

# **Diplomarbeit**

an der  
Hochschule für Angewandte Wissenschaften Hamburg  
im Studiengang Medientechnik  
Schwerpunkt A/V-Medien

## **„Experimenteller Vergleich verschiedener Surroundmikrofonanordnungen hinsichtlich ihrer Eignung als Hauptmikrofon“**

Michael Abbing

Oktober 2004

1. Prüfer: Prof. Dr. Ulrich Schmidt
2. Prüfer: Prof. Dr. Robert Mores

# 1 .Einleitung

Der 5.1-(Surround)-Audio-Standard hat sich in den letzten Jahren immer mehr im Heimbereich etabliert und ist auf dem Weg, das 2.0-Stereoformat abzulösen. Mit der DVD-Video als Datenträger und den Codierverfahren Dolby-Digital und DTS ist auch für 5.1-Musikproduktionen ein qualitativ und quantitativ gutes Übertragungsmedium gegeben.

Mit den herkömmlichen Hauptmikrofonsystemen aus der Stereotechnik wurden aber bislang wenige überzeugende Aufnahme-Ergebnisse in 5.1-Musikproduktionen erzielt oder sie waren in ihrer Handhabung zu aufwendig und unrentabel. Und so wurden neue Mikrofonsysteme entwickelt, die gute Aufnahmen von akustischen Darbietungen im 5.1 Format bzw. 3/2-(Stereo)-Format mit relativ wenig Aufwand und in kompakter Form ermöglichen sollen.

Zum einen werden in dieser Diplomarbeit die wichtigsten, zur Zeit auf dem Markt erhältlichen, Surround-Hauptmikrofonsysteme vorgestellt.

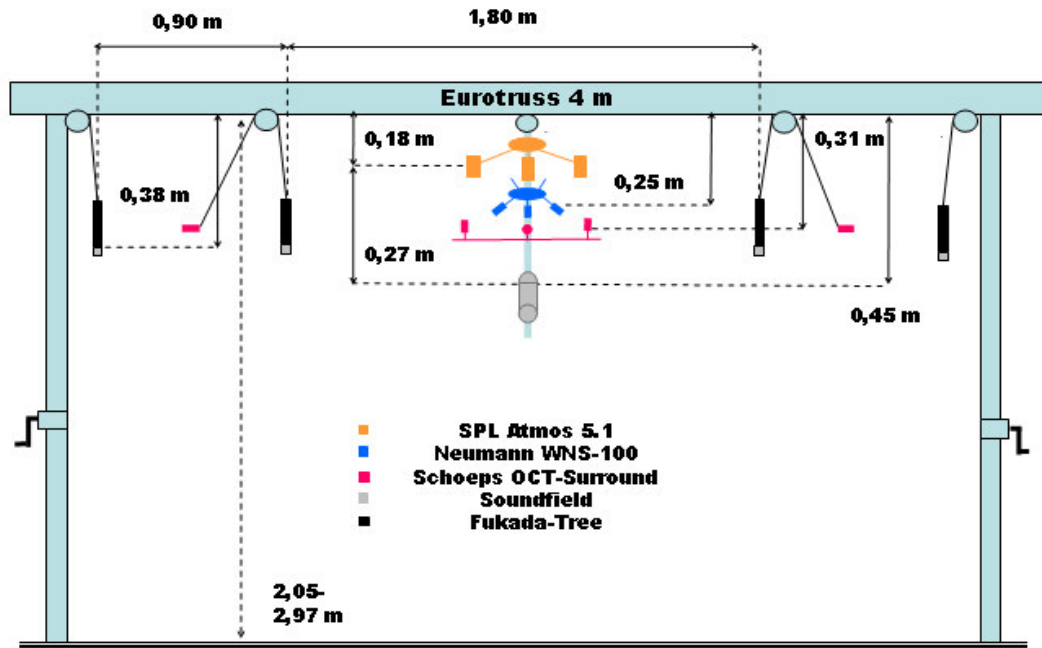
Zum anderen werden 5 dieser Systeme in einem groß angelegten Versuch getestet und als Hauptmikrofon in der Pop-Musik-Produktion beurteilt.

Dabei werden nicht die individuell für jedes Mikrofon geeignetsten Positionen ermittelt und ausgewählt, sondern alle Anordnungen mit möglichst geringer Abstandsdifferenz zueinander in einem selbst konstruierten Array befestigt (siehe Bild 1-1).

Die digitale Signalaufzeichnung der einzelnen Anordnungen erfolgt simultan, mit gleichem „virtuellen Mittelpunkt“ zur Quelle. Dabei werden für jedes Mikrofon die gleichen hochwertigen A/D-Wandler in 96 kHz/24 bit Qualität benutzt und keine Dynamik- oder Klangbearbeitungen vorgenommen.

Zudem werden verschiedene Raumpositionen der Quellen aufgenommen, wobei nur die Quelle im Raum bewegt wird, nicht aber die Mikrofonanordnung.

Die Beurteilung der Aufnahmen erfolgt mit Hilfe festgelegter Begriffsdefinitionen nach Vorschlägen der EBU Tech. 3286-E rein subjektiv, wobei nicht nur die möglichst perfekte natürliche Abbildung des Hörereignisses, sondern auch die praktische Verwendbarkeit und der Grundcharakter der Aufnahmen diskutiert werden.

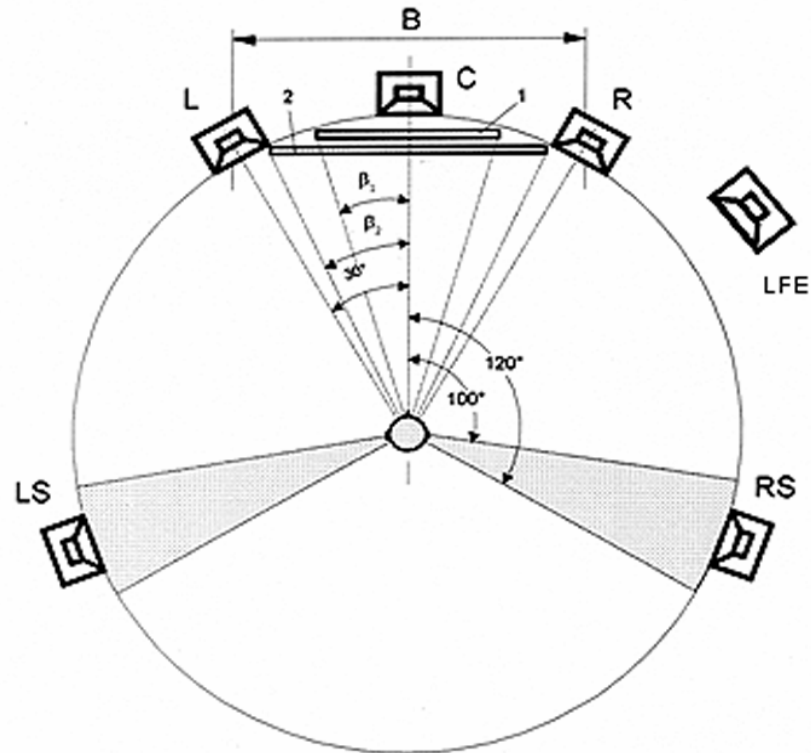


**Bild 1-1 Frontansicht des Mikrofon-Arrays, schematisch**

In dieser Arbeit wird bewusst das 5.1-Format als Mehrkanalformat behandelt, da dieses Format auf Produktions-, wie auch auf Wiedergabeseite, das kommerziellste und weit verbreitetste Format darstellt (siehe Bild 1-2).

Ebenso wird bewusst nicht vom 3/2-Stereo-Format gesprochen, da die fünf Kanäle und der Subwoofer „gleichberechtigt“ behandelt werden. Die in der Literatur als Surround- oder Effektkanäle bezeichneten beiden hinteren Lautsprecher (entsprechend Mikrofone) LS und RS werden in dieser Arbeit als Rear-Lautsprecher (entsprechend Rear-Mikrofone) bezeichnet, und die drei vorderen als Front-Lautsprecher (entsprechend Front-Mikrofone).

Diese Diplomarbeit beinhaltet nicht unbedingt einen Mikrofontest der „klassischen“ Art, dessen Hauptkriterium die exakte naturgetreue Abbildung eines Instrumentes oder Orchesters ist. Sie unterstützt vielmehr den Ansatz in dem Bereich der Pop-Musikproduktion etwas „Neues“ sowohl von künstlerischer, wie auch produktionstechnischer Seite, entstehen zu lassen.



**Bild 1-2 Referenz-Lautsprecher-Anordnung fürs 5.1-Format, vgl. ITU-R BS.775-1 [16]**

akustisches Zentrum	Winkel	Höhe	Neigung
C	0°	1,2 m *)	0° *)
L, R	+/- 30°	1,2 m	0°
LS, RS	+/- (100...120)°	> 1,2 m	≤ 15°

**Bild 1-3 Winkel und Maßangaben der Lautsprecher-Anordnung nach Empfehlungen der ITU-RBS.775-1, aus [16]**

## 2. Psychoakustische Grundlagen

### 2.1 Das Schallfeld in geschlossenen Räumen

Gibt eine Schallquelle in einem geschlossenen Raum ein Signal ab, baut sich ein Schallfeld auf, deren Schallwellen von den Wänden des Raumes und von Raumobjekten reflektiert werden. Ein Zuhörer in diesen Raum stellt ebenfalls ein Raumobjekt dar, den die Schallwellen aus verschiedenen Richtungen und zu verschiedenen Zeiten erreichen.

Der Direktschall der Quelle erreicht den Zuhörer ohne Umwege zuerst, und danach die ersten Reflexionen. Diese werden in kurzer Zeit immer dichter und bilden schließlich den Nachhall. Durch das räumliche Hören (siehe Kapitel 2.2) können wir nun mit der Auswertung des Schallfeldes Informationen über den Raum erlangen.

Das Schallfeld besteht also aus einzelnen Komponenten, welche unterschiedliche Auswirkungen auf den Höreindruck des Zuhörers haben.

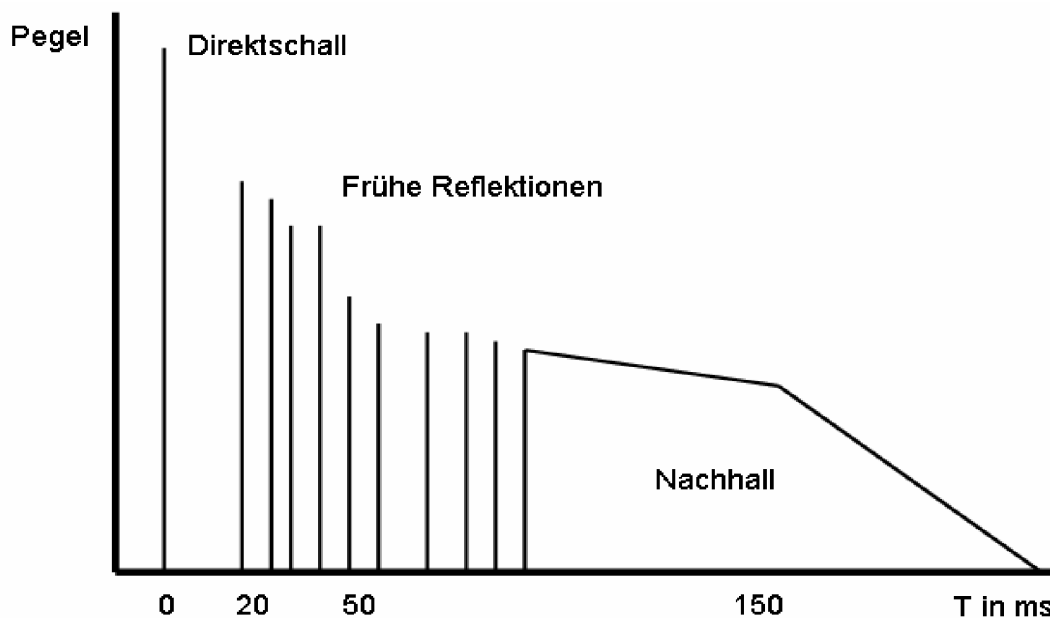


Bild 2-1 Aufbau des idealen Schallfeldes, vgl. [7]

### 2.1.1 Direktschall

Der Direktschall ist die erste Komponente des Schallfeldes, die den Zuhörer erreicht und ist deshalb richtungsbestimmend. Er vermittelt bei musikalischen Ereignissen den Eindruck einer räumlichen Staffelung der einzelnen Instrumente und ist hauptverantwortlich für die „Durchsichtigkeit“ oder „Transparenz“ des Klangbildes.

Die Richtungsinformationen gewinnt das Gehör durch die Auswertung von Unterschieden der beiden Ohrsignale. In erster Linie von interauralen Zeitdifferenzen (interaural time differences - ITD) und interauralen Pegeldifferenzen (interaural intensity differences - IID), wobei das Hörereignis bei gleichen Signalen immer in Richtung des nicht verzögerten und lautereren Signals wandert.

Auch die Auswertung der Außenohrübertragungsfunktion (Head-Related-Transfer-Function -> HRTF-Funktion), also der spezifischen Änderungen, die ein Signal durch den Einfluss des Kopfes und des Außenohres erfährt, ist für die Richtungsbestimmung und fürs räumliche Hören entscheidend.

Man spricht von der *Lokalisation* eines Schallereignisses, im Gegensatz zur aktiven *Ortung*, wie sie z.B. beim Radargerät oder von Fledermäusen benutzt wird. Dazu mehr im Kapitel 2.2..

In recht geringem Ausmaß wird der Direktschall bei Schallfeldern in geschlossenen Räumen auch zur Bestimmung der Schallquellenentfernungen verwendet. Durch die absorbierende Wirkung von Luft kommt es zu einem charakteristischen Höhenverlust und der Lautstärkepegel nimmt bei Entfernungsverdoppelung von der Schallquelle um 6 dB ab, so dass in größeren Räumen Entfernungswahrnehmungen entstehen können. Im Freien treten diese Wahrnehmungen sehr viel häufiger auf [11].

### 2.1.2 Frühe Reflexionen

Die frühen Reflexionen haben je nach Schallereignis verschiedene Auswirkungen auf den Höreindruck, die von ihrer Verzögerungszeit, Intensität und Auslenkung abhängig sind. Der Wert, bis zu welchem Zeitpunkt nach dem Direktschall die frühen Reflexionen als solche betrachtet werden, variiert in der Literatur zwischen 50 und 80 ms.

Sie entscheiden hauptsächlich sowohl über Deutlichkeit und Klarheit des Klangbildes, wie auch über die Wahrnehmung der Räumlichkeit.

Besonders wichtig ist in erster Linie der erste Rückwurf, also die erste Reflexion. Je später er nach dem Direktschall beim Hörer eintrifft, desto größer wird der Raum empfunden. Eine Verzögerung von 30 ms entspricht einem Weg von 10 m, was nur in größeren Räumen möglich ist. Je kürzer sie ist, desto „intimer“ ist der Raumeindruck und desto näher fühlt man sich dem Schallereignis [6].

Je nach Schallereignis wirkt der erste Rückwurf auch direktschallverstärkend, ohne dabei die Deutlichkeit zu beeinflussen. Bei Orchestermusik kann er eine Verzögerungszeit von bis zu 30 ms betragen. Ab 35 ms kann diese Reflexion jedoch als störendes Echo empfunden werden.

Alle späteren Reflexionen werden im Bereich von 50 - 80 ms je nach Schallereignis als direktschallzugehörig wahrgenommen (Sprache bis 50 ms, Orchester bis 80 ms) [7].

Die Einfallsrichtung hat auf die Wahrnehmung der frühen Rückwürfe auch Auswirkungen. Dieser als *interaurale Korrelation* beschriebene Effekt [6] bewirkt bei Schallereignissen, welche direkt aus der Medianebene kommen ( $0^\circ$  von vorne (siehe Kapitel 2.2)), nur eine größere Lautstärkeempfindung des Direktsignals. Je weiter seitlich die Reflexionen eintreffen, desto geringer wird der Wert der interauralen Korrelation und desto mehr wird beim Hörer eine Wahrnehmung des Raumes deutlich.

Neben der Wahrnehmung der Räumlichkeit wird auch der Effekt der Aufweitung des Hörereignisses, welcher in der Literatur häufig als ASW bezeichnet wird (Apparent Source Width), durch die frühen seitlichen

Reflexionen hervorgerufen. Hierbei nimmt die Schallquelle ab einem Lautstärkepegel von ca. 75 – 85dB [9] akustisch einen größeren Raum ein als optisch.

Neben dem Lautstärkepegel ist auch die Zeit des Eintreffens der seitlichen Reflexionen und der Frequenzgang für die Wahrnehmung des Effektes verantwortlich. Sie müssen während der Einschwingzeit der Töne beim Hörer eintreffen, was also nur bei nicht percussiven Signalen bzw. in grossen Räumen möglich ist [8]. Zudem entsteht bei Frequenzen oberhalb von 3 kHz eher der Eindruck einer Verbreiterung der Klangquelle, bei tieferen Frequenzen eher eine Aufweitung in die Tiefe.

### **2.1.3 Späte Reflexionen und Nachhall**

Die späten Reflexionen ab ca. 150 ms nach dem Direktschall bilden den so genannten *Nachhall*.

Bei ausreichend lautem Nachhall fühlt sich der Hörer von den Hörereignissen eingehüllt und der Eindruck einer Lautstärkeerhöhung des Gesamtereignisses macht sich bemerkbar [6].

Für den Umhüllungseffekt ist, wie schon bei den frühen Reflexionen, die seitliche diffuse Komponente des Nachhalls besonders wichtig.

Die Umhüllung entsteht durch Schwankungen der interauralen Zeit- und Pegeldifferenzen. Diese Schwankungen entstehen durch Interferenzen von direktem und reflektiertem Schall.

Schwanken sie mit einer Frequenz von bis zu 3 Hz, werden sie als Schallquellenbewegungen interpretiert. Ab 3 Hz aufwärts wirken sie umhüllend. Bei Signalfrequenzen unter 300 Hz sind Schwankungen der interauralen Zeitdifferenzen erforderlich. Bei Frequenzen über 300 Hz sind sowohl die interauralen Zeit- wie auch Pegeldifferenzen erforderlich.

D. Griesinger [8] bezeichnet den Nachhall als sogenannten „Background Stream“, welcher vom Gehör getrennt vom „Foreground Stream“ wahrgenommen wird.



Die beiden „Streams“ können vom Gehör miteinander verglichen werden, so dass z.B. Informationen über Lage und Entfernung der Schallquelle gewonnen werden können.

Die Stärke der Umhüllung ist von der absoluten Lautstärke des „Background Streams“ abhängig, und nicht von dem Verhältnis von direkter zu zurückgeworfener Energie (R/D-Balance). Hat der Nachhall bei einem Schallereignis einen bestimmten Pegel, wirkt er bis zu einem gewissen Grad umhüllend. Steigen nun Direktschall und Nachhall im gleichen Verhältnis an, so entsteht alleine durch den höheren Nachhallpegel ein größerer Umhüllungseindruck.

Der Grad der empfundenen Umhüllung ist aber auch vom Schallereignis abhängig. So wird er bei musikalischen Darbietungen oft nur in den Pausen bemerkt, jedoch bei Sprache und einzelnen Schallquellen fast ununterbrochen. Das liegt daran, dass das Ohr nach dem Ende des Ereignisses eine gewisse Zeitspanne benötigt, um den „Background Stream“ wahrzunehmen. Bei Sprache beträgt diese Zeitspanne ca. 50 ms, bei Orchestermusik ca. 350ms [8].

## 2.2 Räumliches Hören

### 2.2.1 Lokalisation

Die *Lokalisation* wird nach J. Blauert [1] als „Zuordnungsgesetz oder –regel (Operator) zwischen dem Ort eines Hörereignisses (z.B. bezüglich Richtung und/oder Entfernung) und einem bestimmten Merkmal oder bestimmten Merkmalen eines Schallereignisses oder eines anderen, mit dem Hörereignis korrelierten Ereignisses. [...]“ definiert.

Die Unterscheidung der Begriffe *Schall-* und *Hörereignis* ist dabei wichtig, da das *Hörereignis* das vom Hörer wahrgenommene Ereignis beschreibt, und das *Schallereignis* den physikalischen Vorgang des Hörens.

In Experimenten mit Kopfhörerbeschallung ermittelte Blauert [1] physikalische Werte für die Auslenkung von Hörereignissen.

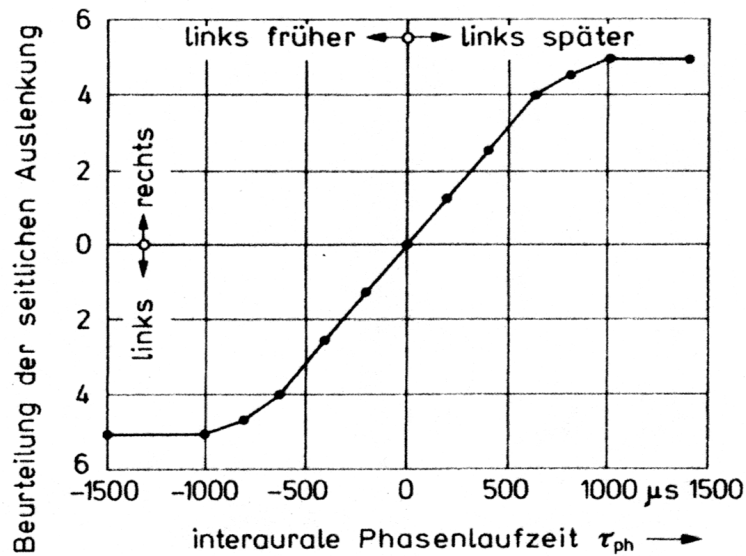
Da die Hörereignisse bei Kopfhörerbeschallung direkt am Kopf bzw. im Kopf entstehen, spricht man hier von „Lateralisation“ [1] im Gegensatz zur „Lokalisation“ bei Lautsprecherwiedergabe, wo das Hörereignis außerhalb des Kopfes stattfindet.

Diese „Lateralisationsexperimente“ geben laut Theile [13] „[...] Aufschlüsse über die Auswertung interauraler Signalunterschiede [...], jedoch gibt es keine abgesicherte Hypothese für die Verallgemeinerung dieser Erkenntnisse auf das räumliche Hören“. Sie sind daher eher als Modellvorstellung für das Auftreten von Hörereignissen mit seitlichen Auslenkungen geeignet.

Blauert [1] fand heraus, dass ein Signal, das keine signifikanten Anteile oberhalb von 1,6 kHz enthält, hauptsächlich durch interaurale Zeitdifferenzen (ITD) ausgewertet wird. Zudem wird ein Signal, wenn es an beiden Ohren den gleichen Pegel hat, in die Richtung desjenigen Ohres ausgelenkt, bei dem es früher eintrifft.

Bis zu einer Differenz von 630  $\mu$ s verläuft die Auslenkung linear, danach flacht die Kurve ab, bis bei ca. 1 ms Laufzeitunterschied die maximale Auslenkung

erreicht ist (siehe Bild 2-2).

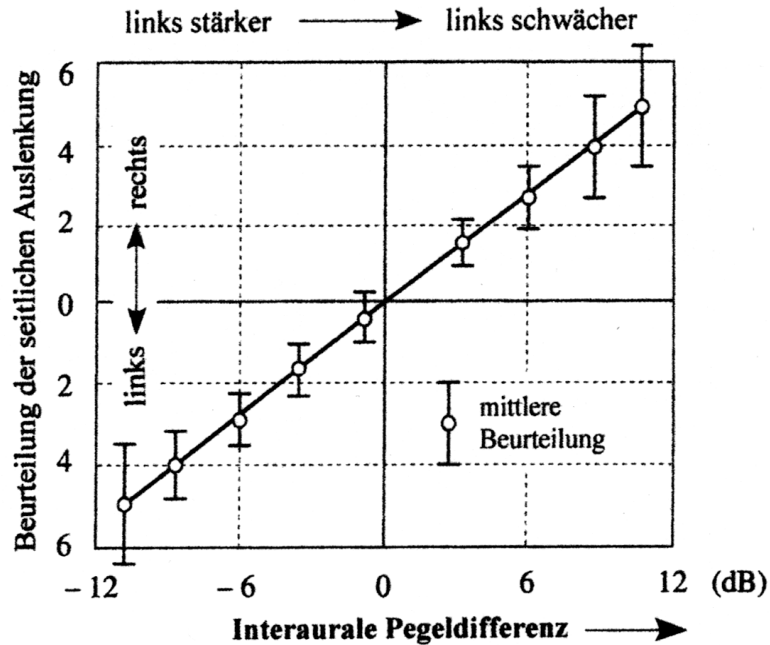


**Bild 2-2 Seitliche Auslenkung des Hörereignisses als Funktion der ITD, aus [1]**

Enthält das Signal vorwiegend Anteile über 1,6 kHz, werden hauptsächlich die interauralen Pegeldifferenzen (IID) ausgewertet. Besteht keine Laufzeitdifferenz zwischen den Signalen an den Ohren, dann wird das Hörereignis in Richtung desjenigen Ohres ausgelenkt, bei dem das Signal stärker ist. Eine Angabe des für Maximalauslenkung benötigten Pegelunterschiedes ist schwierig, da ab einer Differenz von ca. 8 – 10 dB das Hörereignis „breiter“ wird, also nicht mehr so scharf „laterisiert“ ist [1] (siehe Bild 2-3). Erschwerend kommt hinzu, dass die für eine bestimmte seitliche Auslenkung erforderliche Pegeldifferenz frequenzabhängig ist. Sie ist bei ca. 2 kHz am geringsten und steigt für höhere und tiefere Frequenzen an. Die in der Literatur angegebenen Werte für die maximale seitliche Auslenkung bei Breitbandrauschen liegt bei ca. 10 dB Unterschied der beiden Ohrsignale.

Diese Ergebnisse sind rein experimentell und die in den Versuchen verwendeten Signale kommen im natürlichen Schallfeld kaum vor. Daher hat das Ohr fast immer die Möglichkeit, Zeit- und Pegeldifferenzen gleichzeitig

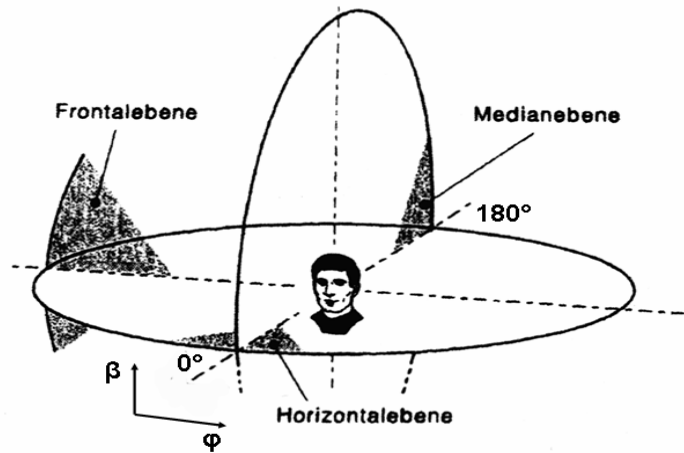
auszuwerten. Auch anhand der Tatsache, dass Pegelunterschiede im natürlichen Schallfeld frequenzabhängig sind, weil hohe Frequenzen bei seitlichem Einfall stärker vom Kopf abgeschattet werden, kann das Ohr weitere Richtungsinformationen gewinnen.



**Bild 2-3 Seitliche Auslenkung des Hörereignisses als Funktion der IID [1]**

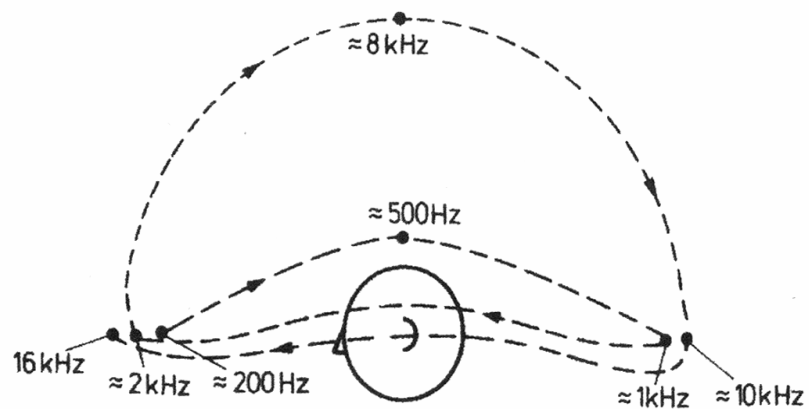
Ähnlich funktioniert die Unterscheidung, ob Schall von vorne bzw. von hinten kommt. Befindet sich die Schallquelle z.B. hinter dem Hörer, dann erfährt ihr Signal bestimmte lineare Verzerrungen, mit deren Hilfe das Gehör die Schalleinfallrichtung bestimmen kann. Ein Spezialfall ist, wenn sich die Schallquelle in der *Medianebene* befindet, der Ebene, die den Kopf genau in der Mitte zwischen den Ohren halbiert (siehe Bild 2-4).

Dem Gehör fehlen in dieser Ebene zur Bewertung des Schallereignisses die wesentlichen interauralen Merkmale zwischen den Ohren, so dass vor allem die jeweiligen monauralen Merkmale der beiden Ohren bewertet werden.



**Bild 2-4 Das Kopfbezogene Koordinatensystem, vgl.[10]**

J. Blauert [1] beobachtete ebenfalls, dass sich ein Hörereignis, welches er direkt vor einem Lautsprecher sitzend in Form eines sich von 200 Hz bis 16 KHz erhöhenden Sinussignals abspielte, mehrmals auf einer Bahn über seinen Kopf hinweg von vorn über oben nach hinten und zurück bewegte (siehe Bild 2-5).



**Bild 2-5 Bahn der Hörereignisrichtung, stark schematisiert, aus [1]**

In weiteren Versuchen mit Terzbandrauschen wurden Personen von vorne, hinten und oben aus der Medianebene beschallt. Es ließen sich „richtungsbestimmende Bänder“ [1] ermitteln, woraus zu schließen war, dass das Gehör die Hörereignisrichtung auf Grund der Frequenz der Ohrsignale (Schallsignale in den Gehörgängen) bestimmt. Schallereignisse in der Medianebene werden demnach entsprechend ihres spektralen Anteils geortet. Die Frequenzbänder um 1 kHz und 10 kHz werden als Schalleinfall von hinten wahrgenommen, um 8 kHz als Schalleinfall von oben und 250 bis 500 Hz und ca. 2,5 bis 4 kHz als Schalleinfall von vorne.

### **2.2.2 Lokalisationsunschärfe**

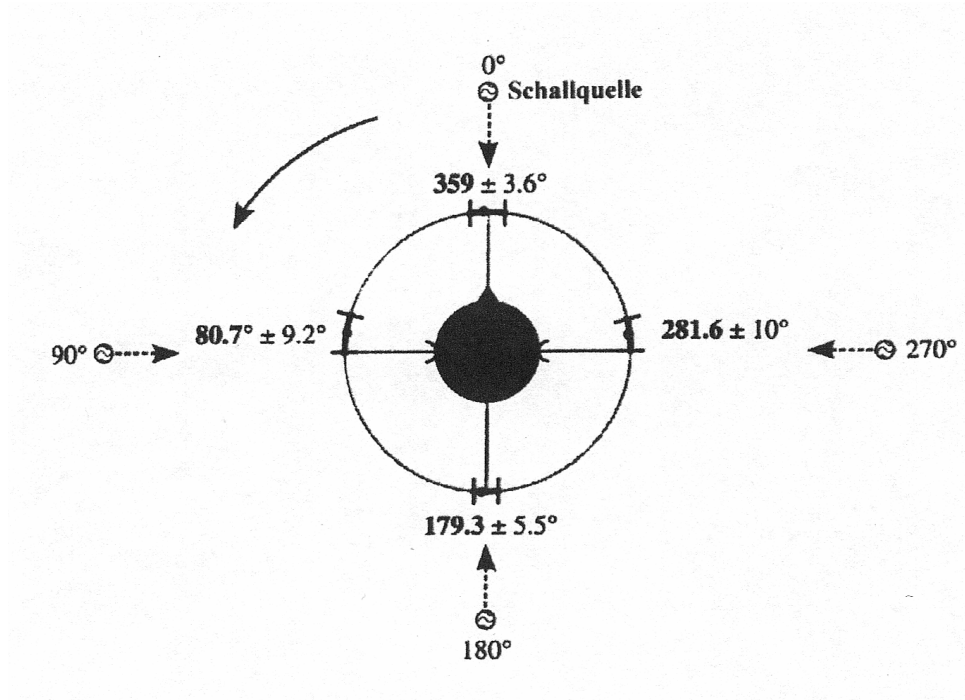
Die *Lokalisationsunschärfe* ist von Blauert [1] als „[...] die kleinste Änderung eines bestimmten Merkmals oder bestimmter Merkmale des Schallereignisses oder eines anderen, mit dem Hörereignis korrelierten Ereignisses, die gerade zu einer Ortsänderung des Hörereignisses (z.B. bezüglich Richtung und/oder Entfernung) führt.“ definiert.

In der Horizontalebene kann das Ohr die Richtungswahrnehmungen sehr viel besser auflösen als in der Medianebene. Besonders gut zeichnet sich der Bereich mit  $\beta = \varphi = 0^\circ$  aus, d.h. vorne mittig, wobei die größten Unschärfen über dem Kopf auftreten.

So liegt z.B. die Lokalisationsunschärfe bei Breitbandrauschen in der „Vorne“-Position bei  $\pm 3,6^\circ$ , vergrößert sich dann mit zunehmendem Winkel bis  $\varphi = 90^\circ$  auf ca.  $\pm 10^\circ$  und nimmt dann bis  $\varphi = 180^\circ$  wieder ab auf ca.  $\pm 5,5^\circ$  (siehe Bild 2-6).

Insgesamt hängt die Lokalisationsunschärfe von vielen Parametern ab, vor allem von der Signalbeschaffenheit wie Art, Frequenz und Dauer, aber auch vom Alter der Versuchspersonen.

Einen Spezialfall der *Lokalisationsschärfe* definiert der Begriff *Abbildungsschärfe*, der die Lokalisationsschärfe von Phantomschallquellen bei Lautsprecherwiedergabe beschreibt (siehe Kapitel 2.4.1).



**Bild 2-6 Lokalisationsunschärfe in der Horizontalebene, aus [12]**

### 2.2.3 Prezedenzeffekt

Einen ebenfalls wichtigen Einfluss auf die Wahrnehmung der Hörereignisrichtung hat der Prezedenzeffekt (auch „Hass-Effekt“ oder „Gesetz der ersten Wellenfront“ genannt). Er bewirkt, dass zeitlich vorseilende Signale gegenüber zeitlich nacheilenden bevorzugt werden, auch wenn jene einen höheren Pegel aufweisen.

Dieser Effekt ist signalabhängig und ab einer Laufzeitdifferenz von ca. 2 ms wirksam. Ab diesem Wert ändert sich bei Erhöhung der Differenz kaum etwas in Bezug auf die Wahrnehmung von Testsignalen. Erst ab einer Differenz von 20 ms setzt der Echoeffekt ein (das später eintreffende Signal wird getrennt wahrgenommen), der eine Veränderung der Wahrnehmung bewirkt.

Bei der Lautsprecherstereofonie wird dieser Begriff ebenfalls zur Richtungsdarstellung von Klangkörpern benutzt [14].

## 2.3 Lokalisation bei Lautsprecherwiedergabe

### 2.3.1 Summenlokalisation

Wird ein Signal über Lautsprecher wiedergegeben, entstehen 2 verschiedene Arten von Schallquellen: *Ersatz- bzw. Realschallquellen* und *Phantomschallquellen*.

*Ersatzschallquellen* werden genau an dem Ort lokalisiert, wo sie sich befinden, also direkt am Lautsprecher. Bei der Stereowiedergabe existieren demnach zwei Ersatzschallquellen L und R, die im Vergleich zu Phantomschallquellen sehr präsent erscheinen und präziser lokalisierbar sind [9]. Bei der 5.1-Wiedergabe existieren 5 Ersatzschallquellen; der 0.1-Subwoofer ist durch sein tiefpassgefiltertes Signal unterhalb von 120 Hz kaum lokalisierbar.

*Phantomschallquellen* (siehe Kap. 2.4) befinden sich zwischen den Lautsprechern, so dass also mindestens 2 Lautsprecher für deren Entstehung verfügbar sein müssen.

Der wahrgenommene Ort der Phantomschallquellen entsteht durch *Summenlokalisation*, welche nach Blauert [1] dann vorliegt, wenn „ [...] zwei oder mehrere Schallquellen kohärente oder teilkohärente Signale abstrahlen und nur ein Hörereignis auftritt, dessen räumliche Attribute von den Orten und abgestrahlten Signalen aller beteiligten Schallquellen abhängen.“.

Die *Summenlokalisation* tritt auf, wenn geringe Pegel- und Zeitunterschiede zwischen den von den Lautsprechern abgestrahlten Signalen vorliegen. Die Theorie sagt aus, dass die Signale an den Lautsprechern sich genau so addieren, dass die resultierenden Ohrsignale bei Beschallung denen einer realen, am Ort der Phantomschallquelle befindlichen, Schallquelle gleichen [1].

Diese Theorie ist jedoch nach Theile [18] nur für schmalbandige Signale anwendbar. Da jedoch gerade bei der Musikwiedergabe eher breitbandige



Signale in einem überlagerten Schallfeld vorherrschen, hat Theile ein „Assoziationsmodell“ vorgeschlagen, mit dem die psychoakustischen Eigenschaften des menschlichen Gehörs bei der Lokalisation im überlagerten Schallfeld erklärt werden.

### **2.3.2 Assoziationsmodell nach Theile**

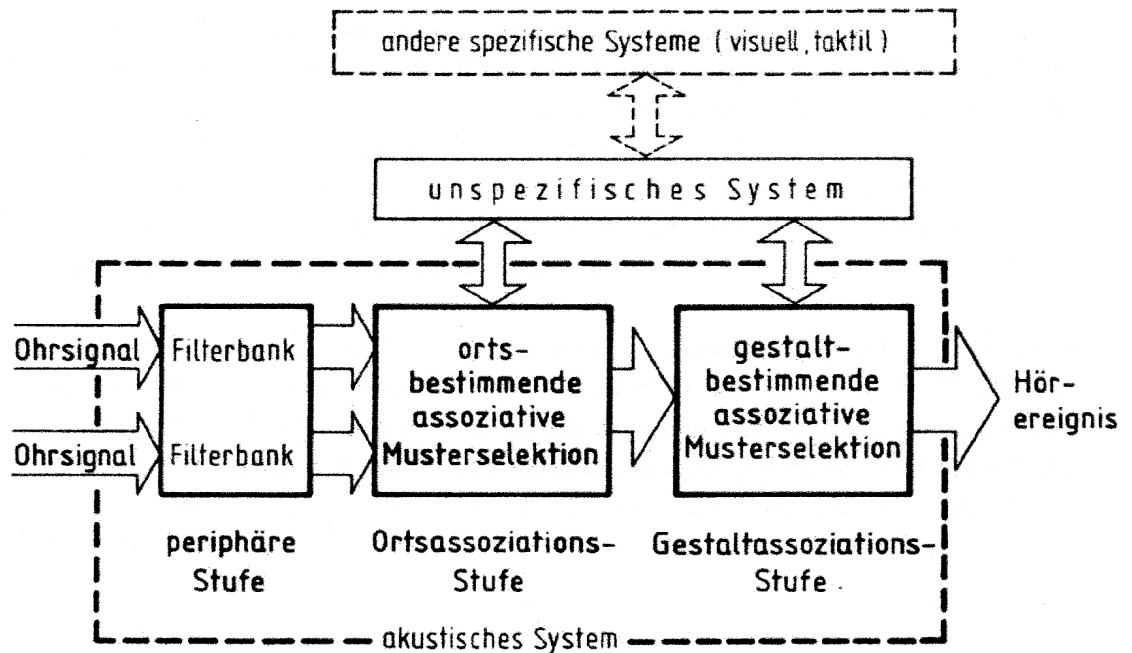
Das *Assoziationsmodell* [18] beruht auf dem Grundprinzip, dass bei der Lokalisation von bestimmten Signalen Assoziationsvorgänge eine wichtige Rolle spielen. Dabei nehmen zwei Assoziationsstufen unterschiedliche Aufgaben getrennt voneinander wahr (siehe Bild 2-7):

Die *Ortsassoziationsstufe* vergleicht bestimmte Signalanteile des Gehörten mit erlernten Mustern, die jeweils durch Erfahrungen bestimmten Schallquellenorten zugeordnet sind. Lokalisationsreize kommen zustande. Die resultierenden Hörereignisorte werden in einem Mustervorrat durch bekannte Merkmale wie Laufzeitunterschiede und spektrale Zusammensetzung unterschieden. Eine Selektion der Lokalisationsreize durch eine Art steuerbares Filter, welches spektral invers zur Wirkung des Außenohres in der erkannten Hörereignisrichtung arbeitet, erlaubt es, verschiedene Signale voneinander zu trennen und mehrere Lokalisationsreize gleichzeitig zu verarbeiten. Nach der Entfernung dieser „richtungsabhängigen Klangfärbungen“ und Kammfiltereffekte werden die Signale an die gestaltbestimmende Assoziationsstufe weitergeleitet.

In der *Gestaltassoziationsstufe* werden die Signale unabhängig von ihren räumlichen Informationen inhaltlich ausgewertet. Dies geschieht wieder durch den Vergleich verschiedener gespeicherter Muster, die durch Hörerfahrungen gewonnen wurden.

Sind die beiden Ohrsignale hinreichend ähnlich, verschmelzen sie zu einer „gemeinsamen mittleren räumlichen Information“ [18], einer Phantomschallquelle.

Ist ein Signal wesentlich lauter und/oder eilt dem anderen erheblich voraus, so wird hauptsächlich dieses Signal zur Auswertung herangezogen und das andere weitestgehend ausgeblendet. (*Lokalisationsreizdominanz*).  
 Werden die beiden Signale an gleichen Orten lokalisiert, haben aber unterschiedliche Gestalt, so werden sie getrennt voneinander ausgewertet.



**Bild 2-7 Das Funktionsprinzip des Assoziationsmodells, aus [18]**

## 2.4 Parameter einer Phantomschallquelle

Die schon im Kapitel 2.3.1 erwähnten Phantomschallquellen entstehen, wenn „[...] mehrere Schallereignisse einem gemeinsamen Hörereignis zugeordnet sind, so dass Hörereignisort und Schallquellenorte nicht übereinstimmen.“ [18].

Die Entstehungsart der Phantomschallquelle bestimmt hauptsächlich die unterschiedlichen Eigenschaften wie Abbildungsschärfe, Entfernungseindruck und Klangfarbe. Dabei können durch unterschiedliche Arten entstandene Phantomschallquellen am selben Ort große Unterschiede aufweisen.

Bild 2-8 zeigt Beispiele einer auf verschiedene Arten hergestellte Phantomschallquellen bei einer Position von etwa  $-7,5^\circ$  im Frontbereich (L-C-R) einer 5.1-Anordnung.

In der Produktionspraxis geschieht dies z.B. bei einzelnen Signalen mit Hilfe von Panoramaeinstellungen am Mischpult (meist Pegeldifferenzen) oder Hinzufügen von Delay- und Halleffekten (Laufzeitdifferenzen).

Basis	Art der Signaldifferenz(-en)	Signaldifferenz(-en)
L-R	Pegelunterschied	$\Delta L(L/R) = 3 \text{ dB}$ ; $\Delta t(L/R) = 0$ ;
L-R	Laufzeitunterschied	$\Delta L(L/R) = 0 \text{ dB}$ ; $\Delta t(L/R) = -0,2 \text{ ms}$ ;
L-R	komb. Pegel- und Laufzeitunterschiede	$\Delta L(L/R) = 1,5 \text{ dB}$ ; $\Delta t(L/R) = -0,1 \text{ ms}$ ;
L-C	Pegelunterschied	$\Delta L(L/C) = -6 \text{ dB}$ ; $\Delta t(L/C) = 0$ ;
L-C	Laufzeitunterschied	$\Delta L(L/C) = 0 \text{ dB}$ ; $\Delta t(L/C) = 0,4 \text{ ms}$ ;
L-C	komb. Pegel- und Laufzeitunterschiede	$\Delta L(L/C) = -3 \text{ dB}$ ; $\Delta t(L/C) = 0,2 \text{ ms}$ ;
L-C-R	komb. Pegel- und Laufzeitunterschiede	$\Delta L(L/C) = -3 \text{ dB}$ ; $\Delta t(L/C) = 0,2 \text{ ms}$ ; $\Delta L(C/R) = 6 \text{ dB}$ ; $\Delta t(C/R) = 1,0 \text{ ms}$ ;
L-C-R	komb. Pegel- und Laufzeitunterschiede	$\Delta L(L/R) = 1,5 \text{ dB}$ ; $\Delta t(L/R) = -0,1 \text{ ms}$ ; $\Delta L(L/C) = 9 \text{ dB}$ ; $\Delta t(L/C) = -1,0 \text{ ms}$ ;

**Bild 2-8 theoretische Aufstellung der möglichen Signaldifferenzen für die Entstehung einer Phantomschallquelle bei  $-7.5^\circ$ , aus [14]**

### **2.4.1 Abbildungsschärfe**

Der Begriff *Abbildungsschärfe* beschreibt die Lokalisationsschärfe einer Phantomschallquelle zwischen den Lautsprechern, deren Optimum die Wiedergabe in einem einzelnen Lautsprecher darstellt. Sie kennzeichnet die räumliche Auflösung einer Aufnahme und ist neben der Entstehungsart auch von der Hörposition und der Art des Signals abhängig.

Bei durch Pegeldifferenzen entstandenen Phantomschallquellen nimmt sie mit wachsendem Pegelunterschied zu.

Bei durch Laufzeitdifferenzen entstandenen Phantomschallquellen nimmt sie mit steigendem Zeitunterschied ab.

Durch die Kombination der beiden Arten kann sie erhöht werden, da für die selben Auslenkungen gegenüber den laufzeitbasierten Auslenkungen ein geringerer Zeitunterschied notwendig ist.

### **2.4.2 Klangfarbe**

Die Wahrnehmung von Phantomschallquellen existiert in natürlichen Schallfeldern nicht. Die bei der Lautsprecherwiedergabe auftretenden Phänomene wie frequenzabhängige Auslöschungen/Addierungen von Signalanteilen, wie z.B. Kammfiltereffekte, fallen daher schnell auf und machen sich als *Klangfarbenveränderung* bemerkbar. Starke Abweichungen von der Klangfarbe führen auch zur Beeinträchtigung der Wahrnehmung des Hörereignisortes [14].

Bei verkleinerten Lautsprecherbasen können die Wahrnehmungen von Klangfarbenveränderungen stark zunehmen, so dass diese z.B. beim 5.1.-Standard durch die Halbierung der L/R-Basis durch den zusätzlichen Center-Kanal unter bestimmten Umständen in erhöhtem Maße auftreten können. Zudem lässt der zusätzliche Centerkanal ein „dreikanaliges Klangbild“ entstehen, d.h. dass die Phantomschallquelle durch die 3 Lautsprecherbasen LC, RC und LR dreimal dargestellt wird, was Klangfarbenbeeinträchtigungen zur Folge haben kann [14] (siehe Kapitel 3.2).

### **3 .Hauptmikrofontechniken**

In diesem Abschnitt werden verschiedene bestehende Hauptmikrofontechniken erläutert, sowohl stereophone wie auch mehrkanalige.

Der Begriff „*Hauptmikrofon*“ wird in der Literatur oft verschiedenartig verwendet und umschrieben, doch es werden prinzipiell folgende Forderungen an ein Hauptmikrofon gestellt, sowohl für stereophone wie auch für mehrkanalige Verfahren:

- Es soll den Raumeindruck abbilden, indem es in angenehmer Balance sowohl die Reflektionen mit Nachhall, wie auch den Direktschall aufnimmt (R/D-Balance).
- Es soll eine Richtungsabbildung der Klangquellen erzeugen, so dass bei der Wiedergabe für den Zuhörer ein Richtungseindruck der einzelnen aufgenommenen Quellen möglich wird.

Um den Forderungen möglichst gerecht zu werden, haben sich im Laufe der Jahre viele verschiedene Techniken bewährt, die für viele verschiedene Aufnahmesituationen entwickelt wurden. Hier werden nur die Grundprinzipien der wichtigsten Hauptmikrofontechniken erläutert.

#### **3.1 Stereophone Hauptmikrofonverfahren**

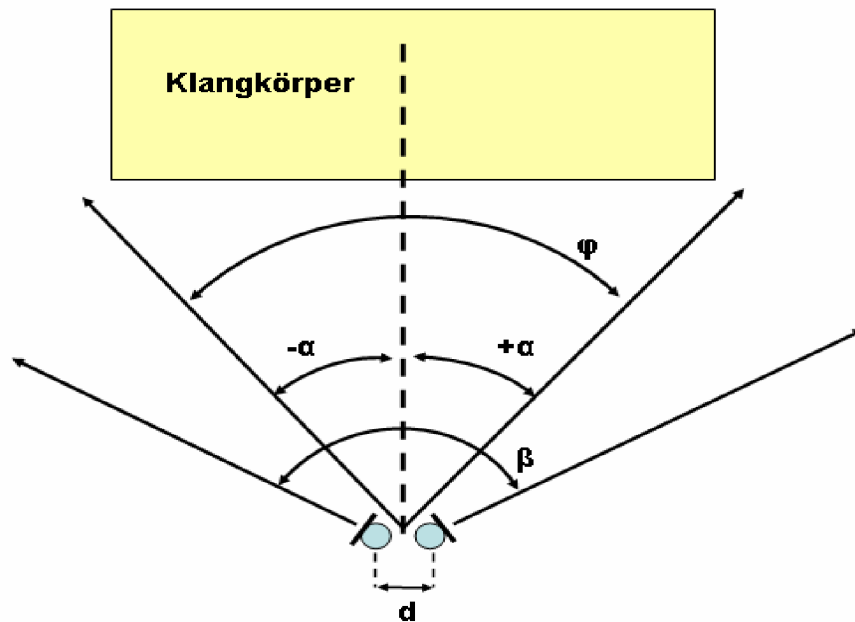
Bevor die Anordnungen besprochen werden, sind noch einige Begriffsdefinitionen zu klären.

Der *Aufnahmewinkel* entspricht dem Winkelbereich, innerhalb dessen die Schallquellen so aufgenommen werden, dass sie zwischen den beiden Lautsprechern lokalisiert werden. Alle Ereignisse die links und rechts außerhalb des Bereiches liegen, werden nicht mehr getrennt abgebildet und überlagern

sich im jeweiligen Lautsprecher. In der Literatur wird dieser Winkel oft mit  $\pm\alpha$  bezeichnet und soll hier der Einfachheit halber nur  $\varphi$  genannt werden.

Der *Hauptachsenwinkel*  $\beta$  entspricht hier dem Winkel, den die beiden Mikrofone bezogen auf ihre senkrechten Mittelachsen beim entgegengesetzten Auseinanderdrehen nach Außen bilden.

Der *Abstand* zwischen den beiden Mikrofonskapseln wird hier  $d$  genannt.



**Bild 3-1** Aufnahmewinkel  $\varphi$  und Hauptachsenwinkel  $\beta$

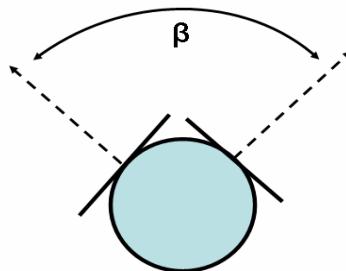
### 3.1.1 Intensitätsstereophonie

Bei der „*Intensitätsstereophonie*“ werden ausschließlich Pegeldifferenzen zwischen den Mikrofonen aufgenommen.

Erreicht wird dies, indem man richtende Mikrofone koinzident, d.h. unmittelbar übereinander, anordnet, und sie gegeneinander in einem bestimmten Winkel aus der Hauptachse verdreht (siehe Bild 3-2). Eine solche Aufstellung wird auch *XY-Anordnung* genannt.

Bestimmend für den Aufnahmewinkel  $\varphi$  der Anordnung sind der Hauptachsenwinkel  $\beta$  und die Richtcharakteristika der verwendeten Mikrofone.

Auf der Wiedergabeseite sind die Kennzeichen solcher Aufnahmen eine gute Lokalisation und Abbildungsschärfe der Klangquellen. Durch die hohe Korrelation zwischen den Kanälen ist der Raumeindruck jedoch oft unbefriedigend, da ein großer Teil des Diffusfalls in Mono wiedergegeben wird [3].



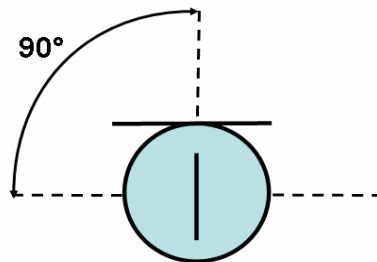
**Bild 3-2 XY-Anordnung schematisch**

Eine weitere Variante der „Intensitäts-Stereophonie“ stellt die *MS-Anordnung* dar. Hier werden ein Mikrofon mit beliebiger Richtcharakteristik (M) und ein Mikrofon mit Achtercharakteristik (S) koinzident angeordnet. Die  $0^\circ$  Richtungsachsen der beiden Mikrofone sind jedoch um  $90^\circ$  gedreht, so dass das M ein Mittensignal erzeugt und S ein Seitensignal, wobei die Ausrichtung von S der Ausrichtung der Gesamtanordnung entspricht (siehe Bild 3-3).

Durch eine geeignete Matrizierung ergeben sich aus den M- und S-Signalen die für den linken und den rechten Kanal. Bei Addition von M+S erhält

man das L-Signal und bei Subtraktion M-S das R-Signal.

Die Pegeldifferenz zwischen den Kanälen für seitlichen Schalleinfall ist abhängig von der gewählten Charakteristik des Mitten-Mikrofons und des Mischungsverhältnisses der beiden Mikrofone. Der Charakter der Aufnahmen ist daher von den Verhältnissen dieser einzelnen Parameter abhängig und kann bewusst nachträglich beeinflusst werden.



**Bild 3-3 MS-Anordnung schematisch**

### 3.1.2 Laufzeitstereophonie

Eine Auslenkung von Phantomschallquellen zwischen den Lautsprechern kann auch durch reine Laufzeitunterschiede zwischen 2 Mikrofonen erzeugt werden. Dieses wird erreicht, indem 2 Mikrofone, räumlich von einander getrennt, in einer nicht koinzidenten *AB-Anordnung* nebeneinander aufgestellt werden (siehe Bild 3-4).

Bestimmend für den Aufnahmewinkel  $\varphi$  des Systems ist der Abstand  $d$  zwischen den beiden Mikrofonkapseln und der Winkel  $\beta$  zwischen den beiden Hauptachsen der Kapseln. Der Abstand  $d$  liegt in der Regel zwischen 40-80 cm oder mehr und  $\beta$  variiert zwischen 0 und 90° [3].

Vergrößert man den Abstand  $d$  zwischen den Mikrofonen, so verkleinert sich der Aufnahmewinkel  $\varphi$ , so dass bei einem bestimmten großen  $d$  keine Abbildung mehr zwischen den Lautsprechern stattfinden kann. Man spricht von

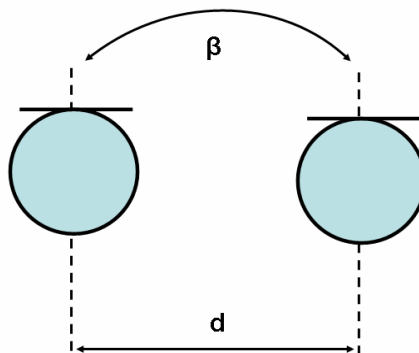


einem „Mittenloch“, bei dem die Hörereignisse größtenteils entweder dem einen oder dem anderen Lautsprecher zugeordnet sind [5].

Häufig werden für die AB-Anordnung Kugelmikrofone benutzt, die auf Grund ihrer nicht idealen Richtcharakteristik parallel ( $\beta = 0^\circ$ ) angeordnet werden sollten, um möglichst frequenzabhängige Pegelunterschiede, welche zu Klangfarbenverfälschungen führen können, zu vermeiden und Einbußen bei der Lokalisationsschärfe von Phantomschallquellen zu verhindern [4].

Bei Benutzung von gerichteten Mikrofonen z.B. Nieren, kann der Winkel  $\beta$  durchaus Werte bis  $90^\circ$  annehmen, wobei man dann schon von Äquivalenzstereophonie (siehe Kapitel 3.1.3) spricht.

Zu den guten Eigenschaften der Laufzeitstereophonie zählen hauptsächlich der sehr gute Raumeindruck mit räumlicher Tiefe und die gute Übertragung tiefer Frequenzen im Vergleich zur Intensitätsstereophonie, wobei die Lokalisation nicht so präzise und stabil ist.



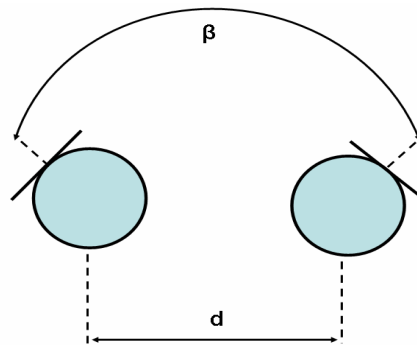
**Bild 3-4 AB-Anordnung schematisch**

### **3.1.3 Äquivalenzstereophonie**

Nutzt man mit einer Mikrofonanordnung gleichsinnig Laufzeit- und Pegelunterschiede aus, so spricht man bei dieser Technik von *Äquivalenzstereophonie*. Gleichsinnig bedeutet in diesem Zusammenhang,

dass die Laufzeit- und Pegeldifferenzen in die gleiche Richtung führen und sich deren Wirkungen überlagern [1].

Dieses wird erreicht, indem man 2 richtende Mikrofone sowohl räumlich von einander trennt, als auch deren Richtungen gegeneinander nach außen verdreht (siehe Bild 3-5).



**Bild 3-5 Beispiel einer Äquivalenzanordnung, schematisch**

Typische Werte sind hierbei ein Mikrofonabstand  $d$  von 5-30 cm und ein Hauptachsenwinkel  $\beta$  von 0-180°. Deren Abhängigkeiten zum Aufnahmewinkel  $\varphi$  der Gesamtanordnung, sind in den so genannten „Williamskurven“ (siehe Bild 3-6) aufgeführt, durch die in der Praxis einfach und schnell das gewünschte  $\varphi$  ermittelt werden kann.

Eine häufig benutzte Beispielaufstellung ist die *ORTF-Anordnung*, bei der sich bei einem Mikrofonabstand  $d$  von 17,5 cm und einem Hauptachsenwinkel von 110° ein Aufnahmewinkel von ca.  $\varphi = 100^\circ$  ergibt.

Durch die Vermischung der beiden Stereophoniearten können sich deren positive Eigenschaften, wie einerseits eine gute Richtungsabbildung und andererseits eine gute Abbildung des Raumeindrucks, zu einem guten Kompromiss führen. Da dem Gehör zwei Arten von Lokalisationsreizen zur Verfügung stehen, ist die Abbildungsschärfe und Stabilität dieser Anordnung besonders groß.

## Hauptachsenwinkel $\beta$

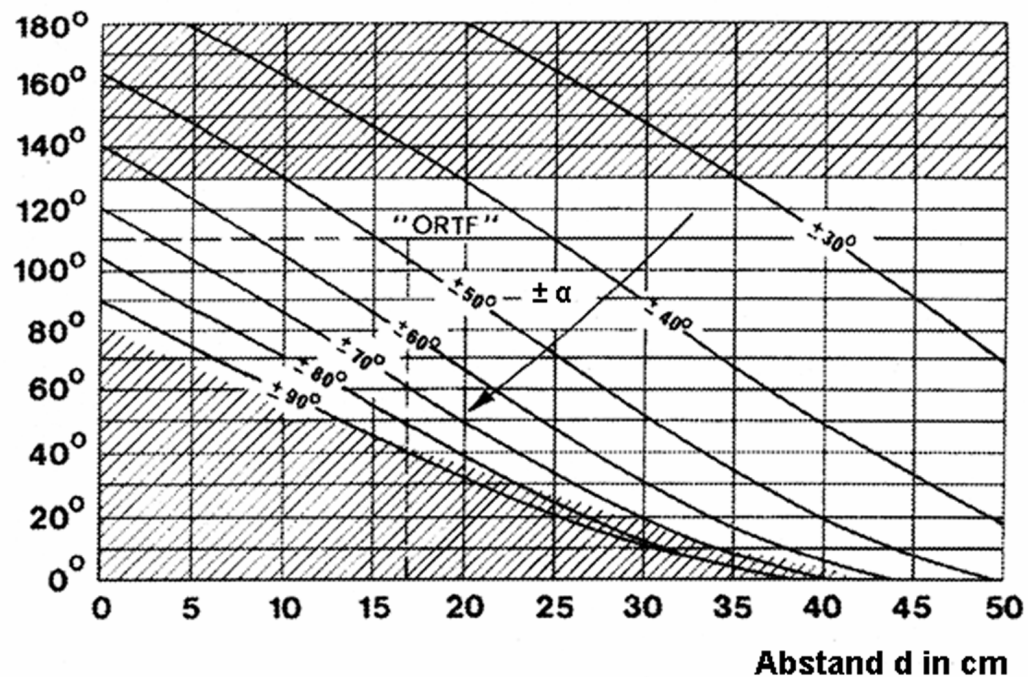


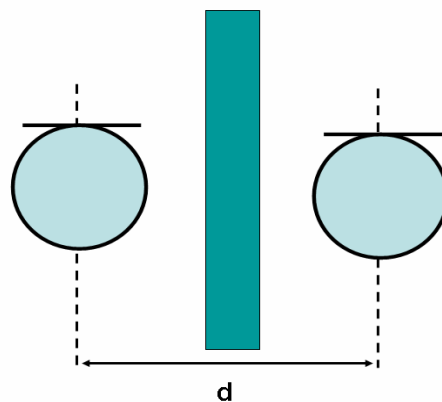
Bild 3-6 "Williams Kurven", vgl. [3]

### 3.1.4 Trennkörperstereophonie

Bei der *Trennkörperstereophonie* befindet sich zwischen zwei nicht koinzident angeordneten Mikrofonen mit Kugelcharakteristik (Druckempfänger) ein fester Trennkörper (siehe Bild 3-7). Dadurch werden neben den Laufzeitunterschieden auch Pegelunterschiede zwischen den Empfängern provoziert, um eine verbesserte Lokalisation zu erhalten.

Die an der Trennfläche auftretenden Schallreflexionen und -beugungen führen jedoch zu unterschiedlichen Effekten, die bei der Wiedergabe mit Lautsprechern trotz Entzerrungsmaßnahmen als unangenehme Klangverfälschungen auftreten können. Zudem ergeben sich in der Aufnahmepraxis Einschränkungen durch den in der Regel festen und nicht besonders variablen Aufnahmebereich der Anordnungen [4].

Einen Spezialfall der Trennkörperstereophonie stellt die *Kunstkopfstereophonie* dar (siehe Bild 3-8). Der Trennkörper ist hierbei der Form eines menschlichen Kopfes nachgebildet und die Mikrofone befinden sich in einem quasi Gehörgang an den Ohrpositionen. Diese Nachbildung der „menschlichen Hörsituation“ ist hauptsächlich für die Schallwiedergabe mit Kopfhörern optimiert, kann aber auch bei Lautsprecherwiedergabe gute Ergebnisse erzielen.



**Bild 3-7 Beispiel einer Trennkörperstereophonie-Anordnung, schematisch**



**Bild 3-8 Kunstkopfanordnung Neuman KU-100, Prospekt-Foto**

## 3.2 Mehrkanalverfahren

Bei dem 5.1-Format werden, wie schon am Anfang beschrieben, 5 Lautsprecher bestehend aus 3 Front- und 2 Rearlautsprechern und ein Subwoofer verwendet. Die Benutzung von 5 Mikrofonen zur Herstellung eines 5.1-Hauptmikrofons ist daher sehr naheliegend. Doch auch andere Techniken sind möglich und haben sich bewährt.

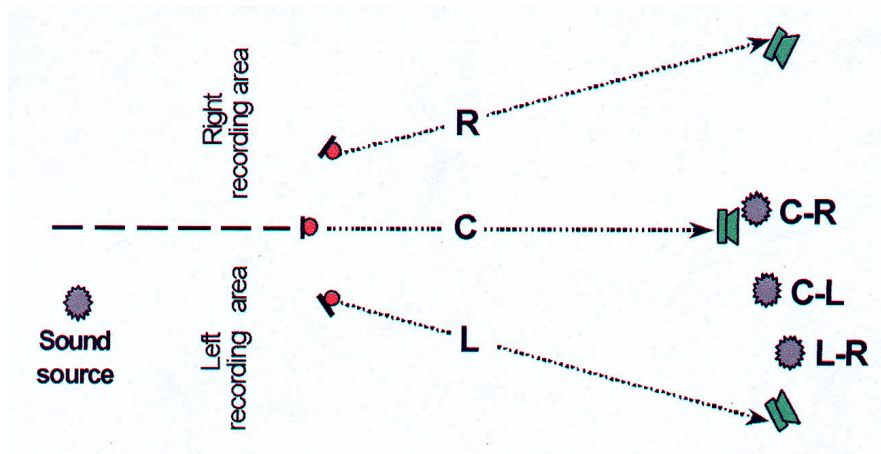
In dieser Arbeit werden die wichtigsten, zur Zeit auf dem Markt erhältlichen und in der Praxis verwendeten 5.1- bzw. 3/2-Stereo-Hauptmikrofonanordnungen beschrieben und in 3 Hauptkategorien aufgeteilt:

- Zweiteilige Systeme bestehend aus *L-C-R-Hauptmikrofon* mit zusätzlichem *Atmomikrofon*
- *Multimikrofonsysteme*: Komplettsysteme bestehend aus 5 oder mehr Mikrofonen in einem festen Array (ggf. mit zusätzlichen Kugelmikrofonen zur Bassanhebung)
- Auf *MS-Technik* basierende *Systeme*

Die aus der Stereotechnik bekannten Regeln und Erkenntnisse lassen sich nur bedingt auf diese Systeme übertragen. Der zusätzliche Centerkanal kann eine bessere Richtungsstabilität und damit auch eine größere Abhörzone im Gegensatz zur Zweikanaltechnik bieten [13]. Zudem kann die Abbildungsstabilität und die Klangfarbe der Phantomschallquellen unter bestimmten Umständen verbessert werden [14].

Der zusätzliche Centerkanal kann jedoch in einer Hauptmikrofonanordnung auch problematisch werden. Denn wie im Bild 3-9 ersichtlich, erzeugen nicht nur die Mikrofonpaare L/C und C/R eine Phantomschallquelle, sondern auch das Paar L/R. Es ergeben sich also 3 Phantomschallquellen, obwohl nur eine erwünscht ist. Dieses Phänomen kann die beschriebenen Vorteile, wie schon im Kapitel 2.4.2. angedeutet, negativ beeinträchtigen, so dass sich die Wiedergabe von Abbildungsschärfe und Klangfarbe verschlechtern kann.

Die Vorteile können also nur bei ausreichender Kanaltrennung wirklich genutzt werden, so dass idealerweise eine Schallquelle im linken Bereich nicht vom rechten Mikrofon mit aufgezeichnet wird [13].



**Bild 3-9 Dreifache Phantomschallquellenbildung, aus [13]**

Die Forderungen an ein Mehrkanal-Hauptmikrofon sind im Grunde genommen mit denen der Stereophonen identisch. Die Möglichkeiten der Umsetzung sind durch die zusätzlichen Kanäle um einiges gestiegen und damit auch anspruchsvoller. Verschiedene und neue Ansätze der Signalaufnahme sind entstanden und müssen bezogen auf ihre Anwendung untersucht werden.

### 3.2.1 L-C-R-Systeme

Die Benutzung eines L-C-R-Systems für das 5.1.-Format erfordert immer ein zusätzliches Atmomikrofon, um alle Wiedergabekanäle abzudecken (siehe Kapitel 3.2.3). Das L-C-R-System sollte dabei möglichst den Direktschall des Ereignisses aufnehmen, das Atmomikrofon, wie der Name schon sagt, den Diffusschall bzw. die indirekten Schallanteile.

Zudem sollte, wie schon erwähnt, eine möglichst hohe Kanaltrennung erreicht werden, so dass die Probleme der dreifachen Phantomschallquellenbildung und

deren negative Beeinflussungen auf die Abbildungsschärfe und die Klangfarbe, vermieden werden können [5].

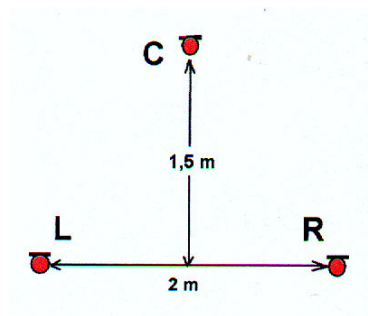
### Decca-Tree

Die als „Decca-Tree“ bekannte Mikrofonanordnung wurde schon in den 50’er und 60’er Jahren von der Firma DECCA in Großbritannien für stereophone Zweikanalaufnahmen von klassischer Musik entwickelt. Hierbei wurden hauptsächlich Kugelmikrofone in relativ großem Abstand zueinander in Form eines Dreiecks aufgestellt (siehe Bild 3-10).

Der Hauptachsenwinkel  $\beta$  des L und R Mikrofons variierte je nach Aufnahme-situation zwischen 0 und 90° und auch die Länge der Hypotenusen des Dreiecks zwischen 1,1 bis 1,6m [19].

Durch das vorhandene Centermikrofon ist dieses System auch für die heutige 5.1-Anwendung geeignet und liefert ein stabiles Klangbild mit offenem, räumlichem Charakter.

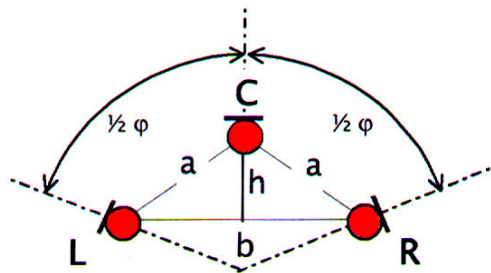
Bedingt durch die weiten Abstände der Mikrofone untereinander findet jedoch keine konkrete Abbildung zwischen den Lautsprechern statt. Ohne zusätzliche Stützmikrofone werden quasi nur 3 Hörereignisse, die den 3 Lautsprechern L/C/R zugeordnet sind, wahrgenommen, aber mit hoher Stabilität.



**Bild 3-10 Beispiel für Decca-Tree-Anordnung, aus [13]**

### INA 3

Das Verfahren INA (*Ideale Nieren Anordnung*) ist in den 90'er Jahren von U. Hermann und V. Henkels [5] entwickelt worden. Sie gehen dabei von einer zweiteiligen Betrachtung der stereophonen Teilzone L/C und R/C aus, so dass jeweils 2 Mikrofonpaare in einer Dreiecksanordnung den halben Aufnahmewinkelbereich  $\frac{1}{2} \varphi$  abdecken (siehe Bild 3-11). Zudem wird festgelegt, dass der Hauptachsenwinkel  $\beta$  dem halben Aufnahmewinkel  $\frac{1}{2} \varphi$  entspricht. Somit wird eine Überlappung der beiden Aufnahmebereiche in der Mitte vermieden und auch die Kanaltrennung, welche durch die Verwendung von gerichteten Nierenmikrofonen gegenüber den Kugeln beim Decca-Tree verbessert werden soll, ist recht gut.



**Bild 3-11 INA 3-Anordnung, vgl.[5]**

Nach den bekannten Williamskurven können nun die Abstände zwischen den Mikrofonen in Abhängigkeit vom gewünschten Aufnahmewinkel  $\varphi$  ermittelt werden. Würde man also einen Aufnahmewinkel  $\varphi$  von  $120^\circ$  wünschen, so ergeben sich die Abstände  $a = 53$  cm,  $b = 92$  cm und  $h = 27$  cm [5].

### „Doppel-Stereo“ nach Gernemann

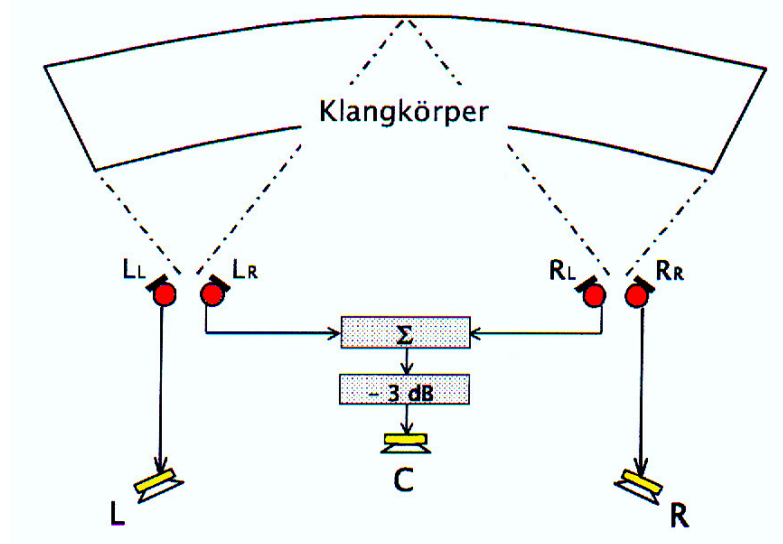
Das „Doppel-Stereo“ System nach Gernemann besteht im Prinzip aus zwei separaten zweikanaligen Hauptmikrofonen, die den Gesamtaufnahmebereich in 2 Teilbereiche aufteilen. Das Centersignal wird durch Addition der beiden



inneren Mikrofonsignale gewonnen, und dessen Pegel muss entsprechend um 3 dB gesenkt werden (siehe Bild 3-12).

Die Überlappung der beiden Anordnungen muss, so gut es geht, vermieden werden, indem der Aufnahmewinkel der beiden Anordnungen und deren Abstand zueinander variiert wird. Erst wenn die beiden Systeme durch ihren Abstand und ihrer Ausrichtung als unabhängig gelten, kann die maximale Abbildungsschärfe erreicht und ein Übersprechen des einen Systems aufs andere vermieden werden.

Daraus wird deutlich, dass diese Anordnungen nur bei großen Klangkörpern sinnvoll sind.



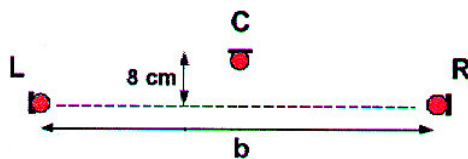
**Bild 3-12 „Doppel-Stereo“ nach Gernemann, vgl. [13]**

### OCT nach Theile

Das auf der Grundlage der in Kapitel 2.3.2 beschriebenen psychoakustischen Prinzipien entwickelte OCT-System (Optimised Cardioid Triangle) von Günther Theile stellt ebenfalls eine Dreiecksanordnung dar. Die Basis L/R ist mit dem Abstand  $b$  jedoch um ein Vielfaches größer als die mit  $h$  bezeichnete Mittelsenkrechte. Sie beträgt 8 cm, wobei  $b$  in der Praxis Werte zwischen 40

und 80 cm annimmt (Bild 3-13).

Als L- und R-Mikrofon werden zur noch besseren Kanaltrennung Super-Nieren-Mikrofone verwendet, die entgegengesetzt mit einem jeweiligen Hauptachsenwinkel von  $\beta = 90^\circ$  angeordnet sind. Da die Super-Nieren-Mikrofone nachweislich den tiefen Frequenzbereich schwächer übertragen, werden häufig zusätzliche Kugeln auf den L und R Positionen angebracht. Die Signale der Kugeln werden ab 100 Hz tiefpassgefiltert und nach Gehör den L und R Signalen beigemischt. Häufig wird das Signal des Nierenmikrofons für den Centerkanal ab 100 Hz hochpaßgefiltert, so dass die tiefen Frequenzen zur Verbesserung des Klangbildes nur von L und R übertragen werden [13]. Das OCT-System besticht durch seine gute Abbildungsschärfe, Klangfarbentreue und hohe Stabilität bei der Lokalisation von einzelnen Hörereignissen, was auch nicht zuletzt durch die speziell entwickelten Mikrofone der Firma Schoeps möglich wird.



**Bild 3-13 OCT-Anordnung ohne Kugeln, aus [13]**

### 3.2.2 Atmo-Mikrofonanordnungen

Der Begriff „Atmomikrofon“ steht für vierkanalige Raummikrofone, die in Verbindung mit einem dreikanaligen LCR-Hauptmikrofon sehr variable Aufstellmöglichkeiten bieten.

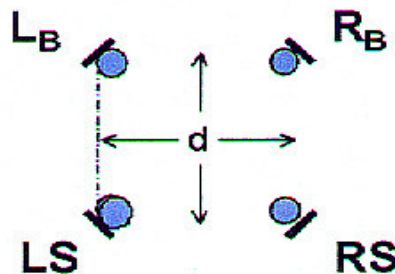
Dabei sollen sie ausschließlich diffuse Schallanteile aufnehmen, die den Raumeindruck erhöhen sollen.

Die beiden als  $L_B$  und  $R_B$  bezeichneten Mikrofone werden dabei den Frontmikrofonen  $L$  und  $R$  hinzugemischt und die  $L_S$  und  $R_S$  Mikrofone entsprechen den beiden Rear-Kanälen.

Wegen der fehlenden festen Beziehung zum Frontmikrofon, können die Atmomikrofone in Bereichen untergebracht werden, wo sowohl ein optimaler Raumeindruck vorherrscht, als auch Störgeräusche (z.B. aus dem Publikum) wenig ins Gewicht fallen. Der Anwendungszweck ist hauptsächlich für große Konzertaufnahmen bestimmt.

### IRT-Kreuz

Das von Günther Theile entwickelte „IRT-Kreuz“ kann grundsätzlich mit Mikrofone von Nieren- bis Kugelcharakteristik bestückt werden. Dabei ist bei einem Abstand von  $d = 20 - 25$  cm mit Nierenmikrofonen eher eine realistische und relativ direkte Abbildung des Klangbildes zu erwarten und bei einem  $d = 60$  cm mit Kugeln ein eher diffuses und „verschmiertes“ Klangbild [11]. Durch die geringen Ausmaße und das geringes Gewicht ist es auch für mobile Einsätze gut geeignet.

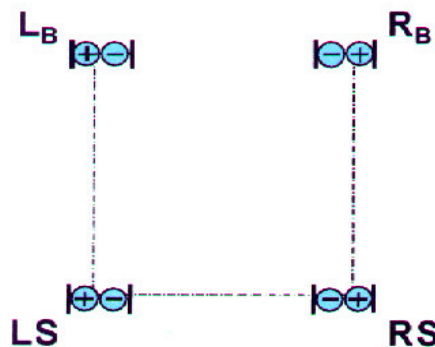


**Bild 3-14 IRT-Kreuz , aus [13]**

## Hamasaki-Square

Das sogenannte "Hamasaki-Quadrat" besteht aus vier Mikrofonen mit Achter-Charakteristik, die, mit mehreren Metern Abstand zur Schallquelle, in einer quadratischen Anordnung aufgestellt werden. Sie sind dabei so ausgerichtet, dass ihre Nullebenen zu den Musikern zeigen. Die Seitenlängen des Quadrates können je nach Anwendung und Raumgröße zwischen 1 - 4 m variieren (siehe Bild 3-15).

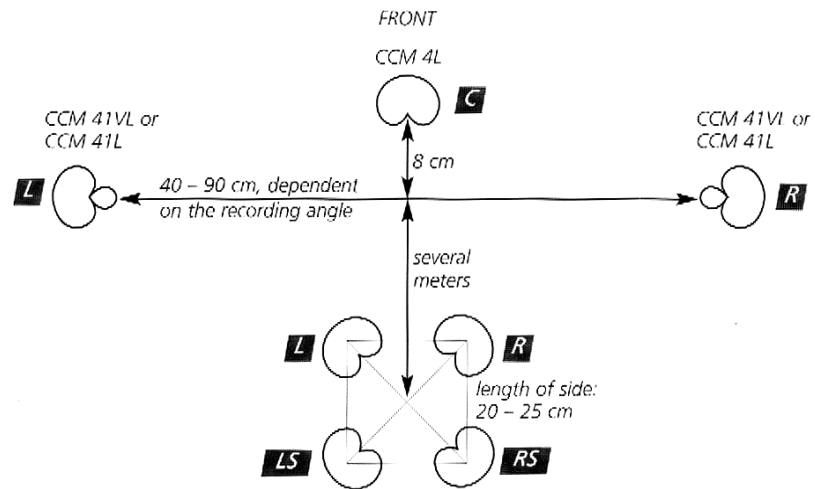
Auf diese Weise wird fast nur indirekter Schall aufgezeichnet, und durch die Größe des Quadrates entstehen in den seitlichen Segmenten AB-Stereobasen. Zusätzlich wird die Polarität der rechten Mikrofone gegenüber der der linken umgekehrt, um die Korrelation zwischen linker und rechter Hälfte weiter zu verringern. Das Resultat ist eine sehr räumliche und homogene Wiedergabe des diffusen Schallfeldes.



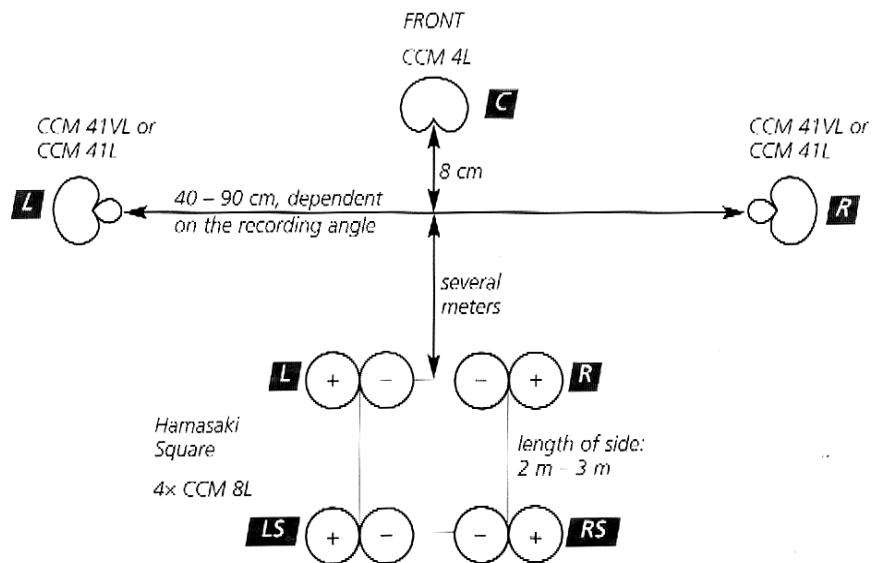
**Bild 3-15 Hamasaki-Square, aus [13]**

## Kombinationen

Die angesprochenen Kombinationen von LCR-Frontmikrofonanordnungen und Atmomikrofon ist in den folgenden zwei Bildern 3-16 und 3-17 als Beispiel ersichtlich.



**Bild 3-16 Kombination OCT mit IRT-Kreuz, aus [16]**



**Bild 3-17 Kombination OCT mit Hamasaki-Square, aus [16]**

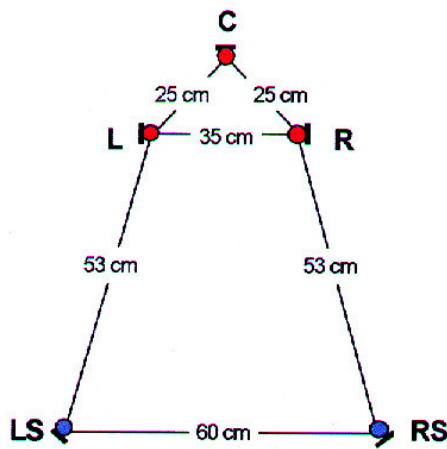
### 3.2.3 Multimikrofon-Systeme

Als Multimikrofonssysteme werden hier Mehrkanal-Hauptmikrofonanordnungen beschrieben, die sowohl die Front-, wie auch die Rearkanäle, in einer nicht koinzidenten Anordnung mit einbeziehen.

Somit werden keine zusätzlichen externen Mikrofone benötigt und eine einfachere Handhabung, in manchen Fällen aber auch eine geringere Flexibilität, gewährleistet.

#### INA 5

Die INA 5 Anordnung besteht im Kern aus einer INA 3 Anordnung mit einem Aufnahmewinkel von  $180^\circ$  und zusätzlichen Rear-Mikrofonen LS und RS. Die Entwickler Hermann und Henkels [5] gehen davon aus, dass eine Abbildung zwischen allen Lautsprechern stattfinden soll und somit ein Gesamtaufnahmewinkel von  $360^\circ$  optimal wäre. Mit Hilfe der Williamskurven wurden die Abstände der Mikrofone berechnet, so dass das Frontdreieck mit einer Basis von  $b = 35 \text{ cm}$  recht klein ist. Die Rear-Mikrofone sind mit einem Winkel von  $\pm 150^\circ$  bezogen auf die  $0^\circ$  Senkrechte ausgerichtet in einem Abstand von  $60 \text{ cm}$  zueinander (siehe Bild 3-18).



**Bild 3-18 INA 5 Anordnung, aus [13]**

Der Anordnung wird eine mäßige Praktikabilität bei Orchester-Musikaufnahmen in großen Räumen bescheinigt, da der Aufnahmewinkel von  $180^\circ$  in der Front theoretisch eine sehr geringe Entfernung zum Klangkörper bedingen würde. Zudem sollen seitliche Phantomschallquellen recht instabil abgebildet werden, so dass eine Richtungsabbildung, wie in der Front, nicht möglich ist [13]. Inwieweit diese Probleme bei der Aufnahme einzelner Instrumente bzw. kleiner Klangkörper in relativ kleinen Räumen ins Gewicht fallen, werden die Versuchsaufnahmen mit dem Atmos 5.1.-System zeigen, welches eine INA 5 Anordnung darstellt (siehe Bild 3-19).



**Bild 3-19 INA 5 System "Atmos 5.1" von der Firma SPL, Prospektfoto**



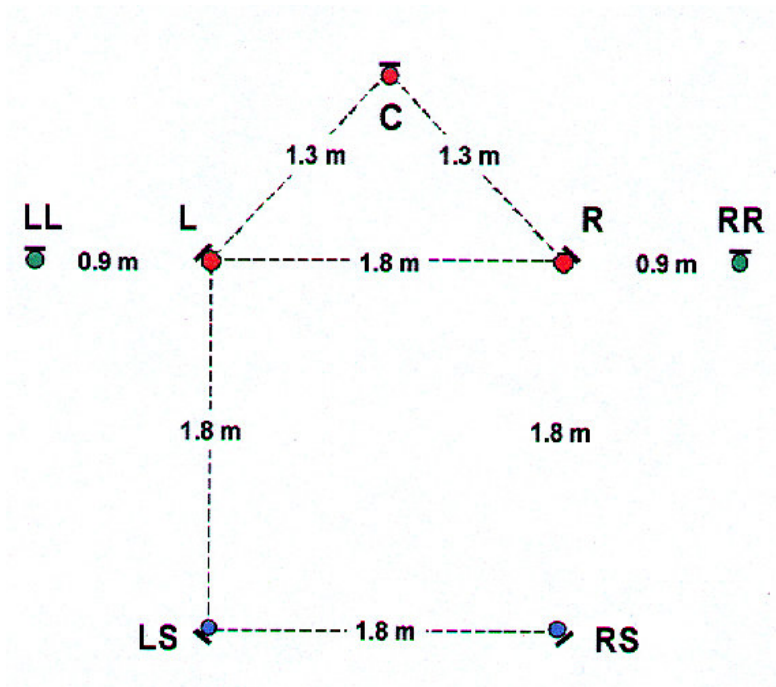
**Bild 3-20 INA 5 System der Firma Microtech Gefell, Prospektfoto**

## Fukada-Tree

Der von A. Fukada entwickelte Fukada-Tree stellt im Prinzip eine Weiterentwicklung und 5.1-Optimierung des Decca-Trees dar. Die drei Frontmikrofone sind mit einer L/R-Basisbreite von  $b = 1,80$  m ähnlich wie beim Decca-Tree relativ weit auseinander, wobei hier alle drei Kugeln durch Nierenmikrofone ersetzt wurden. Die Mikrofone L und R sind mit einem Hauptachsenwinkel von  $\beta = \pm 45^\circ$  ausgerichtet, ebenso wie die Rear-Mikrofone LS und RS, welche wiederum  $1,80$  m von den Frontmikrofonen entfernt sind (siehe Bild 3-21).

Diese Anordnung ist durch seine Ausmaße hauptsächlich für große Räume konzipiert, wobei die durch Nieren ersetzten Frontmikrofone den Pegel der diffusen Raumanteile gering halten sollen.

Wie es sich in relativ kleinen Räumen und bei einzelnen Instrumenten verhält, werden die Versuchsaufnahmen zeigen.



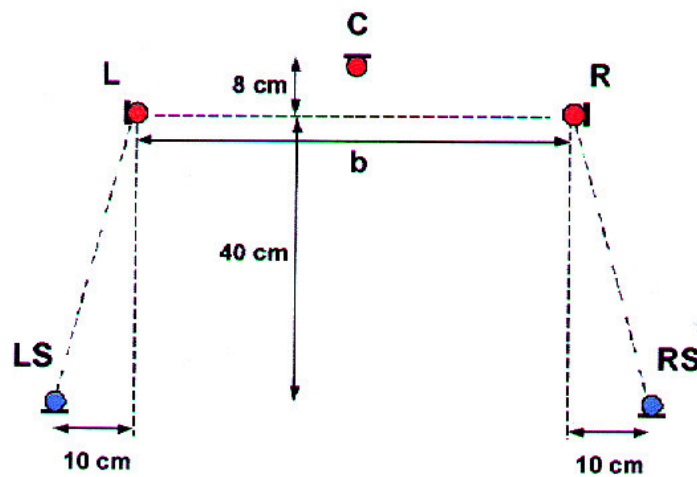
**Bild 3-21 Fukada-Tree Anordnung, aus [13]**



## OCT-Surround

Das OCT-Surround-System ist eine auf 5 Mikrofone erweiterte OCT-Anordnung. Als zusätzliche Mikrofone sind 2 Rearmikrofone LS und RS (Nierencharakteristik) hinzugekommen, die exakt um  $180^\circ$  entgegengesetzt zum Centermikrofon ausgerichtet sind (siehe Bild 3-20).

Sie sollen damit so wenig Direktschall wie möglich aufnehmen (bei Schalleinfall von Vorne) und sollen in Kombination mit den L und R Mikrofonen im Idealfall nur die frühen Reflektionen aus den seitlichen und hinteren Bereichen übertragen [13].



**Bild 3-22 OCT-Surround-Anordnung, aus [13]**

Diese Anordnung ist, wie viele andere auch, trotz der relativ geringen Ausmaße auf die möglichst naturgetreue Aufnahme von Orchestern in großen Räumen optimiert. Wie es sich in kleinen Räumen verhält bei relativ dichten Positionen zur Schallquelle, wird in den Versuchsaufnahmen deutlich werden.

## Williams MMA

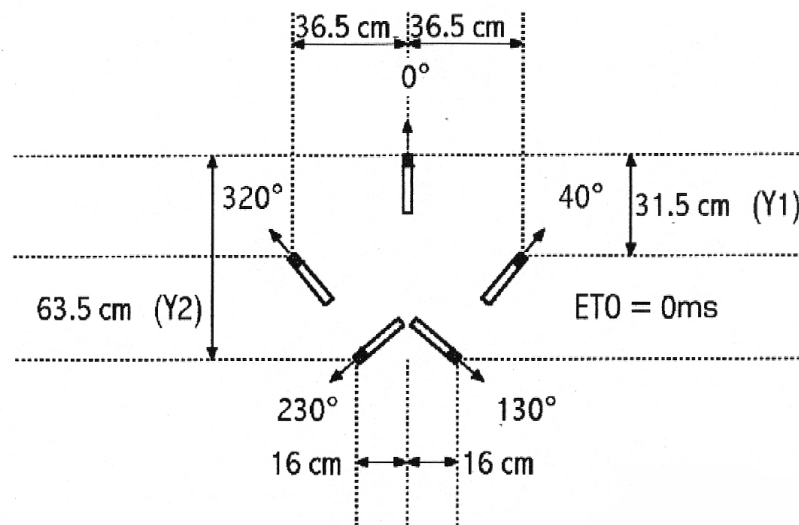
Eine ganze Reihe von Anordnungsmöglichkeiten bietet M. Williams mit seinem MMA (Multichannel Microphone Array) an. Die Grundtheorie geht dabei von einer Aufteilung des 360°-Aufnahmebereichs in 5 einzelne Sektoren aus. Dazu gehören u.a. die gewünschte Front-, Seiten- und Rückabdeckung des aufzunehmenden Klangfeldes und damit die Balance aus Intensitäts- und Laufzeitstereofonie. Allen 220 von Williams vorgeschlagenen Varianten ist gemein, dass sie unter Verwendung von Nierenmikrofonen darauf optimiert sind, eine gleichmäßige Reproduktion des Klangfeldes zu gewährleisten („Critical Linking“).

Je nach gewünschter vorderer, seitlicher und rückwärtiger „Schallabdeckung“ sind verschiedene Kombinationen aus den L-C-R-Frontmikrofonen und den LS/RS Rear-Mikrofonen möglich.

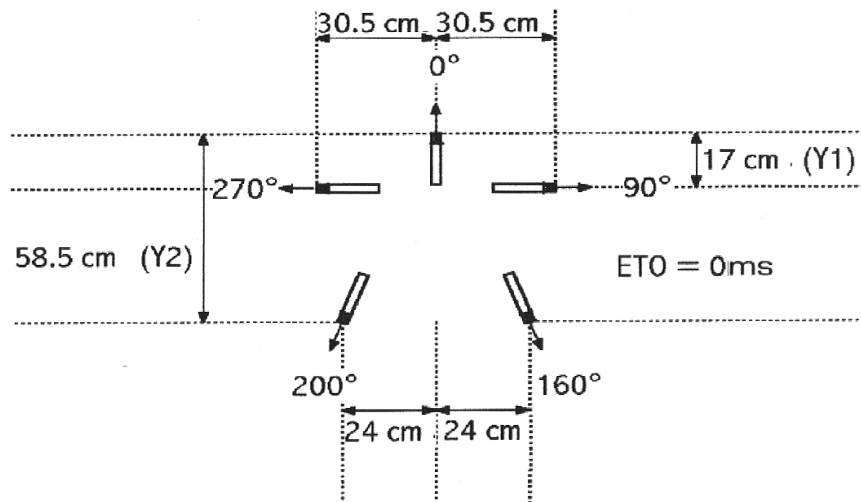
Diese Varianten sind in Tabellentafeln festgehalten und decken nahezu jede mögliche Aufnahmesituation ab [15].

Zwei Beispielanordnungen sind in den Bildern 3-23 und 3-24 dargestellt.

Für die Versuchsaufnahmen in dieser Arbeit wurde jedoch keine dieser Kombinationen ausprobiert.



**Bild 3-23 Williams MMA-Array Beispiel 1, aus [15]**



**Bild 3-24 Williams MMA-Array Beispiel 2, aus [15]**

#### Neumann WNS-100

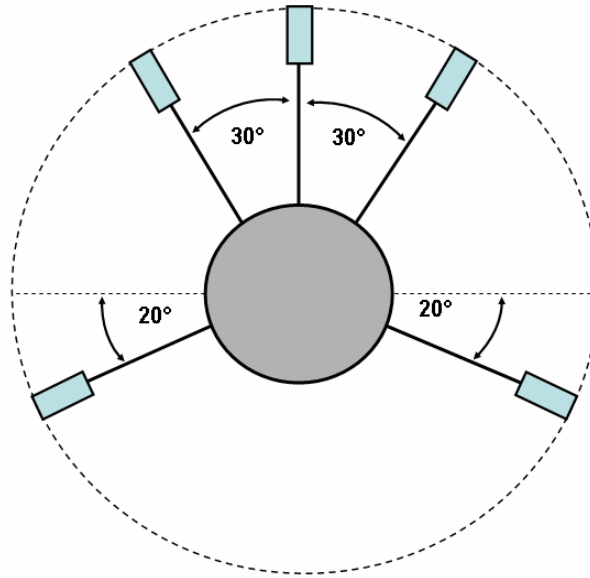
Dieses System ist ein Prototyp der Firma Neumann, das nach eventuellen Nachbesserungen in „Serienproduktion“ gehen soll.

Die recht kleine Anordnung entspricht ihrer Ausrichtung nach den Lautsprecherpositionen der ITU-R BS. 775-1-Norm, wobei die Mikrofonpositionen denen der Lautsprecher entsprechen, jedoch 180° gedreht ausgerichtet (siehe Bild 3-25).

Die drei vorderen Mikrofone sind mit Schwanenhälsen befestigt und flexibel ausrichtbar, wogegen die beiden Rear-Mikrofone starr befestigt sind.

Der geringe Gesamtdurchmesser (ca. 45 cm) und das geringe Gewicht der Anordnung machen auch einen mobilen Einsatz für Atmoaufnahmen ohne großen Aufwand möglich (siehe Bild 3-26).

Inwieweit dieses System auch für Musikaufnahmen geeignet, ist werden die Vergleichstests zeigen.



**Bild 3-25 Neumann WNS-100-Anordnung, schematisch**

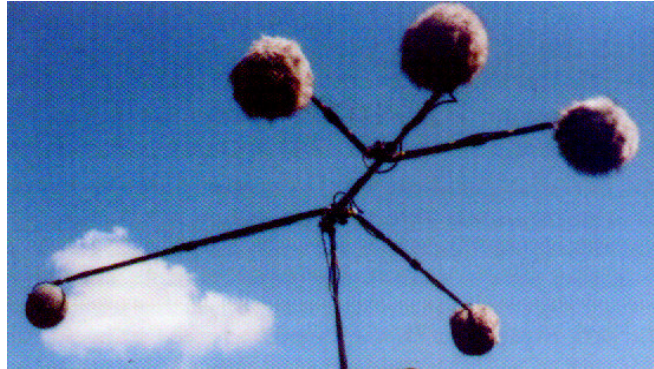


**Bild 3-26 Neumann WNS-100 montiert auf einer Sony DSR-300 DV-Cam**

### A-Ray-System

Die A-Ray-Mikrofonhalterung der Firma Ambient Recordings erlaubt eine flexible Gestaltung der Mikrofonanordnungen.

Durch die jeweils um ca. 180° drehbaren und ausfahrbaren Einzelarme können die Abstände zwischen den Mikrofonen zwischen 0,5 m und 2 m variieren (siehe Bild 3-27). Somit lassen sich z.B. klassische INA 5- Anordnungen einstellen, aber auch viele der nach Williams vorgeschlagenen MMA-Kombinationen realisieren.



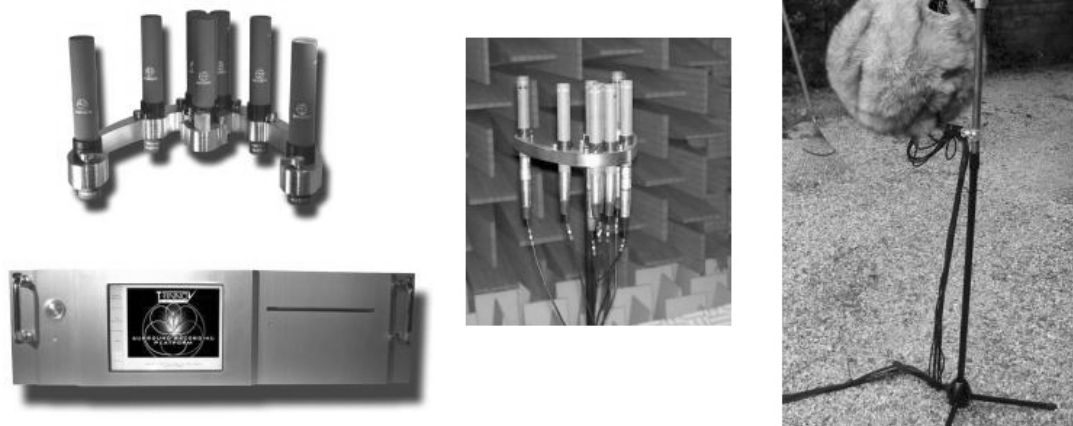
**Bild 3-27 A-Ray Mikrofonhalterung, Prospektfoto**

#### SRP-System von Trinnov-Audio

Einen sehr eigenen Weg geht das von Trinnov-Audio entwickelte SRP-System. Es besteht aus einem Array von 8 Mikrofonen, die auf einer relativ kleinen Metallkonstruktion angebracht sind. Die 8 Signale werden nach digitaler Wandelung mittels Signalprozessoreinheit verändert und als 5.1-Format ausgegeben. Die Mikrofone, die Vorverstärker und die A/D-Converter sind dabei nicht festgelegt und können frei gewählt werden (siehe Bild 3-28).

Die Grundtheorie dieses Systems basiert auf dem „räumlichen Sampeln“ von Schallfeldern. Dabei wird nicht aus nur einem Mikrofonsignal ein entsprechender Wiedergabekanal generiert, sondern für jeden einzelnen Kanal werden die Informationen aller Mikrofone benutzt. Somit erhält jeder Kanal mehr Rauminformationen und eine höhere räumliche Auflösung lässt sich herstellen (HSR - High Spatial Resolution).

Konkret macht sich nach Herstellerangaben neben den hervorragenden unverfälschten Klangeigenschaften eine optimale 360° Darstellung der Phantomschallquellen bemerkbar und ein größerer Sweet-Spot-Bereich. Leider war dieses Gerät zur Zeit der Versuchdurchführung noch nicht verfügbar und konnte daher in dieser Arbeit nicht getestet werden.



**Bild 3-28 SRP-System der Firma Trinnov Audio, Fotos aus Werbeprospekten**

### H2 Pro System von Holophone

Einen ebenfalls sehr eigenen Weg geht das H2-Pro System der Firma Holophone. Es besteht aus einer Art Kunstkopfform, in die 8 Mikrofone implementiert sind (siehe Bild 3-29).

Dabei sind neben den 5 standardisierten Mikrofonen (L/C/R/LS/RS) auch eine separates LFE-Mikrofon (Frequenzgang 20–110 Hz) für den Bassbereich, ein Rear-Center Mikrofon für die erweiterten 6.1-Formate wie Dolby-EX und DTS-ES und ein Top-Mikrofon (nach oben ausgerichtet) für das IMAX-Format vorhanden.

Die einzelnen Signale werden per Multicorekabel diskret abgeführt und die Miniaturmikrofone werden von der Firma DPA hergestellt.

Leider war dieses Produkt zur Zeit der Versuchsdurchführungen für diese Arbeit nicht verfügbar.



**Bild 3-29 H2-Pro Mikrofonsystem von Holophone, Prospektfoto**

### **3.2.4 Auf MS-Technik basierende Systeme**

Die folgenden Hauptmikrofon-Systeme basieren auf der im Kapitel 3.1.1 beschriebenen MS-Technik, deren Mikrofone koinzident angeordnet sind. Die durch die Technik bedingten „gemischten“ Signale werden während oder nach der Aufnahme mittels herstellereinspezifischer Prozessoren und Switchboxen in die 5.1-kompatiblen Signale aufgetrennt.

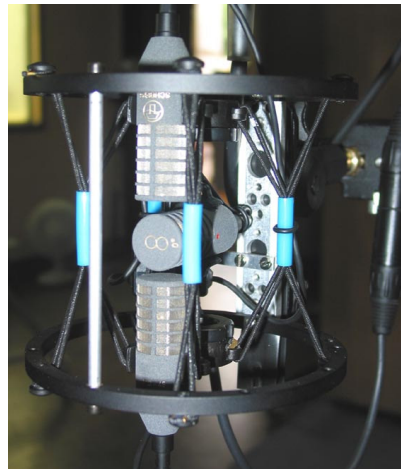
#### Schoeps Doppel-MS

Das Schoeps-Doppel-MS-System ist ein um ein zusätzliches, nach hinten gerichtetes, Nierenmikrofon erweitertes Stereo-MS-System (siehe Bild 3-30). Das nach vorne gerichtete Mittenmikrofon M wird dabei dem Center-

wiedergabekanal zugeordnet und die aus Addition und Subtraktion mit dem Achter-Mikrofon gewonnenen Signale M+S und M-S den Kanälen L und R. Die Rearsignale LS und RS werden aus der Addition und Subtraktion des Achter-Mikrofons mit dem Rear-Mikrofon M-R und M+R gewonnen. Diese Einzelsignalgewinnung erfolgt mit Hilfe eines systemzugehörigen externen Passiv-DMX-Splitters, der per 7-Pol-XLR-Kabel mit dem Mikrofonsystem verbunden ist.

Dieses sehr kompakte und unkomplizierte System ist auch für mobile Einsätze gut geeignet.

Leider konnten in den Versuchen dieser Arbeit durch technische Probleme mit den Vorverstärkern nur sehr begrenzt Aufnahmen mit diesem Mikrofon gemacht werden.



**Bild 3-30 Doppel-MS-System von Schoeps**

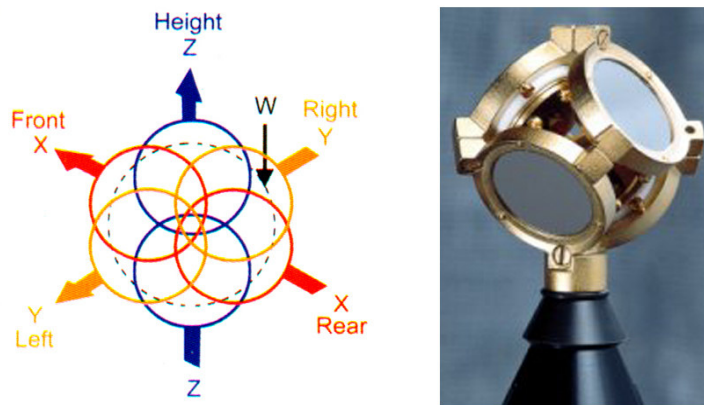
### Soundfield MKV

Das Soundfield MKV-System verfolgt eine ganz andere Philosophie als die bisher beschriebenen Systeme. Sie basiert auf dem Prinzip, das jedes Schallereignis durch 4 „Raumelemente“ beschrieben werden kann (siehe Bild 3-31):



- Dem X-Element für Vorne-Hinten Informationen (Depth)
- Dem Y- Element für Links-Rechts Informationen (Width)
- Dem Z-Element für Oben-Unten Informationen (Height)
- Dem Positionspunkt  $W$ , auf den sich die drei Elemente beziehen

Diese Informationen, die mit Hilfe einer tetraederförmigen Mikrofonanordnung, welche aus 4 Kapseln besteht, und dem MKV-Prozessormodul gewonnen werden, stellen das sogenannte B-Format dar, mit den 4 Signalen X, Y, Z und W.



**Bild 3-31 schematische B-Format-Darstellung und Kapselanordnung des Soundfield MKV Systems, Prospektfotos**

An dem B-Format-Prozessor selbst können diverse elektronische Einstellungen wie Kapselcharakteristika- und virtuelle Positionsveränderungen vorgenommen werden, die im Kapitel 4.5 näher erläutert werden.

Eine zweite Stufe des Soundfield-Systems ist die Gewinnung der 5.1-Format kompatiblen Signale durch den Signalprozessor SP 451. Dieser Prozessor, der auch als Softwarevariante für DAW-Systeme angeboten wird, bildet im Kern eine „konventionelle“ Hauptmikrofonanordnung nach, mit den Mikrofonen L (Pos.  $-45^\circ$ ), C (Pos.  $0^\circ$ ), R (Pos.  $+45^\circ$ ), LS (Pos.  $-135^\circ$ ) und RS (Pos.  $+135^\circ$ ).

Der LFE-Kanal wird dabei von seinem eigenen virtuellen Mikrofon generiert. Die Positionen der Front- und Rearmikrofone lassen sich an dem Gerät in Ihrer „Ausrichtung“ elektronisch verändern.

Die Versuchsaufnahmen werden zeigen, wie gut sich diese Technik in der Praxis bewährt.



**Bild 3-32 Soundfield MKV System, Prospektfoto**

### Schoeps KFM-360

Die erweiterte Version des von der Firma Schoeps hergestellten KFM-360 Kugelflächenmikrofons nach Jerry Bruck ist für Surround-Aufnahmen abgestimmt. Es basiert auf einer Kombination von MS-Technik mit Trennkörper-Stereophonie.

An den Seiten einer schallharten Kugel mit 18 cm Durchmesser sind 2 Kugelmikrofone entgegengesetzt eingebaut, die einen Aufnahmewinkel  $\varphi$  von  $120^\circ$  bilden. Diese werden durch 2 zusätzliche Achter-Mikrofone, die dicht unterhalb der Kugeln angebracht und mit dem Winkel  $0^\circ$  nach vorne ausgerichtet sind, ergänzt.

Die Mikrofonsignale werden der Prozessoreinheit DSP-4 KFM-360 zugeführt und dort in 5.1-kompatible Signale gewandelt. Dabei ergeben sich je Achter/Kugel-Kombination nach entsprechender Addition und Subtraktion die 4 Signale L und LS bzw. R und RS. Das Centersignal C wird durch eine Matrixschaltung nach M. Gerzon [3] aus den beiden Frontsignalen L und R gewonnen. Ein LFE-Kanal für Signale unterhalb 70 Hz wird ebenfalls generiert.

Die Charakteristika der Achter-Mikrofone lassen sich zudem an der Prozessoreinheit elektronisch verändern.  
Für die Aufnahmen in dieser Arbeit war dieses System leider nicht verfügbar.



**Bild 3-33 KFM-360 Mikrofon mit zusätzlichen Achter-Mikrofonen,  
Prospektfoto**

#### Sanken CUW-180

Das von der Firma Sanken entwickelte koinzidente Stereo-Mikrofon CUW-180 besteht aus 2 Nierenmikrofonen in X/Y-Anordnung mit einem einstellbaren Aufnahmewinkel  $\varphi$  von bis zu  $180^\circ$ . Durch Hinzunahme eines zweiten, entgegengesetzt positionierten, Systems wird also eine  $360^\circ$  Abbildung möglich (siehe Bild 3-34).

Wie das Center und das LFE-Signal gewonnen wird, wurde jedoch von Herstellerseite nicht erwähnt und bleibt hiermit offen.

Dieses System stand für die Versuchsaufnahmen leider nicht zu Verfügung.