptanz beim Konsumenten eine wesentliche Rolle. Ambisonic kann kompatibel gestaltet werden.

Ambisonic ist zuallererst selbst - kompatibel. Das heißt, eine beliebige Ordnung eines Systems mit einer beliebigen Anordnung und Anzahl der Lautsprecher ist kompatibel mit einer anderen Anordnung, welche eine unterschiedliche Anordnung und Anzahl der Lautsprecher aufweist. Das kann folgendermaßen verallgemeinert werden: Ein System mit N_1 Lautsprechern und der Ordnung m_1 ist kompatibel mit einem System mit N_2 Lautsprechern und der Ordnung m_2 , wobei N_i und m_i beliebige Interger – Zahlen sein können. Zusätzlich gilt folgende (anfangs erwähnte) Bedingung: $N_i \geq (2m_i + 1)$. Eine Aufnahme, welche für eine Anordnung von acht Lautsprechern und ein System zweiter Ordnung gemacht wurde, könnte ohne weiteres auf einem drei – Lautsprecher – System erster Ordnung wiedergegeben werden. Das funktioniert deshalb, weil es für jedes System beliebiger Ordnung bestimmte Regeln gibt, welche sich auf die Wiedergabe von Aufnahmen mit bestimmter Ordnung beziehen. Eine

²³ Alle vorangegangenen Annahmen und freies Schallfeld vorausgesetzt

Aufnahme erster Ordnung z.B. enthält die Signale *W*, *X* und *Y*. Ein Signal zweiter Ordnung besteht aus zwei weiteren hinzugefügten: *U* und *V*. Ein System erster Ordnung generiert seine Lautsprechersignale basierend auf den ersten drei Signalen. Stünden alle fünf Signale zur Verfügung, würden jene zwei zusätzlichen, welche die Information zur zweiten Ordnung beinhalten, einfach ignoriert. Die Wiedergabe wäre zwar nicht so gut wie eine komplette Wiedergabe zweiter Ordnung, aber sie ist möglich. Das heißt eine Quelle höherer Ordnung kann auf einem System niedrigerer Ordnung wiedergegeben werden.

Das Umgekehrte ist ebenfalls gültig: Wird eine Quelle erster Ordnung auf einem System zweiter Ordnung wiedergegeben, dann werden die Signale *U* und *V* einfach als Null aufgefasst. Sie werden in der gesamten Versorgung der jeweiligen Lautsprecher einfach nicht aufscheinen. Das System zweiter Ordnung wird jedem Lautsprecher ohne Probleme ein Signal erster Ordnung zur Verfügung stellen. Das Klangmaterial wird nicht anders klingen, als es das auf dem System erster Ordnung getan hat, welches z. B. zum Mischen verwendet wurde. Tatsächlich kann das Resultat sogar besser gefallen, da ein System zweiter Ordnung üblicherweise aus mehreren Lautsprechern besteht als ein System erster Ordnung. Ist das der Fall, so wird das System mit der höheren Anzahl an Lautsprechern das bessere Resultat bei der Wiedergabe erzielen.

Es kann also festgehalten werden, dass Programm – Material höherer Ordnung auf Systemen niedrigerer Ordnung mit einer Einbuße an Lokalisationsschärfe wiedergegeben werden kann – jedoch keine größere Einbuße, als wenn die ursprüngliche Aufnahme bereits in dieser Ordnung gemacht worden wäre. Eine Quelle niedrigerer Ordnung kann in dieser Hinsicht eventuell bessere Ergebnisse aufgrund der (üblicherweise) erhöhten Lautsprecher – Anzahl erzielen. Ambisonic ist in sich selbst aufwärts - und abwärts – kompatibel. Ein System kann so viel oder so wenig der zur Verfügung gestellten Information verwenden, wie gewünscht wird. Das implementiert auch die Tatsache, dass Ambisonic kompatibel mit einer Quelle nullter Ordnung, d.h. mit einer monofonen Aufnahme ist. Wird eine solche auf einem Ambisonic – System wiedergegeben, so wird jeder Lautsprecher einfach das selbe Versorgungssignal erhalten.

Es ist von geringer Bedeutung, mit welcher Ordnung Programm – Material aufgezeichnet wird, da verschiedene Ordnungen möglich sind. Große Kinos könnten z. B. Systeme zehnter Ordnung verwenden (wofür allerdings $(10+1)^2 = 121$ Lautsprecher notwendig wären), während die Versionen für die Heim – Verbraucher – Videos auf fünfte Ordnung heruntergemischt und auf einem System erster Ordnung wiedergegeben werden. Der Vorgang des Heruntermischens ist einfach: die Anteile der höheren Ordnung werden einfach nicht miteinbezogen.

Betrachtet man die Möglichkeit zur Durchsetzung des Systems, so könnte es nur in dieser Hinsicht problematisch sein, dass das durchschnittlich verfügbare Programm – Material in einer Ordnung aufgenommen ist, welche unter der Ordnung des durchschnittlich verfügbaren Systems liegt. Das Verständnis aus der Sicht der Konsumenten wäre wohl nicht allzu hoch.

Das Verhältnis dieser beiden Ordnungen muss natürlich aufeinander abgestimmt sein, - in Hinblick auf die zukünftige Entwicklung ist es sicherlich besser, die Aufnahmen im Zweifelsfall in einer höheren Ordnung zu machen.

Stereo

Zur Kompatibilität mit Stereo ist folgendes festzuhalten: Bei Ambisonic – Systemen mit nur zwei Lautsprechern in dem Bereich um $\pm 90^\circ$ um die x – Achse wird das linke Versorgungssignal für den linken Stereo – Kanal verwendet und das rechte Versorgungssignal für den rechten Stereo – Kanal. Besteht das System aus mehreren Lautsprechern, ist die Antwort nicht mehr so einfach.

Wie sollen z.B. die Lautsprechersignale für ein System mit vier Lautsprechern in den vorderen beiden Quadranten aus den linken und rechten Kanälen einer herkömmlichen Stereo – Quelle gewonnen werden? Hierbei werden die hinteren Lautsprecher nicht berücksichtigt. Da Zwei - Kanal - Stereo kein Surround System ist, wird auch kein Versuch unternommen Lautsprecher zu verwenden, welche sich in einem Bereich befinden, in welchem niemals eine Zwei – Kanal – Stereo Wiedergabe angestrebt werden würde. Daher werden also nur Lautsprecher in den vorderen zwei Quadranten ($\pm 90^\circ$) berücksichtigt. GERZON widmete sich diesem Thema in einer seiner vielen Veröffentlichungen [70]. Er entwickelte ein ganzes System von schrittweisen Aufwärts – und Abwärts – Konversions – Matrizen. Mit diesem System könnten beispielsweise die Signale von drei Front – Lautsprechern erweitert werden für fünf oder reduziert werden für zwei Kanäle. Für diese Fragestellung ist momentan also die Möglichkeit der Erweiterung eines zwei - Kanal – Systems interessant. Die Verwendung der GERZON – Matrizen um die zwei Stereo – Kanäle auf drei, vier, fünf oder mehrere Kanäle zu erweitern ist hier sicherlich von Vorteil²⁴. Die Auswirkungen auf die Lokalisierbarkeit einer solchen Erweiterung wurden mit den in [69] beschriebenen Methoden noch nicht getestet, es wird aber vermutet, dass Verbesserungspotential bezüglich der Größe der Hörfläche gegeben ist. In diesem Bereich ist weiterhin Arbeitsbedarf vorhanden.

Als Beispiel sei hier eine Enkodierungsform, die UHJ – Kodierung angeführt (vgl. [73]). Sie wird zur Erzeugung von zwei – Kanal – Stereo kompatiblen Ambisonic – Aufnahmen verwendet. Die Signale W , X und Y werden dazu in der folgenden Form matriziert:

²⁴ Eine Implementierung dieser Möglichkeit in den Ambisonic – Decoder würde auch dessen Wertigkeit erhöhen.

$$\text{Links} = (0.0928 + 0.255j)X + (0.4699 - 0.171j)W + (0.3277)Y$$

$$\text{Rechts} = (0.0928 - 0.255j)X + (0.4699 + 0.171j)W - (0.3277)Y$$

Tabelle 10: UHJ - Enkodierungs - Matrix, Zwei - Kanal - Stereo aus Ambisonic erster Ordnung, aus [73]

Die beste Referenz für UHJ – Dekodierung ist wiederum bei GERZON zu finden [74].

3.4.3 SYSTEMERWEITERUNG UND SPEICHERMEDIUM

Bei Ambisonic ist es möglich von einem Grundsystem auszugehen, und es später, Schritt für Schritt, auszubauen. Ein solches Grundsystem wäre z.B. eines erster Ordnung mit drei Lautsprechern. Das ist der erforderliche Minimumbedarf. Zu einem späteren Zeitpunkt können dann mehrere Lautsprecher hinzugefügt und das System für eine höhere Ordnung erweitert werden. Das erlaubt eine breite Flexibilität auf Seiten der Konsumenten. Das System ist also später komplett erweiterungsfähig und erlaubt es dennoch jegliches Material bereits im ‚Grundzustand‘ zu spielen.

Für die Speicherung des Programmmaterials sind, abhängig von der Ordnung des aufzuzeichnenden Materials, verschieden viele ‚Aufzeichnungskanäle‘ bzw. ‚Spuren‘ notwendig. Erste Ordnung horizontal benötigt drei, zweite Ordnung horizontal benötigt bereits fünf vollwertige²⁵ Audiokanäle. War also diese Technik noch bis vor Kurzem Anwendern, welche im Besitz von Mehrspurmaschinen waren vorbehalten, so eröffnen sich durch die Einführung der DVD neue Möglichkeiten. Die DVD – Audio besitzt laut den im März 1999 veröffentlichten Spezifikationen ([72],(Version 1.0)) bereits die Möglichkeit zur Speicherung von sechs diskreten Kanälen bei benutzerdefinierter Abtastfrequenz und Wortbreite. Dadurch steht der Aufzeichnung, Speicherung und Distribution von Programmmaterial zweiter Ordnung nichts mehr im Weg.

²⁵ D.h. mit einer verfügbaren Bandbreite von 20 Hz – 20 kHz.

3.4.4 AMBISONIC AUFNAHMEN

Aufnahmen für Material erster Ordnung sind im Moment relativ einfach zu machen. Es existieren, im Handel erhältliche, Mikrofone, welche am Ausgang vier Signale erster Ordnung anbieten (das zusätzliche Signal beinhaltet Höheninformation). Diese Signale müssen diskret aufgezeichnet werden und können dann über ein Ambisonic – System wiedergegeben werden. Aufnahme und Wiedergabe erster Ordnung sollte also problemlos funktionieren. Ein Ambisonic – Mikrofon erster Ordnung entspricht einer Anordnung von drei zueinander orthogonalen Mikrofonen mit Achtercharakteristik und einem Mikrofon mit Kugelcharakteristik. Die größte Schwierigkeit dabei liegt darin, dass diese Mikrofone möglichst koinzident angeordnet werden sollen, damit sie auch gewährleisten können das Schallfeld in einem Punkt aufzunehmen. Mit höherer Ordnung ist es momentan nicht möglich Mikrofone herzustellen, da mit steigender Ordnung auch die Komplexität der benötigten Richtcharakteristika der Mikrofone steigt.

Software zur Erzeugung von Studio -, bzw. Computer – generiertem Ambisonic – Material stellt z.B. der ‚Spatialisateur‘, welcher am IRCAM in Paris (vgl. [76]) entwickelt wurde, dar. Diese Software erlaubt die Abbildung von Schallquellen mittels verschiedenster Lautsprecher – Aufstellungen. Ein anderes Programm, VSpace, entwickelte FURSE [75] und stellt es auch unter [75] zum Herunterladen zur Verfügung. Dies seien nur einige wenige Hinweise um Tätigkeiten in diesem Bereich anzudeuten.

4 HÖRVERSUCH

4.1 ALLGEMEINES

Die erarbeitete Theorie aus der Literatur auch in die Praxis umzusetzen und zu vergleichen ist ein wesentlicher Bestandteil dieser Diplomarbeit. Zu diesem Zweck wurden zwei Mikrofonverfahren zur Aufnahme von Surround - Sound ausgesucht und damit Aufnahmen gemacht. Die Wiedergabe dieser Aufnahmen ist im Format 5.1 bzw. im Format 3/2 möglich. Ausschnitte der Aufnahmen wurden einander in einem Hörversuch gegenüber gestellt.

Neben den objektiven Kriterien wie Richtcharakteristika, Laufzeiten, Aufnahme- und Aufstellungswinkel bei den im Kapitel 3 beschriebenen Mikrofonkonstellationen und den zugehörigen Wiedergabekonfigurationen sind die subjektiven Urteile zur Bewertung eines bestimmten Systems unerlässlich. Die anatomischen Voraussetzungen und die Verarbeitung der Ohrsignale einer jeden einzelnen Versuchsperson bilden dabei die Unterschiede, aber auch die Grundlagen für die individuellen Bewertungen. In der Psychoakustik wird deshalb zwischen Schallereignis und Hörereignis, welchem erst eine individuelle Bewertung folgen kann, unterschieden. Zur Beschreibung eines Schallereignisses werden also physikalische Parametern herangezogen, wogegen für die subjektive Sinneswahrnehmung das Hörereignis entscheidend ist (vgl. [80]).

Bei der elektroakustischen Übertragungskette Mikrofon – Speichermedium – Lautsprecher steht am Ende stets der Hörer. Die Qualität der Kette wird also nicht ausschließlich durch technische Spezifikationen und physikalische Eigenschaften beschrieben, sondern zu einem großen Teil auch durch das Hörereignis. Da das Hörereignis keine physikalische Größe ist und sich in diesem Sinn nicht „messen“ lässt, ist man auf Aussagen der Hörer angewiesen. Der Hörer (die Versuchsperson) ist somit gleichzeitig „Messobjekt“ und „Messgerät“ [7].

Die Vorgangsweise zur Erstellung, Durchführung und Auswertung der Hörversuche gliedert sich in folgende Punkte:

- Auswahl der entsprechenden Mikrofonkonstellationen
- Auswahl und Aufnahme eines Ensembles
- Festlegung der Kriterien und Entwurf des Versuchsablaufs
- Durchführung der Versuchsreihe
- Auswertung der Ergebnisse

Im Folgenden werden diese einzelnen Punkte detailliert beschrieben.

4.2 AUSWAHL DER MIKROFONVERFAHREN

Um dem aktuellen Entwicklungsstand auf dem Gebiet der Surround – Aufnahmetechnik gerecht zu werden, musste zuerst eine Literaturrecherche bezüglich der momentan existierenden Verfahren durchgeführt werden. Wesentliche Beiträge dazu finden sich in den Tagungsbände der letzten Tonmeistertagungen und in den Veröffentlichungen der letzten AES (Audio – Engineering – Society) – Conventions. Sie (siehe Kapitel 5) bildeten eine gute Grundlage.

Die Entscheidung zugunsten der hier schwerpunktmäßig behandelten Systeme fiel durch die Teilnahme an der 108ten AES – Convention in Paris im Februar 2000. Dort präsentierten G. Theile vom IRT München (Institut für Rundfunktechnik) und M. Williams (Sounds of Scotland) mit G. Le Dû ihre neuen Erkenntnisse bezüglich Mikrofonverfahren für Aufnahmen des Formats 3/2. Während diesen Präsentationen und in den Diskussionen danach wurden einerseits Gemeinsamkeiten bezüglich des Verständnisses der Grundlagen (siehe 3.2.2, 3.2.6.4 bzw. 3.2.6.5), andererseits Unterschiede in den Realisierungen der Konfigurationen der beiden Wissenschaftler deutlich. Um den Unterschieden in der Praxis näher zu kommen, fiel die Entscheidung auf die Optimized – Cardioid – Triangle - (OCT) Anordnung nach Theile (siehe 3.2.6.1) und auf ein Multichannel – Microphone – Array (MMA) nach Williams (siehe 3.2.6.2).

Für den Prototypen der OCT – Anordnung, welcher von G. Theile in Zusammenarbeit mit der Fa. Schoeps GmbH entwickelt wurde, konnte vor Ort in Paris Interesse angemeldet werden. Es konnte eine Kooperation in der Form, dass die Fa. Schoeps GmbH den Prototypen zur Verfügung stellte, erwirkt werden.

Um die beiden Mikrofonkonstellationen zu vergleichen, war es notwendig bei der Aufnahme für die beiden Konstellationen ähnliche Bedingungen zu schaffen. Das war durch die Realisierung der hinteren beiden Kanäle der OCT – Anordnung gegeben. Sie wurden nach Rücksprache mit G. Theile ähnlich denen des Multichannel – Microphon - Arrays ausgelegt (siehe Seite 82, bzw. Abb. 45).

Auf das Vorhaben dieses Hörversuchs angesprochen, konnte M. Williams zwei seiner neu präsentierten Konstellationen (siehe [12], Figure 19 und Figure 22) empfehlen. Eine Version mit Aufnahmewinkel von 120° und eine zweite Version mit Aufnahmewinkel von 110° . Da der damals verfügbare Prototyp der OCT – Anordnung leider noch ein fixes Gestänge mit einer Breite von 70 cm und somit einen nicht veränderbaren Aufnahmewinkel von 100° aufwies, fiel die Entscheidung zugunsten der Williams – Version von 110° Grad (siehe Abb. 45). Damit waren entsprechend den Gegebenheiten auch bei diesem Parameter nicht zu große Unterschiede gewährleistet. Ob und inwiefern der verbleibende Unterschied der Aufnahmewinkel der beiden Konstellationen von 10° Auswirkungen zeigen würde, war erst durch die Auswertung der Versuchsreihe zu beantworten.

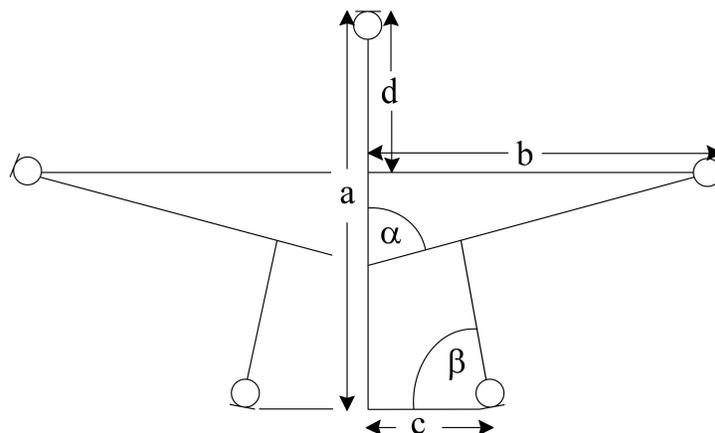


Abb. 45: Multichannel Microphone Array nach Williams, schematische Darstellung (vgl. [12], Figure 22).

Die zur Mittelachse symmetrische Anordnung von Mikrofonen mit Nierencharakteristik wird durch folgende Werte beschrieben: $a = 54\text{cm}$; $b = 50.05\text{cm}$; $c = 19,75\text{cm}$, $d = 22,81\text{cm}$; $\alpha = 75^\circ$; $\beta = 80^\circ$; linkes und rechtes hinteres Mikrofon müssen zusätzlich zur natürlichen Laufzeit bei der Nachbearbeitung noch um 1.02ms verzögert werden. Vorderer Aufnahmewinkel (FCA – Front Coverage Angel): 110° . Verwendete Mikrofone: 5 Stück AKG C 480 B mit den Aufsätzen CK 61 – ULS (Nierencharakteristik). Frequenzgang und Polardiagramm siehe **Anhang A**.

Für die OCT - Anordnung nach Theile (vgl. Abb. 15) wurden die hinteren Kanäle ähnlich wie in Abb. 45 realisiert. Die Werte dafür lauten (wieder mit Bezug auf das Center – Mikrofon): $a = 50\text{cm}$, $c = 25\text{cm}$, $\beta = 80^\circ$. Verwendete Mikrofone (zur Verfügung gestellt von der Fa. Schoeps GmbH) siehe Tabelle 11.

Position	Mikrofontyp
Center	CCM 4 L (N)
Links / Rechts Vorne	CCM 41 VL (SN)
Links / Rechts (tiefpassgefiltert bei 100Hz)	CCM 2 SL (K)
Links / Rechts Hinten	CCM 4 L (N)

Tabelle 11: Verwendete Mikrofone bei der OCT - Anordnung und den zugehörigen hinteren Kanälen. Buchstaben in Klammern bezeichnen die entsprechende Richtcharakteristik (N)...Niere, (SN)...Superniere, (K)...Kugel. Frequenzgänge und Polardiagramme siehe **Anhang A**.

4.3 AUFNAHMEN

Nach der Festlegung der zu testenden Mikrofonkonstellationen war es zuerst notwendig ein geeignetes Ensemble (mehrere verschiedene Instrumentengruppen sollen vertreten sein) und einen geeigneten Aufnahmeraum zu finden. Für die Aufnahme stellte sich das Ensemble „Tanguango“, aus Graz zur Verfügung. Das Ensemble setzt sich wie folgt zusammen: zwei 1. Geigen, zwei 2. Geigen, zwei Bratschen, ein Cello, ein Klavier und drei Akkordeons.

Als Aufnahmeraum wurde der Saal 301 in der Hofgasse 10 (Nebengebäude der Universität für Musik und darstellende Kunst Graz, Skizze der Aufnahme Raumes siehe Anhang B, Abb. 69) verwendet.

Um bei den Aufnahmen zu den Hörversuchen einen möglichst reibungslosen Ablauf zu garantieren, mussten die verwendeten Komponenten im Vorfeld erprobt werden. Die Handhabung des Multichannel Microphone Array, genauer dessen exakte Aufstellung entsprechend den berechneten Daten stellte sich als nicht sehr einfach heraus. Um die entsprechenden Winkel und Abstände der Mikrofone zueinander am Aufnahmeort einfach und schnell einzustellen, wurde vorher eine Schablone auf Karton im Maßstab 1:1 angefertigt. Diese Schablone wurde dann auf die Mikrofone samt Stative aufgelegt und das Array entsprechend eingerichtet. Bei der Aufstellung der OCT – Anordnung gab es dahingehend keine Schwierigkeiten, da der Prototyp mit einem fixen Gestänge, auf welchem die Mikrofone zu befestigen waren, geliefert wurde. Dadurch war aber auch der Aufnahmewinkel fixiert. Ein Exemplar aus einer eventuelle Serienfertigung würde laut J. Wuttke [82] in der Breite verstellbar sein.

Die Aufstellung der Instrumentengruppen und der Mikrofonkonstellationen in Relation zum Aufnahme Raum ist in Anhang B skizziert. Der Aufstellungsort der Mikrofonkonstellationen erfolgte gemäß den gegebenen Aufnahmewinkeln (OCT: 100°, MMA: 110°).

Vor der Aufnahme des Programmmaterials wurden mit einfachen Mitteln verschiedene Testsignale zur Kontrolle und Orientierung bei der Nachbearbeitung aufgezeichnet. Dies waren z.B. Sprache während des Abgehens des Aufnahmewinkels, Impulse in Form von Klatschen bei den einzelnen Pulten und markanten Positionen (Mitte, linker / rechter Rand) und direktes Ansprechen der einzelnen Mikrofone.

Zur Aufnahme wurden neben den oben beschriebenen Mikrofonen folgendes Equipment verwendet:

Mischpult: Mackie 16-8-2 (EQ neutral, alle Kanäle gleich eingestellt)

Mehrspurmaschine: 2 SONY Digital Audio Recorder PCM 800

4.4 ERSTELLEN DER VERSUCHSREIHE

Bevor die Versuchsreihe zusammengestellt werden konnte, musste die Anzahl der Fragen und den dazugehörigen Hörbeispielen festgelegt werden. Damit keine Ermüdungserscheinung bei den Versuchspersonen auftritt, wurde die maximale Dauer eines Versuches mit ca. 25min festgesetzt.

Bei der Auswahl der Fragen stellten Versuche mit ähnlichem Inhalt, welche in etwa zeitgleich an der FH – Düsseldorf von Rieseberg und Reiss durchgeführt wurden, eine Orientierungshilfe dar (vgl. [83], Arbeit nach Fertigstellung in der Bibliothek der FH Düsseldorf verfügbar). In online - Diskussionen mit Theoretikern und Praktikern im elektronischen Diskussionsforum „surround-sound“ kristallisierten sich relevante Parameter heraus (vgl. [81]). Nachdem die Fragen festgelegt waren (Fragebogen des Hörversuchs siehe Anhang B), wurde das Programmmaterial ausgewählt. Die Dauer der einzelnen Hörbeispiele ergab sich aus der Anzahl der notwendigen Vergleiche und der maximalen Versuchsdauer. Sie beträgt, abhängig vom Programmmaterial, ca. 10 – 15s. Bei der Festlegung der Dauer der Hörbeispiele wurde, soweit möglich, auch Bedacht auf die Musikalität der einzelnen Passagen gelegt. Die absolute Reihenfolge der einzelnen Beispiele und somit die Zuordnung zu den jeweiligen Fragen im bereits feststehenden Fragebogen wurde durch Ziehung ermittelt. Welche der beiden Konstellationen pro Gegenüberstellung am Anfang stand, wurde durch einen Zufallsgenerator bestimmt.

Editiert und gestaltet wurde die Versuchsreihe auf PC – Basis mit dem Harddisk – Recording – System SAMPLITUDE 2496 (Version 5.55).

Zur Einspielung in das Harddisk – Recording – System durchlief das Programmmaterial folgenden Signalweg (siehe Abb. 46)

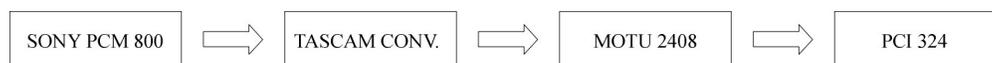


Abb. 46: Blockdiagramm: Signalweg von der Mehrspurmaschine in das HD-Recording – System. Das AES/EBU – Signal von der Sony PCM 800 wird im Tascam – Converter auf das Format TDIF umgewandelt. Damit kann es von der 19 – Zoll – Einheit MOTU 2408 der PCI – Karte 324 (welche die Schnittstelle zum PC darstellt) zur Verfügung gestellt werden.

Im HD – Recording – System erfolgte eine Feinabstimmung der Pegelverhältnisse der aufgezeichneten Signale. Mit Hilfe des Direktschall – Anteiles von aufgezeichneten Testschallen und der Richtcharakteristika der Mikrofone, mit denen das jeweilige Signal aufgenommen wurde, erfolgte eine Einpegelung. Sie wurde entsprechend dem Einfallswinkel des Direktschalles auf das jeweilige Mikrofon vorgenommen. So weisen z.B. die Signale der Mikrofone L und R des MMA bei Schalleinfall von 0° auf das Array (in Richtung der Hauptachse des C – Mikrofons) eine Absenkung des Direktschalls um 3dB bei 1kHz auf (Die

Pegelanzeige erfolgte über das Pegel- und Korrelationsgrad – Anzeigegerät DK – AUDIO MSD 550).

Bei der Erstellung der Versuchsreihe wurden vor Beginn der Beispiele zu den einzelnen Fragen zwei Beispiele als Eingewöhnungsphase vorangestellt. Danach folgten, nach 20s Pause, die Gegenüberstellungen zu den einzelnen Fragen. Nach den Gegenüberstellungen zu einer Frage standen, abhängig vom Programmmaterial, zwischen 17s und 20s als Vorbereitungszeit für die nächste Frage zur Verfügung (siehe Anhang B, Zusatzinformation für die Versuchspersonen).

Nach Beendigung des Gestaltens des Ablaufes der Versuchsreihe in SAMPLITUDE 2496 wurde das Material wieder zurück in die SONY PCM 800 (entsprechend Abb. 46 in umgekehrter Richtung) auf das Format HI 8 gespielt. Vor die Versuchsreihe wurden Einmeß- und Testaufzeichnungen zur Einstellung von Wiedergabeanlagen für diskrete 3/2 – Mehrkanal – Stereophonie entsprechend SSF – 02 auf das Band aufgezeichnet (siehe 7.3 bzw. vgl. [85]).

4.5 DURCHFÜHRUNG DES HÖRVERSUCHS

4.5.1 DER ABHÖRRaum

Der Hörversuch wurde im Regieraum des Tonstudios am Institut für Nachrichtentechnik und Wellenausbreitung durchgeführt. Dieser Regieraum war in [86] Gegenstand von raumakustischen Messungen. Diesen Messungen wurden die Werte für die Bestimmung der Nachhallzeit entnommen (siehe Abb. 47). Aus diesen Werten wurde ein arithmetischer Mittelwert der gemessenen Nachhallzeiten, der Terzbänder von 200 Hz bis 2,5 kHz, von $T_m = 0.34s$. errechnet. Dieser Wert entspricht den Empfehlungen des SSF – 02 (vgl. 3.2.3 und [85]). Die Schwankungen der Nachhallzeitwerte in den Terzbändern, in dem oben genannten Bereich, bewegen sich im Wesentlichen (bis auf die Werte bei 250 Hz, 500 Hz und oberhalb von 6,3 kHz) ebenfalls innerhalb der Toleranzgrenzen der Empfehlungen.

Die Aufstellung und Einmessung der Lautsprecher erfolgte entsprechend den Empfehlungen SSF – 01 [84] und SSF – 02 [85]. Der Abhörpegel L_{List} der gesamten Wiedergabeanlage wurde gemäß den Empfehlungen entsprechend dem Referenzabhörpegel $L_{List Ref}$ auf 85dBA eingestellt. Dem entspricht ein Pegel pro Wiedergabekanal von 78dBA ($L_{List Ref} = 85 - 10 \log(n)$, mit $n =$ Anzahl der Wiedergabekanäle (vgl. [85]). Messsignal war rosa Rauschen (200Hz bis 20kHz). Gemessen wurde mit dem Schallpegelmesser der Fa. Bruel und Kjaer, Typ 2232 nach IEC 651, Typ 1.

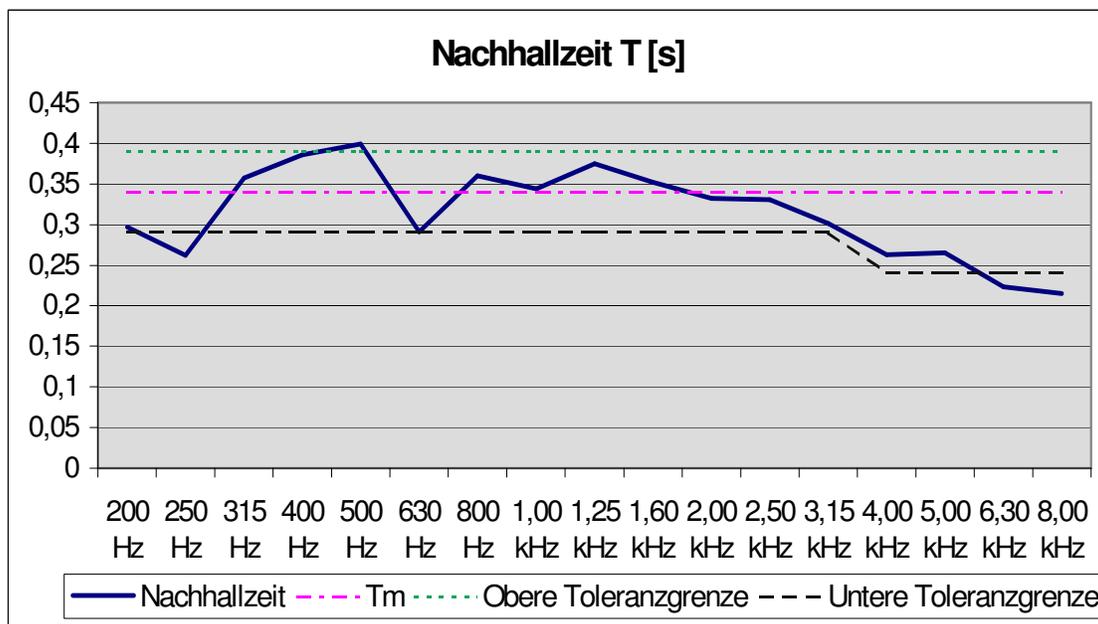


Abb. 47: Nachhallzeit im Abhörraum Nachhallzeit, T_m...Mittlere Nachhallzeit, Obere und Untere Toleranzgrenze (vgl. [84], Untere Toleranzgrenze 'springt' bei 4,00 kHz von (T_m – 0,05)s auf (T_m – 0,1)s).

4.5.2 VERSUCHSPERSONEN

Die Auswahl der Versuchspersonen (VPN), welche am Hörversuch teilnahmen erfolgte gezielt. Es wurden einerseits Personen herangezogen, welche den Studiengang „Elektrotechnik – Toningenieur“ absolvieren. Andererseits nahmen Personen teil, welche alle im Bereich der Akustik tätig sind. Der Einsatz von akustisch nicht versierten Personen wäre neben eventueller Unkenntnis der verwendeten Begriffe möglicherweise auch in Hinblick auf die Versuchssituation problematisch gewesen. Da manche Teilnehmer im Unterbewusstsein glauben könnten, an einer Prüfung teilzunehmen, in der man richtig oder falsch urteilen kann würde diese Unsicherheit das Ergebnis insofern verfälschen, dass nicht mehr das angekreuzt wird, was gehört wurde, sondern das, was geglaubt wird, „richtig“ zu sein (vgl. [80]). Deshalb wurden die Versuchspersonen beim Hörversuch darauf hingewiesen, dass auf ihre persönliche Beurteilung Wert gelegt wird.

Die Anzahl der Versuchspersonen wurde auf 25 festgelegt, um ein repräsentatives Ergebnis zu erhalten.

4.5.3 FRAGESTELLUNG UND VERSUCHSABLAUF

Voraussetzung für einen effektiven Hörversuch ist eine konkrete Vorstellung davon, was mit dem Versuch herausgefunden werden soll. Daraufhin müssen die Fragestellungen an die Versuchsteilnehmer ausgesucht werden. Es ist einerseits sehr wichtig, dass die Fragestellungen zur Zielsetzung passen, andererseits sollen die Versuchspersonen die gestellten Fragen relativ einfach beantworten können ([80]). Deshalb wurden Blätter mit Zusatzinformation für die Versuchspersonen vor dem Versuch zur Einsicht ausgehändigt, um den Ablauf des Versuchs zu erläutern (siehe Anhang B). Es wurde jedoch darauf Wert gelegt, nur jene Information zu geben, die unbedingt benötigt wurde.

Die acht zu beantwortenden Fragen lauteten folgendermaßen (der vollständige Fragebogen befindet sich im Anhang B):

Welches der Beispiele aus den folgenden Gegenüberstellungen ...

1. ...**gefällt Ihnen besser?**
2. ...weist eine **bessere Lokalisierbarkeit** einzelner Instrumente auf?
3. ...vermittelt eine **größere Breite** des aufgenommenen Ensembles?
4. ...vermittelt eine **größere Tiefe (Tiefenstaffelung)**?
5. ...**wirkt lauter?**
6. ...**lässt Proportionen einzelner Instrumente zueinander besser erkennen?**
7. ...**lenkt die Aufmerksamkeit eher auf die Klangfarbe?**
8. ...**weckt Ihr Interesse über längere Zeit und warum?**

Die Versuchspersonen nahmen einzeln am Referenzpunkt der Wiedergabeanlage Platz und bekamen dann die gestaltete Versuchsreihe einmal vorgespielt. Sie hatten die Aufgabe, pro Frage und Gegenüberstellung die Mikrofonkonstellation ihrer Wahl anzukreuzen. Die Zuspieldung der Versuchsreihe erfolgte aus dem Nebenraum per SONY PCM 800. Dies gewährleistete einerseits minimale zusätzliche Geräuschentwicklung im Abhörraum und andererseits gleiche Bedingungen in Bezug auf den Ablauf der Versuchsreihe für alle Versuchspersonen. Jede Versuchsperson konnte demnach die Beispiele einmal hören und musste danach jeweils entscheiden.

Nach der Beantwortung dieser acht hatten die Versuchspersonen noch Angaben zu ihrer Person auf dem letzten Blatt des Fragebogens zu machen.

Aus Zeitgründen wurde nur ein Teil der Versuchspersonen (15 VPN) nach Ende ihres Versuchsablaufes ersucht, die ersten beiden Fragen noch einmal zu beantworten.

4.5.4 AUSWERTUNG DER VERSUCHSERGEBNISSE

Die aus der Versuchsreihe erhaltenen Daten wurden einer statistischen Auswertung unterzogen. Die Ergebnisse daraus lassen eine vergleichende Aussage über spezielle Parameter der zwei Mikrofonkonstellationen zu.

Von den 25 teilnehmenden Versuchspersonen (VPN) gaben 21 VPN vollständig ausgefüllte Fragebögen ab. Zu Beginn wurden aus diesen 21 Fragebögen die Häufigkeiten der Nennungen OCT, bzw. MMA pro Gegenüberstellung ermittelt. Aus diesen Häufigkeiten wurden danach für die Fragen 1 bis 7 die Medianwerte zu jeder Frage ermittelt. Siehe dazu Abb. 48 bis Abb. 51.

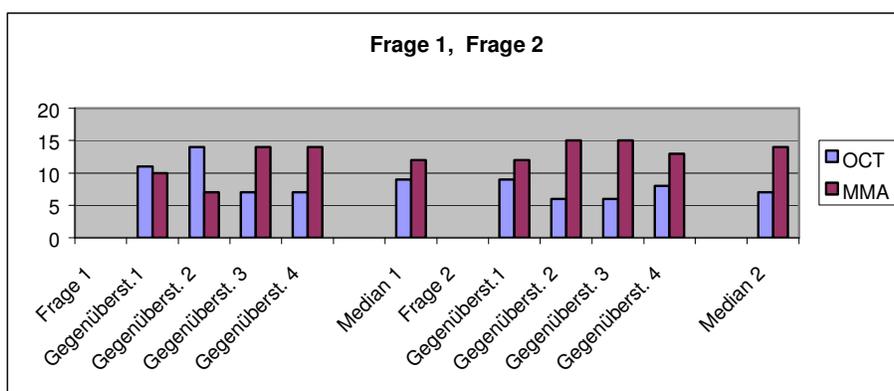


Abb. 48: Häufigkeiten der Nennungen OCT, bzw. MMA zu Frage 1 (...**gefällt Ihnen besser?**) und Frage 2 (...weist eine **bessere Lokalisierbarkeit** einzelner Instrumente auf?). Medianwerte über alle vier Gegenüberstellungen zu den Fragen.

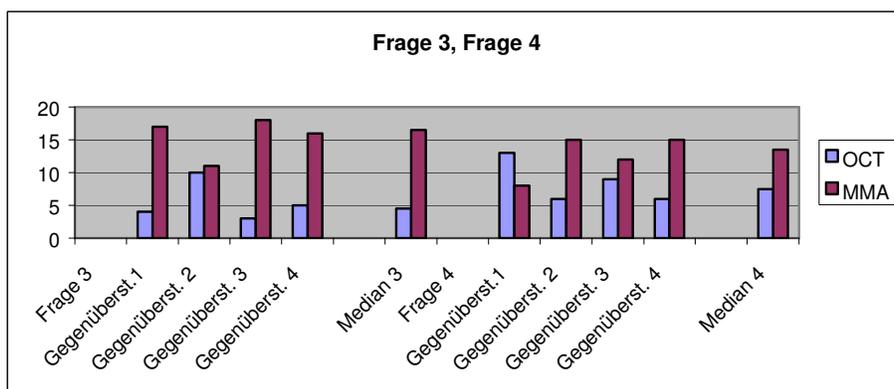


Abb. 49: Häufigkeiten der Nennungen OCT, bzw. MMA zu Frage 3 (...vermittelt eine **größere Breite** des aufgenommenen Ensembles?) und Frage 4 (...vermittelt eine **größere Tiefe (Tiefenstaffelung)**?). Medianwerte über alle vier Gegenüberstellungen zu den Fragen.

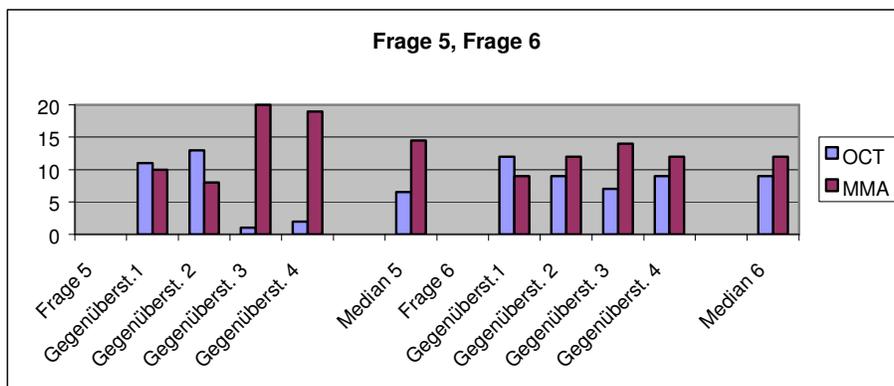


Abb. 50: Häufigkeiten der Nennungen OCT, bzw. MMA zu Frage 5 (...**wirkt lauter?**) und Frage 6 (...**läßt Proportionen einzelner Instrumente zueinander besser erkennen?**). Medianwerte über alle vier Gegenüberstellungen zu den Fragen.

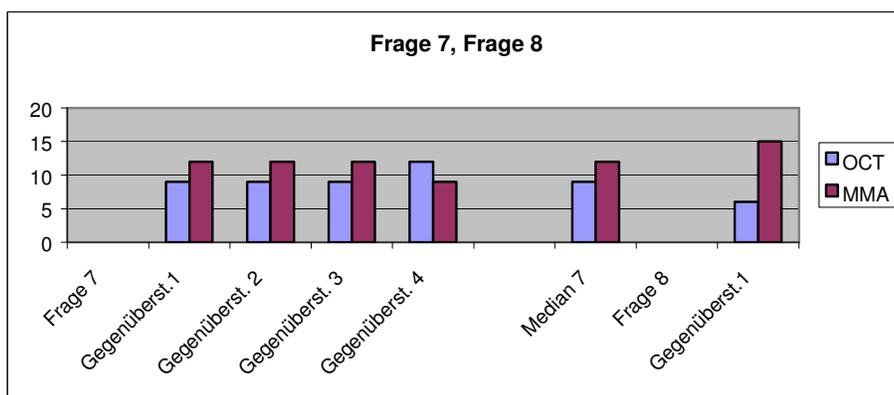


Abb. 51: Häufigkeiten der Nennungen OCT, bzw. MMA zu Frage 7 (...**lenkt die Aufmerksamkeit eher auf die Klangfarbe?**) und Frage 8 (...**weckt Ihr Interesse über längere Zeit und warum?**). Medianwert über alle vier Gegenüberstellungen zu Frage 7.

Die Analyse der Ergebnisse der einzelnen Fragen mittels Binomialtest nach [87, S. 279] liefert folgendes Ergebnis: Die Nullhypothese, nach welcher die Wahrscheinlichkeit des Auftretens der Beurteilung ‚MMA‘ gleich bzw. kleiner der Wahrscheinlichkeit $p = 0,5$ (gleiche Wahrscheinlichkeit für beide Konstellationen) ist, kann bei einseitiger Fragestellung auf dem 5% - Niveau für vier der acht Fragen abgelehnt werden. Es sind dies die Fragen 2 (Lokalisierbarkeit), 3 (Breite), 5 (Lautstärke) und 8 (Interesse über längere Zeit). Anders formuliert: Es kommt bei diesen vier Fragen die Alternativhypothese zum Tragen, nach welcher die Wahrscheinlichkeit des Auftretens der Beurteilung ‚MMA‘ größer als die Wahrscheinlichkeit $p = 0,5$ ist (vgl. Anhang B, 7.4).

Weiter ergibt die Analyse nach dem Chi – Quadrat – Mehrfeldertest nach [88, s. S. 184], dass es bei dem Vergleich der Ergebnisse der Fragen 1 bis 8 auf einem Niveau von $\alpha = 1\%$ keine signifikanten Unterschiede gibt.

Die Angaben der Versuchspersonen zu Frage 8 sind in Tabelle 12 aufgelistet.

Angaben bei Nennung OCT	Angaben bei Nennung MMA
Klangbild/farbe, angenehmer zum Zuhören	Wirkt nicht so lebhaft wie 1 (OCT), aber ‚räumlicher‘
Transparenz, Ortbarkeit der Instrumente, ‚Echtheit der Abbildung‘	Transparenter von den Instrumenten her
Nicht so grob	Macht mehr auf, Musik ist sichtbarer, greifbarer
2 (MMA) ist zu präsent -> Ermüdung	?
Eingehüllt sein, räumlicher	Transparenter, Instrumente besser lokalisierbar
Wärmerer Klang, passt besser zur Musik	Höhere Räumlichkeit, Live - Charakter
	Brillanter, poröser, schärfer
	Homogener
	Fast wie im Ensemble mitten drin
	Bessere Einhüllung
	Voller und brillanter, besserer Raumeindruck
	Mehr Details, Dynamik, mehr Klang, leichte Surround – Empfindung im zweiten (MMA) Beispiel, mehr Raumeindruck, näher am Live - Erlebnis
	Lebendiger
	Klarer, breiter, natürlicher
	?

Tabelle 12: Angaben der VPN zur Frage 8 (Welches der beiden Beispiel aus der nächsten Gegenüberstellung weckt ihr Interesse über längere Zeit und warum?, ‚?‘ kennzeichnet keine Angabe der VPN)

Die Auswertung jener 15 Fragebögen, bei welchen nach Ende des Versuchsdurchlaufes die ersten zwei Fragen noch einmal beantwortet wurden, lieferte die in Abb. 52 und Abb. 53 dargestellten Ergebnisse.

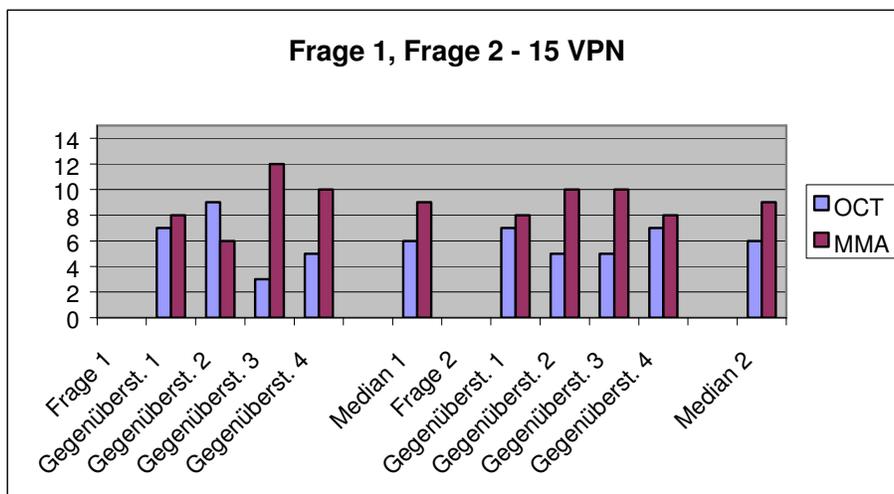


Abb. 52: Häufigkeiten der Nennungen OCT, bzw. MMA zu Frage 1 und Frage 2 jener VPN, welche diese zwei Fragen nach dem Versuchsdurchlauf nochmals beantworteten. Median über alle vier Gegenüberstellungen je Frage.

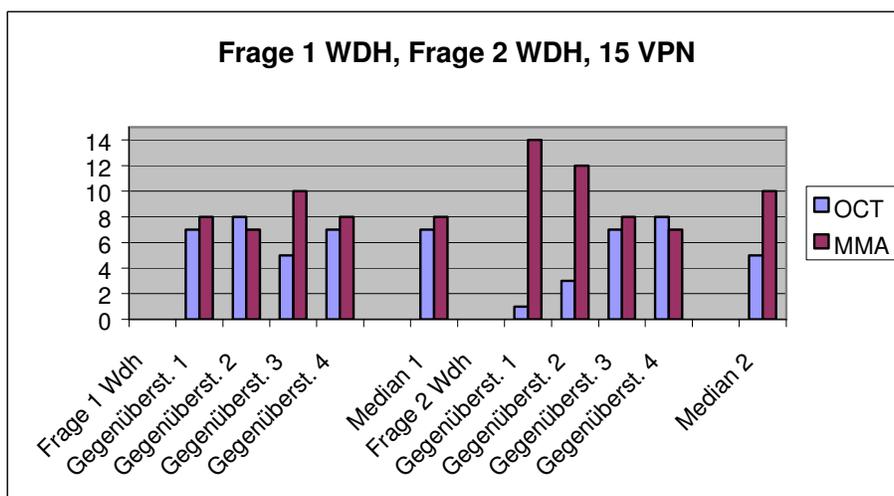


Abb. 53: Häufigkeiten der Nennungen OCT, bzw. MMA zu Frage 1 WDH und Frage 2 WDH. Median über alle vier Gegenüberstellungen je Frage.

Die Analyse mit dem Chi – Quadrat – Mehrfeldertest nach [87, S. 279] ergibt, dass es bei der Wiederholung der Frage 1 und der Frage 2 und dem Vergleich der Ergebnisse auf einem Niveau von $\alpha = 1\%$ keine signifikanten Unterschiede gibt.

Mit Hilfe der Daten von Frage 1, Frage 2 und der Wiederholung dieser beiden Fragen wurden die 15 VPN auf Konsistenz ihrer Antworten hin untersucht. Aus den acht Gegenüberstellungen ergab sich für die 15 VPN eine mittlere Konsistenz von 4,67 (aus Abb. 54). Das heißt bei der Wiederholung der Beantwortung von Frage 1 und Frage 2 wurde pro

VPN im Mittel 4,67 mal die gleiche Mikrofonkonstellation gewählt wie bei der ersten Beantwortung.

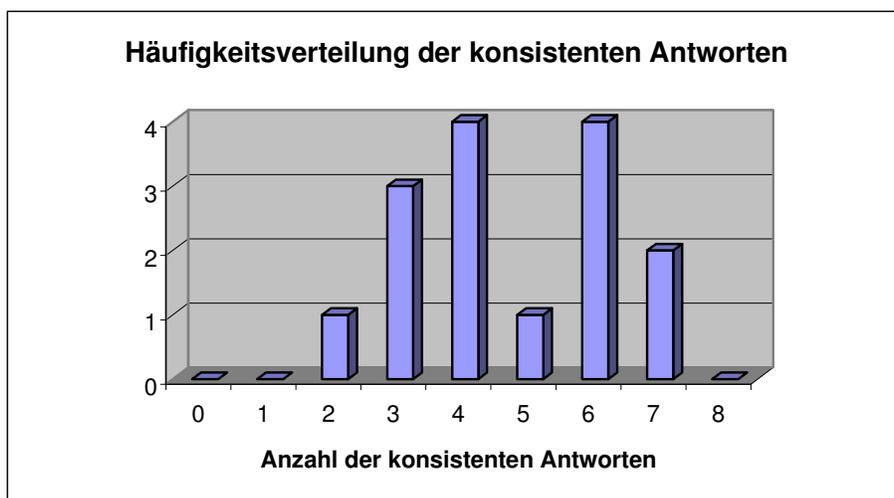


Abb. 54: Häufigkeitsverteilung konsistenter Antworten

Mit Hilfe dieser mittleren Konsistenz und der individuellen Konsistenz jeder einzelnen VPN wurden die Bewertungen der 15 VPN linear gewichtet. Gab eine VPN drei konsistente Antworten wurde sie mit 0,64 gewichtet, gab eine VPN sechs konsistente Antworten wurde sie mit 1,29 gewichtet usw. Die Nennungen der sechs aus Zeitgründen nicht geprüften VPN wurden weiterhin mit 1 gewichtet. Die Ergebnisse sind in Abb. 55 bis Abb. 58 dargestellt.

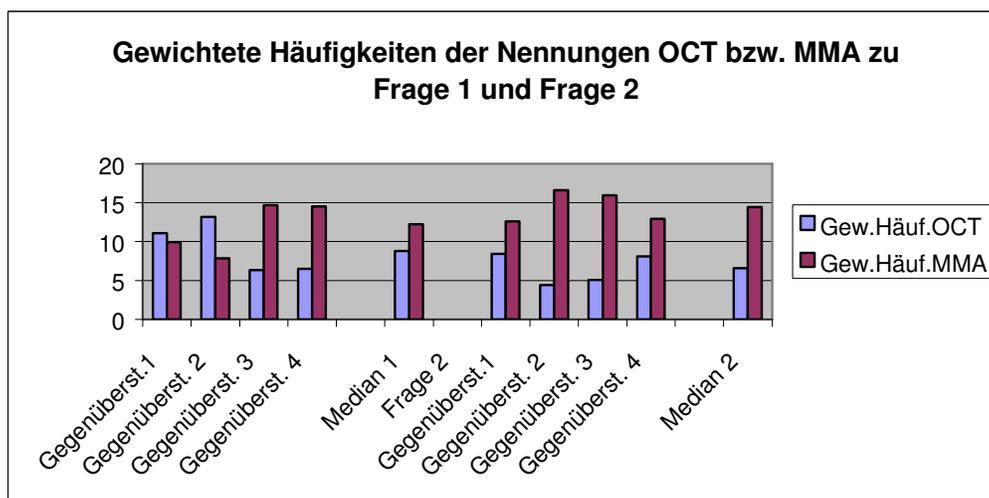


Abb. 55: Gewichtete Häufigkeiten der Nennungen OCT bzw. MMA zu Frage 1 (...**gefällt Ihnen besser?**) bzw. Frage 2 (...weist eine **bessere Lokalisierbarkeit** einzelner Instrumente auf?), Median über alle vier Gegenüberstellungen zu beiden Fragen.

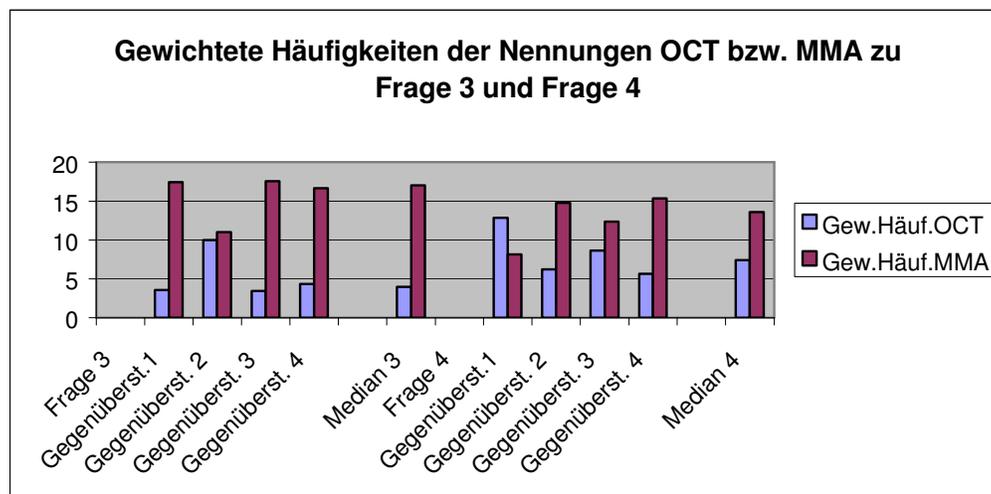


Abb. 56: Gewichtete Häufigkeiten der Nennungen OCT bzw. MMA zu Frage 3 (...vermittelt eine **größere Breite** des aufgenommenen Ensembles?) bzw. Frage 4 (...vermittelt eine **größere Tiefe (Tiefenstaffelung)**?), Median über alle vier Gegenüberstellungen zu beiden Fragen.

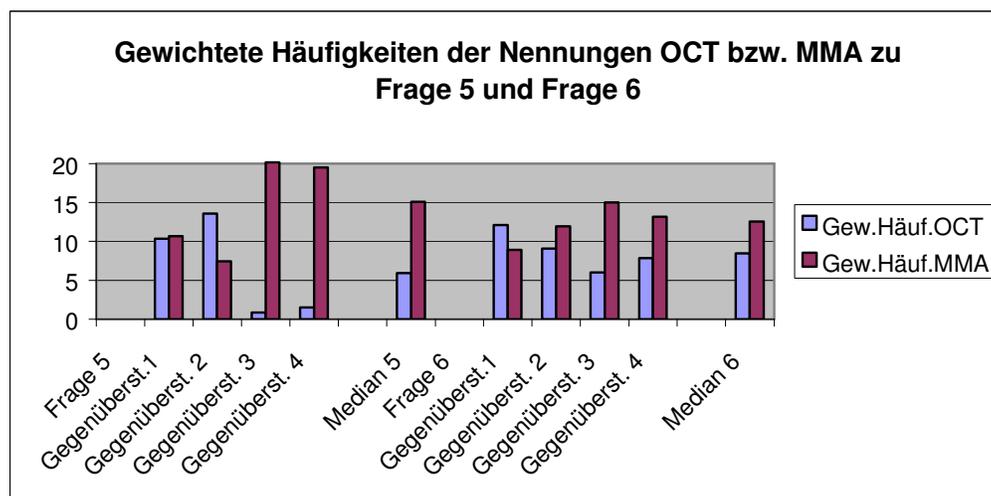


Abb. 57: Gewichtete Häufigkeiten der Nennungen OCT bzw. MMA zu Frage 5 (...**wirkt lauter**?) bzw. Frage 6 (...**läßt Proportionen einzelner Instrumente zueinander besser erkennen**?), Median über alle vier Gegenüberstellungen zu beiden Fragen.

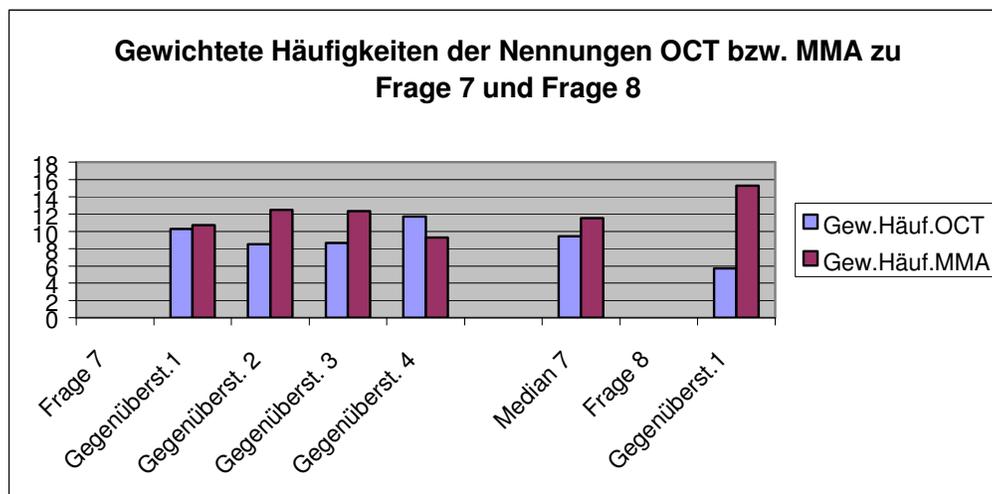


Abb. 58: Gewichtete Häufigkeiten der Nennungen OCT bzw. MMA zu Frage 7 (...lenkt die Aufmerksamkeit eher auf die Klangfarbe?) bzw. Frage 8 (...weckt Ihr Interesse über längere Zeit und warum?), Median über alle vier Gegenüberstellungen zu Frage 7.

Die Analyse der gewichteten Ergebnisse mittels Binomialtest zeigt, dass neben den Fragen 2 (Lokalisierbarkeit), 3 (Breite), 5 (Lautstärke) und 8 (Interesse über längere Zeit) im Vergleich zu den nicht gewichteten Ergebnissen die Nullhypothese (Wahrscheinlichkeit des Auftretens der Beurteilung ‚MMA‘ ist gleich bzw. kleiner der Wahrscheinlichkeit $p = 0,5$) auch noch für Frage 4 (Tiefe (Tiefenstaffelung)) bei einseitiger Fragestellung auf dem 5% - Niveau abgelehnt werden kann.

4.5.5 AUSWERTUNG DER DATEN ZU DEN VERSUCHSPERSONEN

Die Angaben, welche die Versuchspersonen zu ihren persönlichen Daten gemacht haben, ermöglichen eine detailliertere Betrachtung der Ergebnisse. So auch die Zuordnung der Versuchspersonen zu Geschlecht und Alter (siehe Abb. 59).

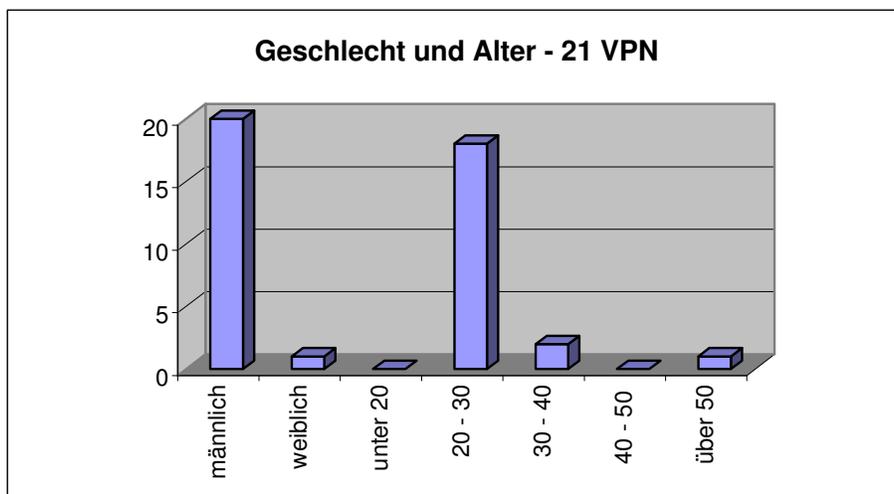


Abb. 59: Ergebnis der Auswertung der Frage nach Geschlecht und Alter

Der relativ große Anteil an männlichen Versuchspersonen ist darauf zurückzuführen, dass der Frauenanteil in dieser technischen Disziplin grundsätzlich sehr gering ist. Der große Anteil an Versuchspersonen mit einem Alter von 20 bis 30 Jahren hat seine Ursache in der großen Teilnahme von Studenten bei dieser Versuchsreihe.

13 VPN gaben an, dass Surround – Erfahrungen vorhanden seien, 8 VPN gaben ‚nicht vorhanden‘ zur Antwort. Detaillierte Auflistung, welche Surround – Erfahrungen die Versuchspersonen angaben, sind aus Tabelle 13 ersichtlich.

Kino	3
Mitwirkung bei diversen Projekten	8
Zuhause	2
Studio	1
DVD (DTS)	1
Dolby Pro Logic	1

Tabelle 13: Angaben zu den persönlichen Surround - Erfahrung der 21 VPN (mehrfach – Angaben möglich)

Die Auswertung der Frage nach der persönlichen Verfassung der Versuchspersonen ergab das in Abb. 60 dargestellten Ergebnis.

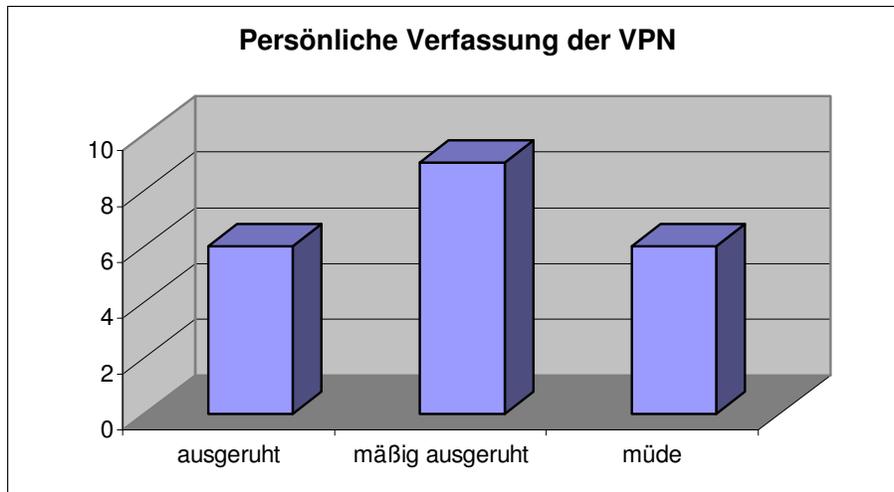


Abb. 60: Ergebnis der Auswertung der Frage nach der persönlichen Verfassung

Das Ergebnis der Auswertung der Frage nach der persönlichen Motivation ist in Abb. 61 dargestellt.

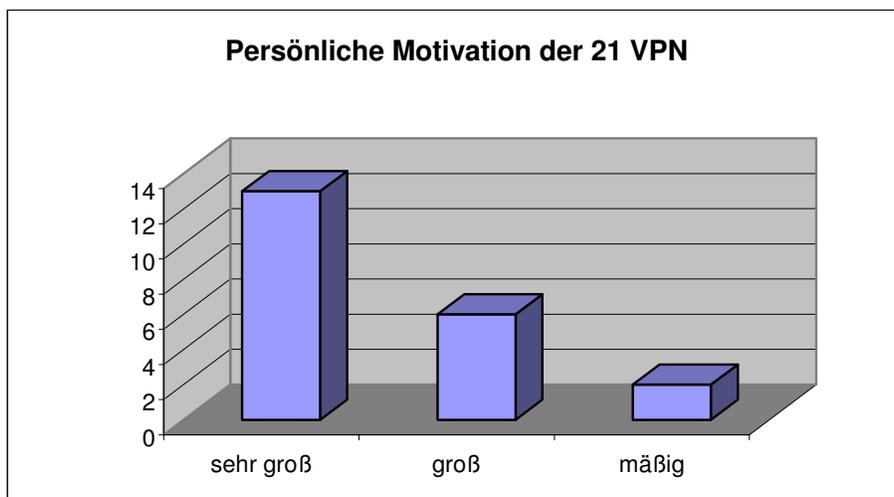


Abb. 61: Ergebnis der Auswertung der Frage nach der persönlichen Motivation

Abschließend hatten die Versuchspersonen die Möglichkeit, ‚Sonstige Anmerkungen‘ zu machen. Diese sind in Tabelle 14 zusammengefasst.

Beispiele zu kurz	3
Surround – Lautsprecher schlecht zu hören	3

Tabelle 14: Sonstige Anmerkungen der 21 VPN

Die Nennungen zu ‚Beispiele zu kurz‘ können als Anregung für die Gestaltung weiterer Versuchsabläufe dienen.

Die Nennungen zu ‚Surround – Lautsprecher schlecht zu hören‘ sind, laut Aussagen der Versuchspersonen, auf eine Erwartungshaltung zurückzuführen, welche durch die Erfahrung mit Spezial – Effekte beim Film (vor allem Kino) geschaffen wird. Da es sich bei dieser Versuchsreihe um Programmmaterial einer Darbietung von vorne handelt, treten die Surround – Lautsprecher nicht spektakulär in Erscheinung. Sie vermitteln Atmosphäre und Einhüllung (vgl. 3.2.2).

4.6 ZUSAMMENFASSUNG

Die Ergebnisse der durchgeführten Versuchsreihe lassen im Vergleich eine Tendenz der Bevorzugung des Multichannel Microphone Array (MMA) nach Williams (vgl. 4.2, Abb. 45) gegenüber dem Optimized Cardioid Triangle (OCT) nach Theile (vgl. 4.2, bzw. 3.2.6.1) erkennen (vgl. 4.5.4). Diese Tendenz ist vor allem bei den Parametern ‚Lokalisierbarkeit‘, ‚Breite‘, ‚Lautstärke‘ und ‚Interesse über längere Zeit‘ erkennbar (vgl. 4.5.4). Gewichtet man die Bewertungen nach der Konsistenz der Versuchspersonen, so gilt dies zusätzlich für den Parameter Tiefenstaffelung.

Inwiefern der unterschiedliche Aufnahmewinkel der verwendeten Konstellationen (MMA: 110° , OCT: 100°) und der daraus resultierende unterschiedliche Aufstellungsort (Differenz zwischen den Centermikrofonen von 60 cm) dieses Ergebnis beeinflusst hat, kann erst durch weitere Versuche mit einem neuen Prototypen der OCT – Anordnung (mit variablem Aufnahmewinkel) beantwortet werden.

5 LITERATURVERZEICHNIS

- [1] Thompson S. P.: ‚On the function of the two ears in the perception of space‘, Philos. Mag. 13, S. 406-416, 1882.
- [2] Batteau D. W.: ‚The role of the pinna in human localization‘, Proc. R. Soc. London Ser. B 168, S 158-160, 1967.
- [3] Hebrank J. H., Wright D.: ‚Are two ears necessary for localization of sound sources on the median plane?‘, Journal of the Acoustical Society of America, 56, pp. 935-938, 1974.
- [4] Köring J.: ‚Anwendungsbezogene Modifikation eines Verfahrens zur richtungsgetreuen Schallwiedergabe über zwei Lautsprecher‘, Diplomarbeit, RWTH Aachen, 1992.
- [5] Schmidt R.F., Thews G.: ‚Physiologie des Menschen‘, 27. Auflage, Springer Verlag Berlin, 1997
- [6] Begault D. R.: ‚3-D Sound for Virtual Reality and Multimedia‘, AP Professional, USA 1994.
- [7] Blauert J.: ‚Räumliches Hören‘, S. Hirzel Verlag Stuttgart, 1974.
- [8] Hugonnet C., Waldner P.: ‚Stereophonic Sound Recording‘ John Wiley & Sons Inc., 1995.
- [9] Webers J.: ‚Tonstudioteknik, Handbuch der Schallaufnahme und –wiedergabe bei Rundfunk, Fernsehen, Film und Schallplatte‘; Franzis Verlag, München, 1985.
- [10] Fellner M., Höldrich R.: ‚Physiologische und Psychoakustische Grundlagen des räumlichen Hörens‘, IEM-Report 03/98, Univ. f. Musik u. darst. Kunst Graz, IEM, Februar 1998.
- [11] Michael Williams, Guillaume Le Dû: ‚Microphone Array Analysis for Multichannel Sound Recording‘, 107. AES - Convention, New York, 1999.
- [12] Michael Williams, Guillaume Le Dû: ‚Multichannel Microphone Array Design‘, 108. AES – Convention, Paris, 2000.
- [13] Michael Williams: ‚The Stereophonic Zoom, A Practical Approach to determining the Characteristics of a Spaced Pair of Microphones‘, 74. AES - Convention, Paris, Preprint 2072 (D1), 1984.
- [14] Michael Williams: ‚Unified Theory of Microphone Suystems for Stereophonic Sound Recording‘, 82. AES – Convention, London, Preprint 2466 (H6), 1987.

-
- [15] Michael Williams: Operational Limits of the Variable M/S Stereophonic Microphone System, 88. AES – Convention, Montreux, Preprint 2931 (K3), 1990.
- [16] Michael Williams: Early Reflections an Reverberant Field Distribution in Dual Microphone Stereophonic Sound Recording Systems, 91 AES – Convention, New York, Preprint 3155 (R4), 1991.
- [17] Guillaume Le Dû, Memoire de Recherche de l'Ecole Nationale Supérieure Louis Lumière: Conception de systèmes de prise de son mulicanaux, 1999.
- [18] Theile, G.: Multichannel Natural Music Recording Based On Psychoacoustic Principles, Handout, 108th AES Convention, Paris, 2000.
- [19] Surround-Sound-Forum – 01-SV-E1, Recommended Practice, Listening Conditions and Reproduction Arrangemets for Multichannel Stereophony (01/00), bzw. deutsch: Empfehlungen für die Praxis SSF – 01-10/98
- [20] ITU-Empfehlung: Recommendation ITU-R BS. 775-1: Multichannel Stereophonic Sound System with and without accompanying Picture (Genf, 1992-1994).
- [21] EBU-Empfehlung: Listening conditions for the assessment of sound programme material. EBU Tech document 3276 - 1997 (second edition).
- [22] ITU-Empfehlung: Recommendation ITU-R BS 1116-1: Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound
- [23] Deutsche Norm: DIN 15996, Ausgabe April 1996: Elektronische Laufbild- und Tonbearbeitung in Film-, Video- und Rundfunkbetrieben. Anforderungen an den Arbeitsplatz (DIN 15996:1996-04).
- [24] Surround-Sound-Forum (VDT/IRT/SRT): Multichannel Recording in 3/2-Format, Parameters for Programme Interchange and Archiving, Alignment of Reproduction Equipment. Recommended Practice SSF- 02/1-E-2 (3-5-99).
- [25] ITU-Empfehlung: Recommendation ITU-R BR.[10C/AADOC. 10/10]: Parameters for International Exchange of Multichannel Sound Recordings (Question ITU-R 215/10), Doc. 10/BL/2, 31 July 1998).
- [26] ITU-Empfehlung: Recommendation ITU 645-2: Test Signals and Metering to be used on International Sound Programme Connections (Genf, 1986-1992).
- [27] EBU-Empfehlung: Exchange of Sound Programmes as digital Tape Recordings. EBU Technical Recommendation R 64 - 1993.
- [28] EBU-Empfehlung: Alignment Level in digital Audio Production Equipment and in digital Audio Recorders. EBU Technical Recommendation R 68 - 1992.
- [29] EBU-Empfehlung: Track allocations and recording levels for the exchange of multichannel audio signals. EBU Technical Recommendation R 91-1998.

-
- [30] SMPTE Recommended Practice: Loudspeaker Arrangements for Audio Monitoring in Television Productions (SMPTE RP 173 - 19).
- [31] SMPTE Recommended Practice: Audio Levels for digital Audio Records on digital Television Tape Recorders (SMPTE RP 155 - 1995).
- [32] Proposed SMPTE Standard for Television: Channel Assignments and Levels on Multichannel Audio Media. ITU Information document ITU-R. 10C/11 and 10-11R/24; 16 March 1998 (E).
- [33] Steinke, G.: „Das Pegelprofil in der Tonstudio- und Rundfunkübertragungstechnik“. In: dB-Magazin, Köln, (1990), H.1/2, S.50-64.
- [34] Blauert, J.: Spatial Hearing - The Psychophysics of Human Sound Localisation. The MIT Press, Cambridge, Mass, 1996, ISBN 0-262 02413-6 .
- [35] Barron, M., Marshall, H.A.: Spatial Impression due to early lateral reflections in concert halls. J. Sound Vibr. 77, 1981, pp. 211-232.
- [36] Herrmann, U., Henkels, V., Braun, D.: Vergleich von fünf Surround Mikrofonverfahren. 20. Tonmeistertagung, 1998, Proceedings (ISBN 3-598-20361-6), pp. 508-517
- [37] Everest, A.; Streicher, R.: The New Stereo Soundbook.. TAB Books, 1992, ISBN 0-8306-3903-9
- [38] Streicher, R.: The Decca Tree in stereo and surround recording. 106th AES Convention, 1999, Workshop Note
- [39] Griesinger, D.: Theory and design of a digital audio signal processor for home use. J. Audio Eng. Soc. 37, 1989, pp. 40-50.
- [40] Theile, G.: On the performance of two-channel and multi-channel stereophony. 88th AES Conference, 1990, Preprint 2932
- [41] Theile, G.: On the naturalness of two - channel stereo sound. J. Audio Eng. Soc.(39), 1991, pp. 761-767.
- [42] Wöhr, M.; Theile, G.; Goeres, H.-J.; Persterer, A.: Room - related balancing technique: a method for optimising recording quality. J. Audio Eng. Soc.(39), 1991, pp. 623-631.
- [43] Horbach, U.: New techniques for the production of multichannel sound. 103rd AES Convention, 1997, Pre-print 4624
- [44] ETSI Technical Report. Draft ETR 154 Rev 3 (June 1999): Digital Video Broadcasting (DVB): Implementation guidelines for the use of MPEG-2 systems. (Annex D: Implementation of Ancillary Data for MPEG Audio)
- [45] Cossette, S.: Metadata issues for ATSC audio. SMPTE Journal, July 1999, pp. 486-490

-
- [46] Media Biz, Film/TV/Radio/Video/Audio, Nr 66, Sommer 2000. www.mediabiz.co.at
- [47] Bruck J.: Die Lösung des 'Surround' - Dilemmas, 19. Tonmeistertagung Karlsruhe, Tagungsband pp. 124, 1996.
- [48] Marks J.: Soundfield of Dreams, Surround Professional, p. 48, January 1999.
- [49] Fukada, A.; Tsujimoto, K.; Akita, S.: Microphone techniques for ambient sound on a music recording. 103rd AES Convention, 1997, Preprint 4540
- [50] Sawaguchi, M.; Fukada, A.: Multichannel sound mixing practice for broadcasting. IBC Conference 1999, Proceedings, in print
- [51] Keibs, L.: Stereo - Ambiophone Schallübertragung (Stereo-ambiophonic sound transmission). Acustica (12), 1962, pp. 118-124
- [52] Karamustafaoglu, A.; Horbach, U.; Pellegrini, R.; Mackensen, P.; Theile, G.: Design and application of a data - based auralisation system for surround sound.
- [53] Theile, G.: On the Theory of Optimum Reproduction of Stereophonic Signals through Loudspeakers and Earphones". (German). Rundfunktechn. Mitt., 1981, pp. 155 – 170
- [54] Holman, T.: Surround Sound Systems used with pictures in cinemas and homes. 8th AES Conference, 1990
- [55] Silzle, A., Theile, G.: HDTV-Multichannel sound: Study on imaging quality of the center loudspeaker (German). 16. Tonmeistertagung, 1990, Proceedings (ISBN 3-598-20357-8), pp. 208-218
- [56] Wöhr, M.; Nellesen, B., Theile, G.: Studies on the selection of the main microphone method (German). 14. Tonmeistertagung, 1986, Proceedings (ISBN 3-598-20355-1), pp. 106-120
- [57] Damaske, P.: Subjective Investigation of Sound Fields. Acustica, Vol.19, 1967/68, pp. 199-213
- [58] Cremer, L.: Zur Verwendung der Begriffe "Korrelationsgrad" und "Kohärenzgrad". Acustica, Vol.35, 1976, pp. 215-218
- [59] Theile, G.: Capabilities and limits of 3/2-stereo recording. 19. Tonmeistertagung, 1996, Proceedings (ISBN 3-598-20360-8), pp. 139-158
- [60] Mitchell D. Entschlüsselt: Der Dolby Digital Code, 'Dolbys Digital Squasher' aus Audio Media, Ausgabe 93, August 1998.
- [61] <http://www.dolby.com/tech/whtppr.html>
- [62] Enter the World of Multichannel. Surround Sound Technologies from Dolby Laboratories, 1999 Dolby Laboratories Inc. S99/12377.

- [63] Catch the next Innovation. Dolby E, Multichannel Audio Coding for DTV Production and Distribution, 1999 Dolby Laboratories Inc. S99/12484.Literatur
- [64] Sonntachi A., Höldrich R.: Konzept zur Schallfeldsynthese und Schallfeldreproduktion. Veröffentlichung zur 50. Jahrestagung der ÖPG, FA – Akustik gemeinsam mit der Österreichischen Akustischen Gesellschaft (AAA)
- [65] P. M. Morse and K. U. Ingard. Theoretical Acoustics. Mc Graw-Hill, New York City, 1968.
- [66] John Vanderkooy and Stan Lipshitz. "Anomalies of Wavefront Reconstruction in Stereo and Surround -Sound Reproduction". 83. AES Convention, Oktober 1987, Preprint 2554 (G8).
- [67] Michael A. Gerzon. "Design of Ambisonic Decoders for Multispeaker Surround Sound". 58. AES - Convention, November 1977 - kein Preprint.
- [68] Michael a. Gerzon. "The Rational Systematic Design of Surround Sound Recording and Reproduction Systems. Part 1. General Theory of Directional Psychoacoustics and Applications". Appendix C of the Comments of National Research Development Corporation to the Federal Communications Commission in the matter of FM Quadrophonic Broadcasting, Docket 21310, Juni 1979.
- [69] Jeffery S. Bamford "An Analysis of Ambisonic Sound Systems of First and Second Order". Thesis at the University of Waterloo, Ontario, Canada, 1995.
- [70] Michael A. Gerzon. "Optimum Reproduction Matrices for Multispeaker Stereo". AES - Journal, 40(7/8):71 - 589, 1992.
- [71] Sonntachi A., Hoeldrich R. : Ambisonic System mit unsymmetrischem Lautsprecherlayout. IEM-Report 09/99, Univ. f. Musik u. darst. Kunst, Graz, 1999.
- [72] http://www.digitalaudioguide.com/faq/dvd-audio/faq_intro.htm Stand: 27.September 2000
- [73] http://www.york.ac.uk/inst/mustech/3d_audio/ambis2.htm Stand: 27.September 2000
- [74] Michael A. Gerzon. "Ambisonics in Multichannel Broadcasting and Video" in the Journal of the AES (Audio Engineering Society), Vol.33 No.11, pp859-871 November 1985.
- [75] <http://www.muse.demon.co.uk/vspace/vspace.html> Stand: 28. September 2000
- [76] <http://www.ircam.fr/departements/recherche/index-e.html> Stand: 28. September 2000
- [77] http://www.thx.com/theatres/sound_digital.html
- [78] http://www.sdds.com/whatis/q_a.html

- [79] ATRAC Adaptive Transform Acoustic Coding for MiniDisc, Reprint from the 93rd AES - Convention, San Fransisco, 1992 October 1-4Literatur
- [80] Graf, F.: Entwicklung eines Aufnahmesystems für psychoakustische Analysen. Diplomarbeit am Institut für Nachrichtentechnik und Wellenausbreitung der TU Graz, März 1999.
- [81] Mail von Peter Hecker vom Freitag, 09. Juni 2000 20:53 an surround-sound@tonmeister.de, Betreff: Re: surround-sound: Hörversuch zu Surround-Hauptmikrophonierungen
- [82] Mail von Schoeps GmbH vom Dienstag, 02. Mai 2000 17:51 an markus.nussbaumer@academy.at, Betreff: Surround-Hauptmikrofon.
- [83] Mail von Christoph Rieseberg <christoph.rieseberg@gmx.de> an surround-sound@tonmeister.de <surround-sound@tonmeister.de> vom Donnerstag, 08. Juni 2000 00:16, Betreff: surround-sound: Hörversuch zu Surround-Hauptmikrophonierungen
- [84] Surround - Sound - Forum: EMPFEHLUNG FÜR DIE PRAXIS SSF – 01, Ausgabe vom 01.10.1998.
- [85] Surround - Sound - Forum: EMPFEHLUNG FÜR DIE PRAXIS SSF – 02, Ausgabe vom 24.10.1998.
- [86] Saurug R. Meßtechnik für Raumakustikdesign. Projektarbeit am Institut für Nachrichtentechnik und Wellenausbreitung, TU - Graz, - April 2000.
- [87] Sachs L. Angewandte Statistik, 8. Auflage, Springer Verlag, 1997.
- [88] Zöfel P.: Statistik in der Praxis, Gustav Fischer Verlag, Stuttgart, Jena, 3. Auflage, 1992.
- [89] Diagramme zur Verfügung gestellt von der Fa. Schoeps GmbH
- [90] <http://www.akg-acoustics.com/english/microphones/spec/spec-ck61.htm>

6 ANHANG A

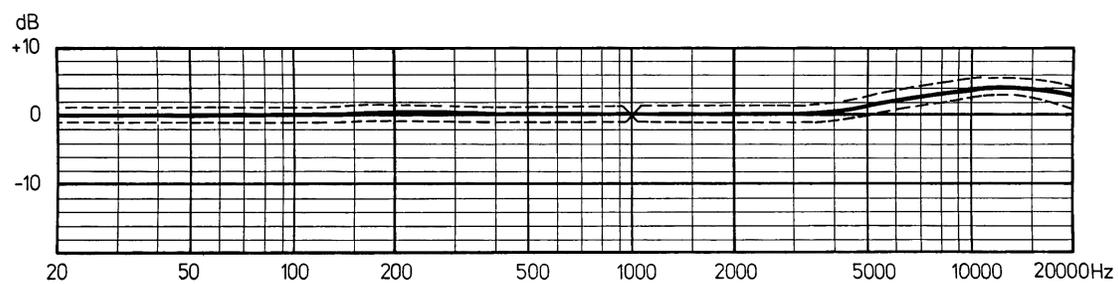


Abb. 62: Mikrofon - Frequenzgang Schoeps MK 2S. Kugelcharakteristik (s.u.), [89].

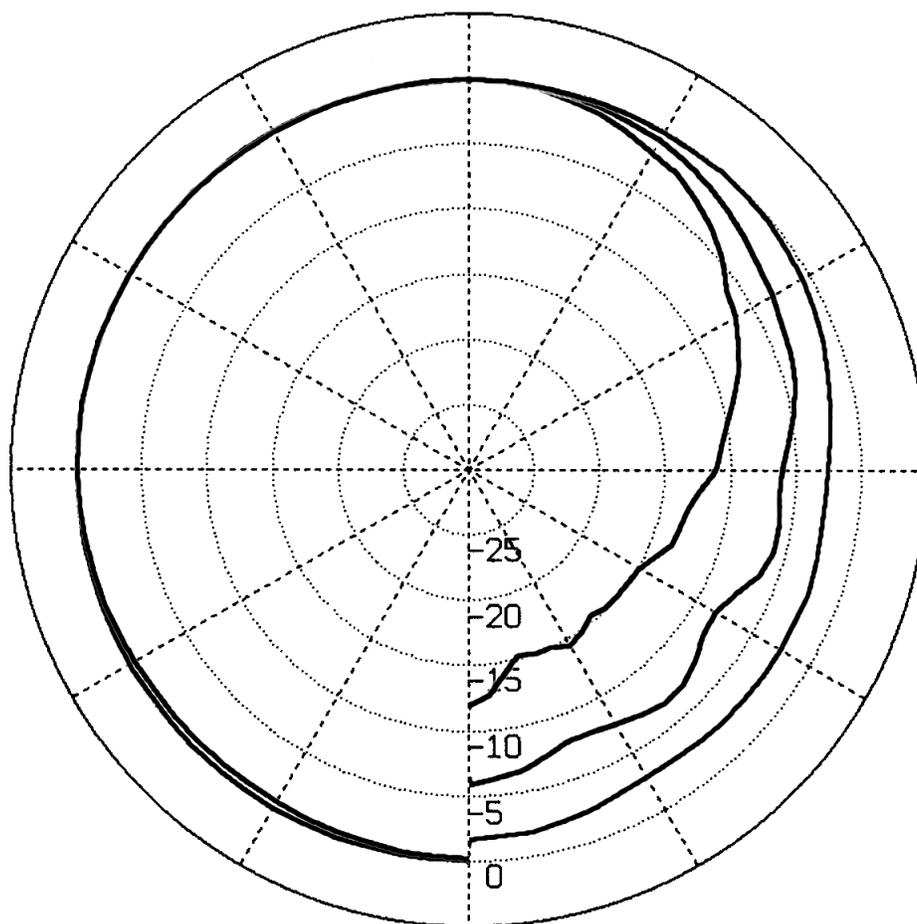


Abb. 63: Polardiagramm Schoeps MK 2S. Zuordnung der Kurven von links beginnend, jeweils von außen nach innen: 1kHz, 2kHz, 4kHz, 8kHz, 16kHz, [89].

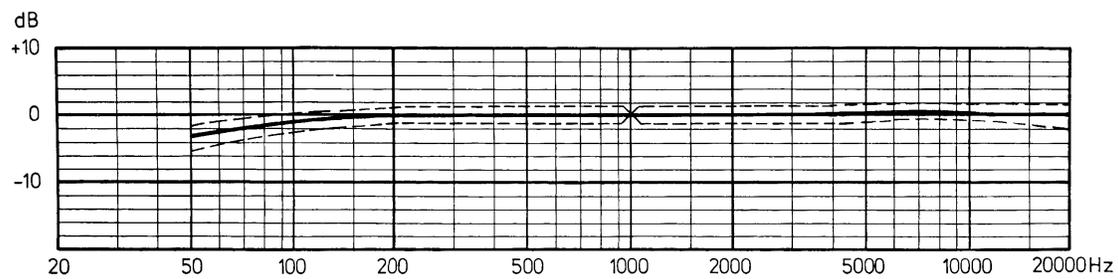


Abb. 64: Mikrofon - Frequenzgang Schoeps MK 4. Nierencharakteristik (s.u.), [89].

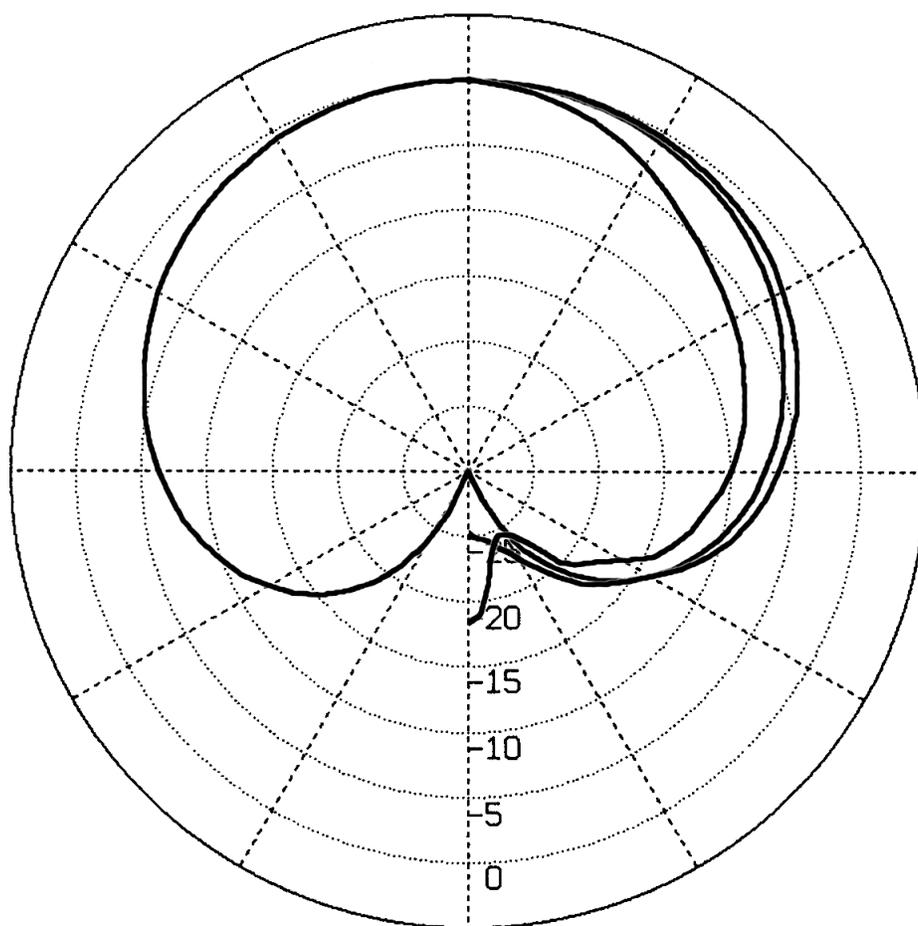


Abb. 65: Polardiagramm Schoeps MK 4. Zuordnung der Kurven von links beginnend, jeweils von außen nach innen: 2kHz, 4kHz, 8kHz, 16kHz, [89].

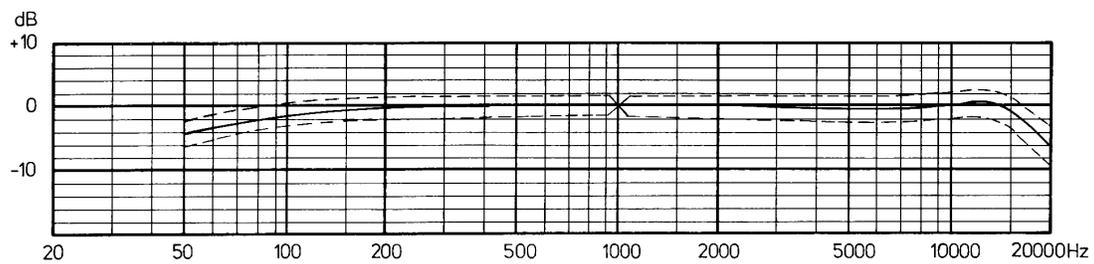


Abb. 66: Mikrofon - Frequenzgang MK 41V. Supernierencharakteristik (s.u.), [89].

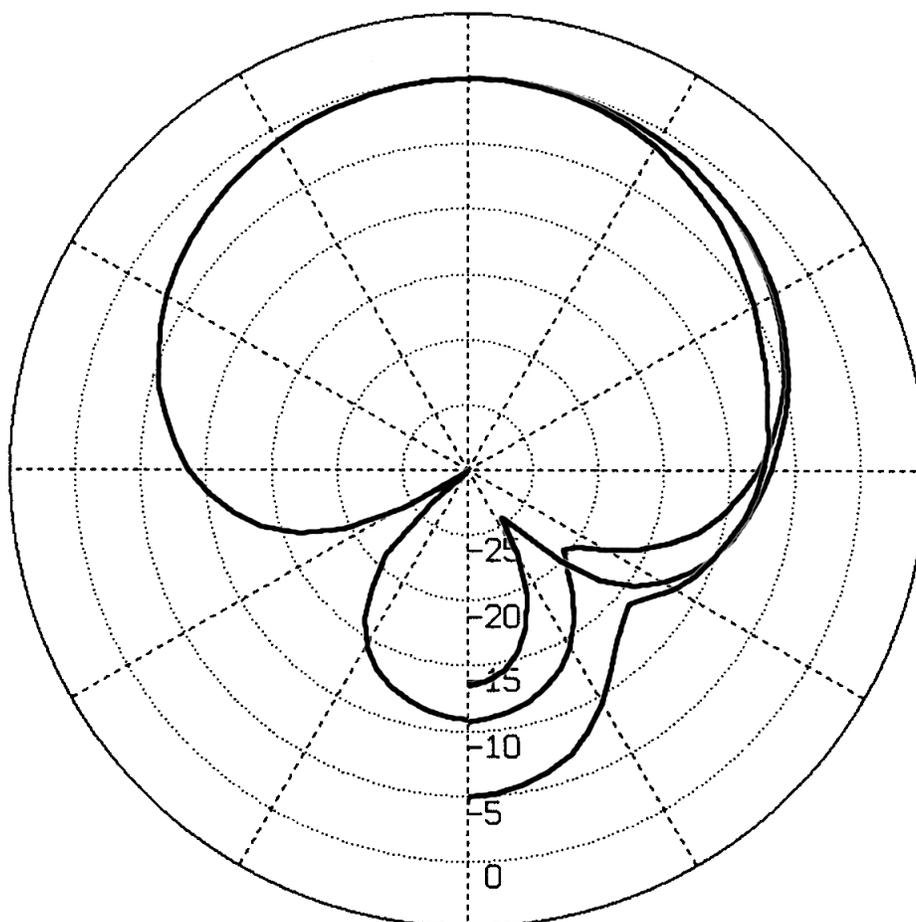


Abb. 67: Polardiagramm MK 41V. Zuordnung der Kurven von links beginnend, jeweils von außen nach innen: 1kHz/2kHz (ident), 4kHz (-13dB bei 180°), 8kHz, 16kHz (-5dB bei 180°), [89].

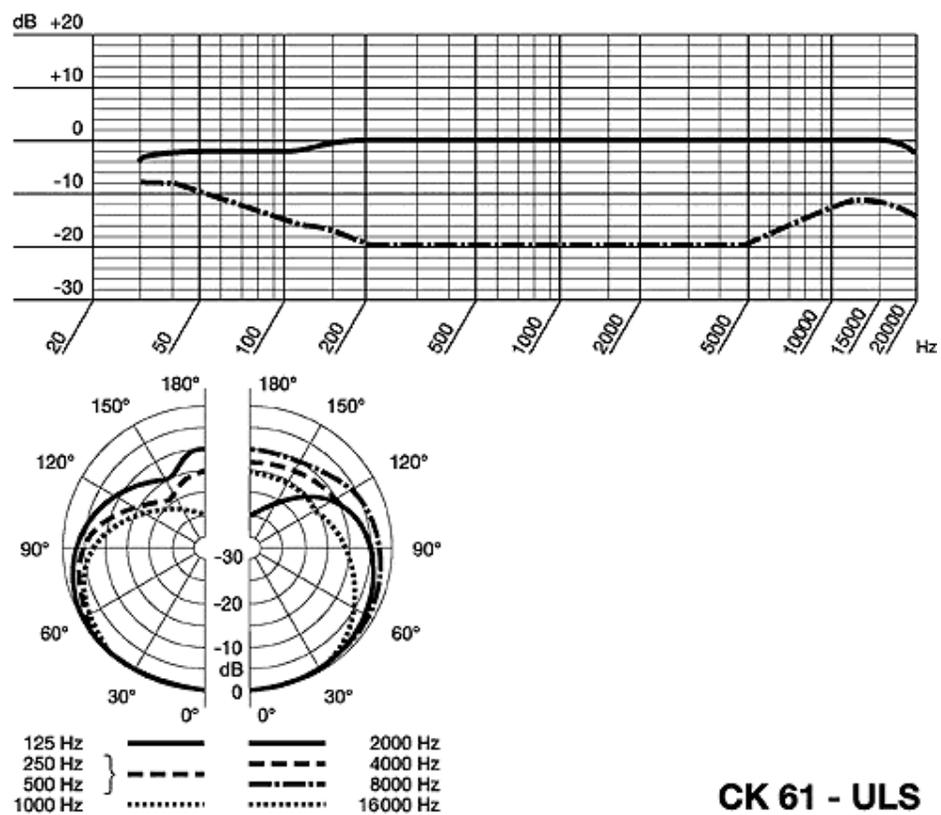


Abb. 68: Frequenzgang und Polardiagramm: CK 61 – ULS, aus [90]

7 ANHANG B

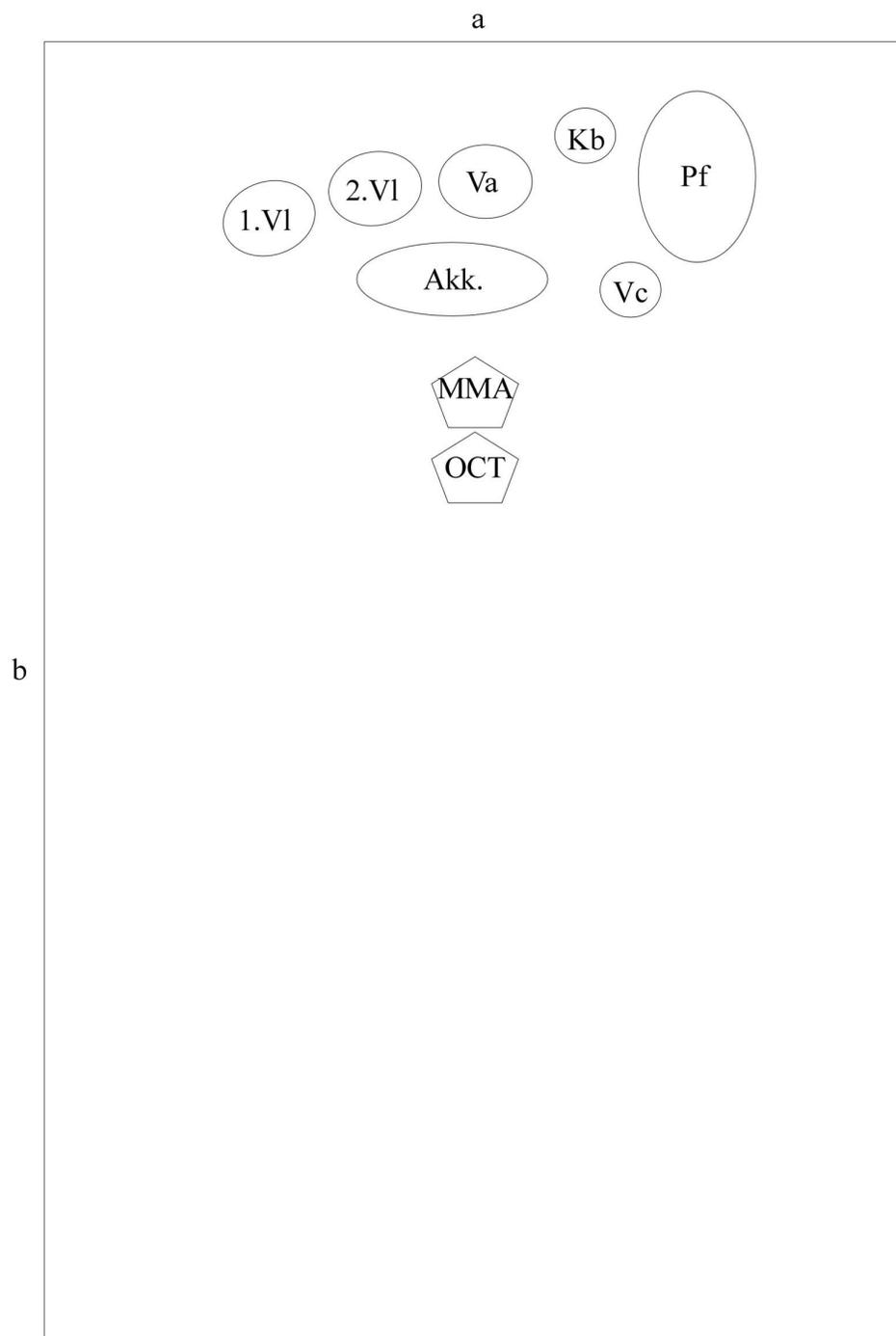


Abb. 69: Skizze des Aufnahmerraumes mit Instrumentengruppen und Mikrofonpositionen.

A = 9,5m, b = 14m. Raumhöhe: 4,1m, MMA = Multichannel Microphone Array, OCT = Optimized Cardioid Triangle incl. der hinteren Kanäle, 1./2.VI: jeweils zwei Violinen, Va: zwei Violen, Vc: ein Violoncello, Kb: ein Kontrabass, Pf: ein Pianoforte, Akk.: drei Akkordeons

Es soll darauf hingewiesen werden, dass dies keine Befragung im Sinne einer Prüfung ist. Wichtig ist Ihre subjektive Bewertung der folgenden Hörproben. Ihre Daten werden vertraulich behandelt.

7.1 HÖRVERSUCH ZU SURROUND - MIKROFONIERUNGS - TECHNIKEN

Es wurde mit verschiedenen Mikrofonierungstechniken ein Tango - Ensemble aufgenommen. Die Techniken werden einander gegenübergestellt und sollen bezüglich verschiedener Qualitäten und Parameter bewertet werden. Sie hören je Beurteilungskriterium vier Gegenüberstellungen. Bitte kreuzen Sie für jede Gegenüberstellung das zutreffende Beispiel an.

Die darauffolgende Frage zur Unterschiedlichkeit wird ebenfalls durch ankreuzen beantwortet.

1. Welches der Beispiele aus den folgenden Gegenüberstellungen **gefällt Ihnen besser?**

Gegenüberstellung	Beispiel	
1	1	2
2	1	2
3	1	2
4	1	2

Die Unterschiede bezüglich dieser Bewertung zwischen den Beispielen empfand ich

groß	mittel	klein
------	--------	-------

2. Welches der Beispiele aus den folgenden Gegenüberstellungen weist eine **bessere Lokalisierbarkeit** einzelner Instrumente auf?

Gegenüberstellung	Beispiel	
1	1	2
2	1	2
3	1	2
4	1	2

Die Unterschiede bezüglich dieser Bewertung zwischen den Beispielen empfand ich

groß	mittel	klein
------	--------	-------

3) Welches der Beispiele aus den folgenden Gegenüberstellungen vermittelt eine **größere Breite** des aufgenommenen Ensembles?

Gegenüberstellung	Beispiel	
1	1	2
2	1	2
3	1	2
4	1	2

Die Unterschiede bezüglich dieser Bewertung zwischen den Beispielen empfand ich

groß	mittel	klein
------	--------	-------

4) Welches der Beispiele aus den folgenden Gegenüberstellungen vermittelt eine **größere Tiefe (Tiefenstaffelung)**?

Gegenüberstellung	Beispiel	
1	1	2
2	1	2
3	1	2
4	1	2

Die Unterschiede bezüglich dieser Bewertung zwischen den Beispielen empfand ich

groß	mittel	klein
------	--------	-------

5) Welches der Beispiele aus den folgenden Gegenüberstellungen **wirkt lauter**?

Gegenüberstellung	Beispiel	
1	1	2
2	1	2
3	1	2
4	1	2

Die Unterschiede bezüglich dieser Bewertung zwischen den Beispielen empfand ich

groß	mittel	klein
------	--------	-------

6) Welches der Beispiele aus den folgenden Gegenüberstellungen **lässt Proportionen einzelner Instrumente zueinander besser erkennen**?

Gegenüberstellung	Beispiel	
1	1	2
2	1	2
3	1	2
4	1	2

Die Unterschiede bezüglich dieser Bewertung zwischen den Beispielen empfand ich

groß	mittel	klein
------	--------	-------

Welches der Beispiele aus den folgenden Gegenüberstellungen **lenkt die Aufmerksamkeit eher auf die Klangfarbe**?

Gegenüberstellung	Beispiel	
1	1	2
2	1	2
3	1	2
4	1	2

Die Unterschiede bezüglich dieser Bewertung zwischen den Beispielen empfand ich

groß	mittel	klein
------	--------	-------

8) Welches der beiden Beispiele aus der nächsten Gegenüberstellung **weckt Ihr Interesse über längere Zeit und warum?**: _____

Gegenüberstellung	Beispiel	
1	1	2

Die Unterschiede bezüglich dieser Bewertung zwischen den Beispielen empfand ich

groß	mittel	klein
------	--------	-------

--- Ende der Beispiele --- Bitte Daten zur Person angeben ---

Zum Abschluss möchte ich Sie um die Beantwortung folgender Fragen durch ankreuzen bitten.

Daten zur Testperson

Geschlecht

m	w
---	---

Alter

Unter 20 J.	Zw. 20 u. 30 J.	Zw. 30 u. 40 J.	Zw. 40 u. 50 J.	Über 50 J.
-------------	-----------------	-----------------	-----------------	------------

Momentane Uhrzeit

8:00 - 10:00	10:00 – 12:00	12:00 – 14:00	14:00 – 16:00	16:00 – 18:00	18:00 – 20:00
--------------	---------------	---------------	---------------	---------------	---------------

Persönliche Verfassung

Ausgeruht	Mäßig ausgeruht	müde
-----------	-----------------	------

Persönliche Motivation

Sehr groß	Groß	mäßig
-----------	------	-------

Persönliche Surround - Erfahrung

vorhanden	Nicht vorhanden
-----------	-----------------

Falls vorhanden, welche:

Sonstige Anmerkungen

Herzlichen Dank für Ihre Mitarbeit!

7.2 ZUSATZINFORMATION FÜR DIE VERSUCHSPERSONEN

- Zu Beginn des Hörversuchs gibt es zwei Beispiele zum Einhören, jedes der beiden ca. 1 min lang
- Danach beginnen die Hörbeispiele zu den entsprechenden Fragen
- Jede Frage wird kurz angesagt (Frage 1, Frage 2, ...)
- Die einzelnen Beispiele haben eine Dauer von ca. 10sec.
- Es soll jede Gegenüberstellung zu jeder Frage beantwortet werden
- Nach den Beispielen zu einer Frage stehen ca. 20 sec zur Vorbereitung auf die nächste Frage zur Verfügung
- Zum Abschluss des Versuchs sollen nach Abgabe des Fragebogens die ersten beiden Fragestellungen nochmals beantwortet werden

7.3 AUFLISTUNG AUFGEZEICHNETER SIGNALE UND ABLAUF DER VERSUCHSREIHE

Nr.	Kanal	Bez.	SPL / dB	SPL / dBA	PPM / dB ($\tau < 0.1\text{ms}$)	PPM / dB ($\tau < 10\text{ms}$)	RMS / dB	Anmerkung
100	L+C+ R+LS/ RS	Sinuston 1kHz			-18	-18	-18	
	L+C+ R+LS/ RS	Sinuston 1kHz			-9	-9	-9	
	L+C+ R+LS/ RS	Sinuston 1kHz			0	0	0	
110	L	Rosa Rauschen 200-20kHz	80	78	-11	-15	-20	Einstellung des Referenz-Abhörpegels je Wiedergabekanal $L_{\text{LIST ref}} = 78\text{dBA}$
	C	Rosa Rauschen 200-20kHz	80	78	-11	-15	-20	
	R	Rosa Rauschen 200-20kHz	80	78	-11	-15	-20	
	RS	Rosa Rauschen 200-20kHz	80	78	-11	-15	-20	
	LS	Rosa Rauschen 200-20kHz	80	78	-11	-15	-20	
111	L+C+ R+LS/ RS	Rosa Rauschen 200-20kHz, inkohärent	87	85	-11	-15	-20	Einstellung des resultierenden Referenz-Abhörpegels $L_{\text{LIST ref}} = 85\text{dBA}$

Tabelle 15: Verwendete Einmeß- und Testaufzeichnungen zur Einstellung von Wiedergabeanlagen für diskrete 3/2 - Mehrkanal – Stereophonie laut SSF – 02 (vgl. [85])

Bezeichnung	Timecode auf dem HI – 8 Band (mm:ss)
Track 100 aus Tabelle 15	1:00
Track 110 aus Tabelle 15	2:00
Track 111 aus Tabelle 15	7:45
Einleitende Hörbeispiele	10:00
Beispiele zu Frage 1	12:10
Beispiele zu Frage 2	14:05
Beispiele zu Frage 3	16:25
Beispiele zu Frage 4	18:45
Beispiele zu Frage 5	21:00
Beispiele zu Frage 6	23:30
Beispiele zu Frage 7	26:00
Beispiele zu Frage 8	28:15

Tabelle 16: Auflistung der Aufzeichnungen auf dem Hi - 8 Band (Ablauf der Versuchsreihe mit Timecode)

7.4 BEISPIEL ZUR ANALYSE DER STATISTISCHEN AUSWERTUNG

Binomialtest (vgl. [87])

Für die Überprüfung mit Binomialtest wurde die Summe der Nennungen MMA über alle vier Gegenüberstellungen gebildet.

Verglichen wird p mit der Wahrscheinlichkeit p_0 , wobei eine Tabelle der Verteilungsfunktion der Binomialverteilung benutzt wird. Für $n p_0 q_0 = n p_0(1 - p_0) > 9$ kann anhand der Standardnormalverteilung nach folgenden Bedingungen entschieden werden:

Hypothese paar: $H_{01} : p \leq p_0$, $H_{A1} : p > p_0$

H_0 wird auf dem 5 % - Niveau abgelehnt für: $x > x_o = n p_0 + 0,5 + 1,645 (n p_0(1 - p_0))^{1/2}$

Mit

x : Summe der Nennungen MMA

x_o : x_{oben}

n : Anzahl der Nennungen insgesamt

p_0 : = 0,5 (gleiche Wahrscheinlichkeit für beide Konstellationen)

wird $x_o = 84 * 0,5 + 0,5 + 1,645 (84 * 0,5 * 0,5)^{1/2} = 50,04$

Für Frage 1 (vgl. 4.5.4) wird $x = 45$.

Somit gilt: $x > x_o$ NICHT und es wird H_0 bei einseitiger Fragestellung auf dem 5 % - Niveau NICHT abgelehnt. Für die Durchführung des Binomialtests für die Fragen 2 bis 7 muss nur noch x ermittelt werden.