

Diplomarbeit im Studiengang Audiovisuelle Medien  
Fachhochschule Stuttgart  
Hochschule der Medien

**Richtungs- und Entfernungabbildung in der Fünfkanal-  
Stereofonie bei der Mischung von Popmusik**

Johannes Pfitzenmaier  
Matrikel-Nr.: 11631

Mai, 2005

1. Prüfer: Prof. Oliver Curdt
2. Prüfer: Prof. Axel Hartz

### **Abstract**

Die Fünfkanal-Stereofonie bietet gegenüber herkömmlichem Zweikanal-Stereo eine von der Lautsprecherwiedergabe bisher noch nicht gekannte Qualität von Höreindrücken und – gerade auch für die Popmusik – neue, kreative Gestaltungsmöglichkeiten. Diese sind aber bisher noch wenig erforscht worden. In dieser Arbeit wird untersucht, wie sich Schallquellen in der Fünfkanal-Stereofonie bezüglich Richtung und Entfernung verhalten und abbilden lassen.

*„Stereo-akustisches Hören ist eine Art des Hörens, die dem Betrachter deutlich mehr Information liefert, als es das Doppelte des monauralen Hörens vermag.“ (Sedee, 1957)*

## Erklärung

---

Hiermit versichere ich, dass ich die Arbeit selbständig und ohne fremde Hilfe verfasst habe. Sämtliche Quellen sind im Text und Anhang nachgewiesen.

---

Johannes Pfitzenmaier

## Danksagungen

---

Bei folgenden Personen, die mich unterstützt und zum Entstehen dieser Arbeit beigetragen haben, möchte ich mich herzlich bedanken:

Bei Prof. Oliver Curdt und Prof. Axel Hartz für ihre herzliche und fachkundige Betreuung, Jörg Bauer, der Studioräumlichkeiten und Equipment zur Verfügung stellte, auch Oliver Neumann und dem MW-Tonstudio für zusätzliche Mikrofonleihgaben, und Klaus Hanselmann, der mir ebenfalls Mikrofone aus seinem privaten Fundus zur Verfügung stellte. Ein großes Dankeschön auch an Peter Ruhrmann und den gesamten Bereich Interaktive Medien für die Bereitstellung von Arbeitsplätzen und Geräten, sowie Zeit und Geduld.

Ebenso möchte ich mich bei Andi Zbik, Axel Hemprich, Michael Deak, Fabian Wendt und Carsten Netz von der Band Schnute bedanken, und natürlich auch bei den anderen Musikern: Ralf Specht (Gitarre), Bianca Poppke (Gesang) und meine Schwester Sarah (Piano).

Dank auch an Tobias Wagner und Matthias Pasedag für zahlreiche Tipps und Anregungen.

Bedanken möchte ich mich auch bei Bianca Schmied und Natascha Matthes für das Korrekturlesen.

Und nicht zu vergessen, alle Probanden, die an den Hörversuchen teilgenommen haben.

Zum Schluß möchte ich auch allen Freunden, Verwandten und Bandkollegen danken, die mich während dieser Zeit motiviert und begleitet haben.

<b>1. Einleitung</b> .....	1
<b>2. Begriffserklärungen</b> .....	3
<b>3. Der 3/2-Stereo Standard</b> .....	5
<b>4. Räumliches Hören</b> .....	8
4.1 Lokalisation natürlicher Schallquellen .....	8
4.1.1 Richtungshören .....	9
4.1.2 Entfernungshören .....	13
4.2 Lokalisation bei elektroakustischer Wiedergabe .....	16
4.2.1 Summenlokalisierung .....	17
4.2.2 Präzedenzeffekt .....	19
4.2.3 Assoziationsmodell von Theile .....	20
4.3 Erzeugung von Richtung und Entfernung in der 2.0-Stereofonie .....	23
4.3.1 Richtungsabbildung .....	23
4.3.2 Entfernungsabbildung .....	26
<b>5. Theoretische Überlegungen zur 3/2-Stereofonie</b> .....	28
5.1 Unterschiede zur 2.0-Stereofonie .....	28
5.2 Probleme bei der Positionierung von Schallquellen .....	30
<b>6. Misch-Experimente</b> .....	36
6.1 Centerlautsprecher .....	36
6.2 Seitliche Abbildung .....	40
6.3 Das Surroundfeld .....	48
6.4 Umhüllung .....	53

## Inhalt

---

<b>7. Surround-Produktion</b> .....	57
7.1 Aufnahme .....	57
7.2 Mischung und Mastering .....	61
<b>8. Hörversuche</b> .....	64
8.1 Vorüberlegungen und Konzeption .....	64
8.2 Ablauf der Hörversuche .....	66
8.3 Auswertung und Ergebnisse .....	68
8.4 Fazit .....	84
<b>9. Ausblick</b> .....	86
<b>Anhang</b> .....	I
Fragebogen zum Hörversuch .....	I
Quellenverzeichnis .....	VIII
Index .....	X
Trackliste der DTS-CD .....	XI

### 1. Einleitung

Surround-Formate sind beim Kinofilm schon seit Jahrzehnten etabliert. Vor einigen Jahren wurde im Heimbereich der 5.1-Standard eingeführt. Die Zahl der Haushalte mit Surroundanlage wächst ständig. Der Siegeszug der DVD-Video gegenüber der bisherigen VHS-Kassette ist in vollem Gange. Praktisch keine neue DVD-Video erscheint ohne 5.1-Ton.

Parallel dazu entstand auch in der Musikproduktion der Wunsch, die neu hinzugekommenen Kanäle zu nutzen. Derzeit sind zwei Formate auf dem Markt, die DVD-Audio von Panasonic/Mitsubishi und die Super Audio CD (SACD) von Sony/Philips. Beide Kandidaten verfügen über mindestens sechs diskrete Audiokanäle und sind somit für die Wiedergabe von fünfkanal-stereofonen Produktionen geeignet. Zudem gibt es bereits Musikproduktionen, die fünfkanalig produziert wurden und auf DVD-Video oder als DTS-CD<sup>1</sup> erschienen sind.

Dennoch führen Surround-Audioproduktionen auch in der Popmusik<sup>2</sup> nach wie vor ein Schattendasein und letztendlich wird der Erfolg der neuen Audioformate davon abhängen, ob die Mehrheit der Popmusik-Hörer, die den überwiegenden Anteil der Musikkonsumenten darstellt, das Surround-Hörerlebnis als deutlichen Zugewinn erkennt. In den Audio-Fachzeitschriften wird häufig die überragende Klangqualität von DVD-A/SACD gegenüber der CD herausgestellt. Diese ist aber für viele Konsumenten im Gegensatz zum Surround-Hörerlebnis wenig nachvollziehbar. Es ist folglich nicht die bessere Auflösung, die den entscheidenden Kaufanreiz bringt, sondern eindeutig das Argument „Surround-Sound“.

---

<sup>1</sup> DTS steht für *Digital Theatre System*, ein ursprünglich für den Kineton entwickeltes Komprimierungsverfahren, ähnlich dem AC-3-Verfahren. DTS-codierter Mehrkanalton kann mit den meisten handelsüblichen DVD/CD-Playern mit dem entsprechenden Decoder wiedergegeben werden.

<sup>2</sup> Der Begriff *Popmusik* wird im folgenden stellvertretend für alle Formen moderner, populärer Musik (Rock, Pop, Elektro, Jazz, Musical, Folklore, Schlager, usw.) verwendet. Entsprechend dazu wird der Begriff *klassische Musik* gebraucht, wenn von sinfonischer Musik, Kammermusik, Geistliche Musik, Oper usw. die Rede ist. Diese Bezeichnungen sind m. E. wertneutraler und zeitgemäßer als die Begriffe Unterhaltungsmusik und Ernste Musik (U- und E-Musik)

Für den Musikkonsumenten kann die Umhüllung durch Schallereignisse ein völlig neues und deutlich gesteigertes Hörerlebnis bedeuten. Bei der Aufnahme und Mischung klassischer Musik wird in der Regel versucht, das Schallereignis möglichst authentisch abzubilden. Die Surroundkanäle werden hauptsächlich dazu benutzt, den aufgenommenen Raumeindruck wiederzugeben. Im Gegensatz dazu bietet die Fünfkanal-Stereofonie in der Popmusik, in der es ja nicht primär um eine realistische Abbildung eines natürlichen Klangbildes geht, ganz neue Möglichkeiten eine Mischung zu gestalten; vergleichbar mit dem Übergang von Mono zur Zweikanal-Stereofonie.

In der Popmusik wird seit den 60er Jahren in Mehrspurtechnik produziert. Dabei werden die Instrumente einzeln mikrofoniert (close miking) und meist auch nacheinander im Overdubverfahren oder zumindest räumlich getrennt aufgenommen. Überlegungen wie in der klassischen Musik, welches der zahlreichen neuen und viel diskutierten Surround-Hauptmikrofonieverfahren angewandt werden soll, spielen - von Ausnahmen (z.B. Schlagzeugaufnahme unter Verwendung von Raummikrofonen) abgesehen - folglich keine Rolle. Oftmals kommen viele Klänge sogar direkt aus Samplern und elektronischen Klangerezeugern. Im Gegensatz zu Klassik-Produktionen ändert sich bei der Popmusik aufnahmeseitig somit kaum etwas gegenüber herkömmlichen zweikanaligen Produktionen.

Während in der Zweikanal-Stereofonie Schallereignisse nur auf der Stereobasis zwischen den beiden Lautsprechern plaziert werden können, besteht in der Fünfkanal-Stereofonie nun die Möglichkeit das Klanggeschehen rund um den Hörer aufzubauen. Im folgenden wird untersucht, wie die zusätzlichen Kanäle beim Mischen genutzt und mit welchen Methoden Schallereignisse abgebildet werden können. Anhand von Erkenntnissen über Abbildungsmöglichkeiten in der Zweikanal-Stereofonie, werden Theorien über Mischverfahren in der Fünfkanal-Stereofonie aufgestellt. Anschließend werden diese Theorien in praktischen Versuchen überprüft und in Form einer Popmusik-Produktion umgesetzt.

### 2. Begriffserklärungen

*Stereofonie* setzt sich aus den griechischen Wörtern *stereo* und *phonos* zusammen. *Stereo* bedeutet räumlich, körperlich und *phonos* steht für Ton, Klang. Stereofonie hat nichts mit zwei Kanälen zu tun, wie fälschlicherweise oft angenommen wird, sondern ist ein Gattungsbegriff für mehrkanalige Tonübertragungsverfahren. Darunter fällt die altbewährte Zweikanal-Stereofonie wie auch die Fünfkanal-Stereofonie.

Der Begriff *Surround-Sound* steht hauptsächlich für *Umhüllung*. Allgemein wird damit eine Erweiterung der Zweikanal-Stereofonie ausgedrückt. Darunter fallen sowohl codierte Verfahren, bei denen die Signale für Center- und Surroundlautsprecher über eine Matrix den Kanälen Links/Rechts hinzugefügt und bei Wiedergabe daraus decodiert werden, als auch diskrete Verfahren, bei denen die Kanäle auf separaten Spuren vorliegen.

Laien sprechen oftmals auch von „3-D-Klang“, was irreführend ist, denn weder in der Zweikanal- noch in der Fünfkanal-Stereofonie werden drei Dimensionen abgebildet. In beiden Verfahren gibt es maximal zwei Dimensionen, Breite und Tiefe. Während jedoch in der Zweikanal-Stereofonie die zweite Dimension, die räumliche Tiefe, nur simuliert werden kann, existiert in der Fünfkanal-Stereofonie zwischen den Lautsprechern wirklich ein reales, zweidimensionales Feld. Desweiteren ist in der Zweikanal-Stereofonie das Klanggeschehen ausschließlich frontal vor dem Hörer, während in der Fünfkanal-Stereofonie das Klanggeschehen rund um den Hörer aufgebaut werden kann.

Unter *Umhüllung* oder *Einhüllung* versteht man die Wahrnehmung eines umgebenden, nicht lokalisierbaren Hörereignisses. Dies kann sowohl durch Nachhall, als auch durch räumlich verteilte Schallquellen (z.B. Applaus, Atmo) hervorgerufen werden.

Der Begriff *Raumeindruck* beschreibt in der Hörakustik den Eindruck eines im Raum befindlichen Hörers, die Empfindung der Größe des Raumes, der Halligkeit, also das Vorhandensein von Diffusschall, und der Räumlichkeit.<sup>3</sup>

*Räumlichkeit*, bezeichnet, beschreibt die räumliche Aufweitung eines Hörereignisses, bei der die Schallquelle eine größere Ausdehnung als ihre visuellen Konturen bekommt.<sup>4</sup> In der englischsprachigen Literatur wird dafür der Begriff *Apparent Source Width* (ASW) verwendet.

Die Bezeichnung *sweet spot* steht für einen räumlich begrenzten Bereich, innerhalb dem die Eigenschaften einer Aufnahme am besten wiedergegeben werden, nämlich annähernd so, wie sie der Tonmeister bei der Mischung erzeugt hat. Die optimale Hörposition liegt dabei im Schnittpunkt der Lautsprecherachsen. Die Größe des *sweet spots* ist von der Art der Entstehung der Aufnahme und Mischung abhängig. Eine gelungene Vergrößerung dieser Zone wird häufig als *sweet area* bezeichnet.

---

<sup>3</sup> Vgl.: Dickreiter, Michael, *Handbuch der Tonstudioteknik*, Band 1, K. G. Sauer Verlag, München 1997, S. 27 f.

<sup>4</sup> Vgl.: Blauert, Jens, *Räumliches Hören. Nachschrift: neue Ergebnisse und Trends seit 1972*, S. Hirzel Verlag, Stuttgart 1974, S. 75

### 3. Der 3/2-Stereo Standard

Die Basis-Wiedergabe-Konfiguration des 3/2-Stereo Standards besteht aus drei Frontkanälen (Links, Center, Rechts) und zwei hinteren Surround-Kanälen (Links-Surround, Rechts-Surround). Alle fünf Kanäle sind hierbei gleichwertig und übertragen den kompletten hörbaren Frequenzbereich. Existiert noch ein zusätzlicher LFE-Kanal (Low Frequency Extension)<sup>5</sup> mit begrenzter Bandbreite, so bezeichnet man dies als 5.1-Format. Diese Konfiguration hat ihren Ursprung im Filmtone, bei dem ein mittlerer „Dialog“-Kanal von primärer Bedeutung ist. Sie stellt einen Kompromiss dar, zwischen der Forderung nach optimaler räumlicher Wiedergabe und der Notwendigkeit der Kompatibilität zu DVD-Filmtone und Zweikanal-Wiedergabe, sowie einer für den Heimanwender praktikablen Lösung.

Für die Aufstellung der Lautsprecher gelten gemäß ITU-R BS. 775-1<sup>6</sup> folgende Empfehlungen (siehe auch Abb. 1):

Die fünf Hauptlautsprecher werden in einem Kreis mit einem Radius von 2 - 3 m freistehend angeordnet. Die Abhörposition befindet sich in der Mitte dieses Kreises. Linker und rechter Lautsprecher stehen je in einem 30° Winkel zum Centerlautsprecher, die hinteren in einem Winkel von 100°-120°. Weiter sollten die fünf Lautsprecher identisch sein und das gesamte hörbare Spektrum wiedergeben. Ist eine kreisförmige Anordnung der Lautsprecher nicht möglich, sieht die Empfehlung vor, dass die Lautsprecher innerhalb des Kreises entsprechend verzögert werden sollten. Um frühe Reflexionen unter 15 ms zu vermeiden, sollten die Boxen in einem Abstand von mindestens 1 - 1,5 m von den Wänden aufgestellt, und die Frontlautsprecher nicht nach unten gewinkelt montiert werden, da sonst Kammfiltereffekte durch Reflexionen mit der Mischpultoberfläche auftreten können. Für den Abhörraum gelten die gleichen Bedingungen wie auch schon für Zweikanal-Studios. Der Raum sollte spiegelsymmetrisch zum Hörer sein, die Nachhallzeit (0,2 - 0,4 s) entsprechend der Raumgröße (>25 m<sup>2</sup>) kurz.

---

<sup>5</sup> In einigen ITU-Dokumenten auch als *Low Frequency Enhancement*, bei SMPTE als *Low Frequency Effect* bezeichnet.

<sup>6</sup> ITU-R BS. 775-1, *Multichannel stereophonic sound system with and without accompanying picture*, Rec., International Telecommunications Union, Geneva, Switzerland (1992/1994)

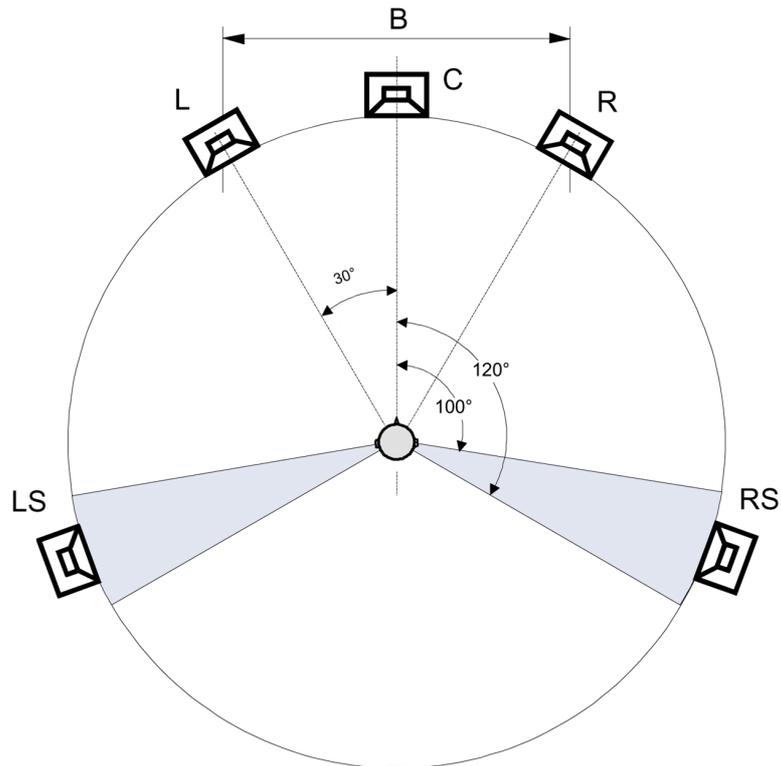


Abb. 1; 3/2-Stereo Anordnung nach ITU-R BS. 775-1

Der im 5.1-Format noch zusätzliche LFE-Kanal verursacht mitunter Verwirrung bei Tonmischungen, die nicht für Kinoanwendungen vorgesehen sind. Er wird hauptsächlich bei der DVD-Video für spezielle tieffrequente Effekte eingesetzt. Für Audioproduktionen wird er eigentlich nicht benötigt, da die fünf Hauptlautsprecher schon den gesamten Frequenzumfang abdecken. Das sollte aber nicht damit verwechselt werden, dass ein beim Konsumenten stehender Subwoofer je nach Verstärker und Frequenzgang der fünf Hauptlautsprecher auch den Bassanteil der Hauptlautsprecher übernimmt. Diese Signalaufteilung wird jedoch vom Verstärker gesteuert, und entsprechende Signale dürfen nicht schon bei der Mischung auf den LFE-Kanal gelegt werden. Jeglicher Tieftoninhalt, der für die Mischung entscheidend ist, sollte daher nicht dem LFE-Kanal, sondern den Hauptkanälen zugeordnet werden. Der LFE-Kanal ist tatsächlich nur für „Effekte“ vorgesehen. Es sollte daher keine Rolle spielen, ob der Konsument in der Lage ist, diese wiederzugeben oder nicht.

Über die optimale Aufstellung eines Subwoofers gibt es unterschiedliche Auffassungen. Die Annahme, dass es egal ist, wo der Subwoofer aufgestellt wird, da tiefe Frequenzen nicht lokalisierbar sind, ist ein Trugschluß und ein ungünstig platzierter Subwoofer kann sehr störend wirken.

Es gibt Messungen, die nahe legen, dass die Anordnung eines einzelnen Subwoofers in einer Raumecke den weitesten und ausgeglichtesten Frequenzgang ermöglicht.<sup>7</sup> Lautsprecher, die in den Ecken platziert werden, weisen eine deutliche Bassverstärkung auf und sind gut an die meisten Raummoden anzukoppeln.

Eigengeräusche, Verzerrungen, und Informationen über 120 Hz, die vom Subwooferort abgestrahlt werden, können ihn allerdings lokalisierbar machen. In solchen Fällen kann es vorteilhafter sein, den Subwoofer zwischen den Frontlautsprechern aufzustellen.<sup>8</sup> Ein direkt unter dem Center platzierter Subwoofer ist jedoch dadurch benachteiligt, dass er am Nulldurchgang seitlicher, stehender Wellen aufgestellt ist. Eine Verschiebung aus der Mitte heraus ist daher meist sinnvoller.

Es gibt Gründe für die Annahme, dass mehrere Tieftoneinheiten, die dekorrelierte Signale erzeugen, eine natürlichere räumliche Wiedergabe erzeugen als eine monaurale Tieftonwiedergabe über ein einzelnes System.<sup>9</sup> Gemäß diesem Vorschlag ist es besser, für die Wiedergabe von monofonem Frequenzinhalt zwei Einheiten mit 90° Phasenverschiebung an den Seiten des Hörers anzuordnen, um die asymmetrischen seitlichen Moden erfolgreich anzuregen und die tief-frequente Räumlichkeit zu verbessern.

---

<sup>7</sup> Nousaine, T., *Multiple subwoofers for home theatre*, J. Audio Eng. Soc. (Abstracts), Vol. 45, S. 1015 (Nov. 1997), Preprint 4558.

<sup>8</sup> Zacharov, N./Bech, S./Meares, D., *The use of subwoofers in the context of surround sound program reproduction*, J. Audio Eng. Soc. (Abstracts), Vol. 46, S.276–287 (Apr. 1998).

<sup>9</sup> Griesinger, David, *Spatial impression and envelopment in small rooms*, <http://world.std.com/~griesngr>, Seitenaufruf vom 26.03.2005

### 4. Räumliches Hören

Zu einem großen Teil geht es bei der Mischung von Popmusik darum, Richtung und Entfernung von meist sehr nah aufgenommenen Instrumenten oder sogar rein elektronisch erzeugten Klängen, nachzubilden oder zu simulieren. Dies ist bei der Fünfkanaal-Stereofonie nicht anders als bei der Zweikanaal-Stereofonie. Um ein überzeugendes Klangbild kreieren zu können, muß daher verstanden werden, wie das menschliche Gehör Richtung und Entfernung wahrnimmt.

#### 4.1 Lokalisation natürlicher Schallquellen

Die akustische Wahrnehmung einer natürlichen Schallquelle kommt durch eine Vielzahl von Übertragungsgliedern zustande. Der Schall gelangt über die Ohrmuschel und den äußeren Gehörgang über das Mittelohr zum Innenohr und wird dort in Nervenimpulse umgewandelt, die anschließend an das Gehirn weitergeleitet und mit gespeicherten Reizmustern verglichen werden. Dabei spielen die Art des Schallereignisses und die visuelle Wahrnehmung eine Rolle, ebenso ob dem Hörer das Signal bekannt ist oder nicht.

Man unterscheidet zwischen dem physikalischen Phänomen *Schallereignis* und der mentalen Interpretation dessen, dem *Hörereignis*. Die Zuordnung zwischen dem Hörereignisort und bestimmten Merkmalen eines Schallereignisses, z.B. dem Schallereignisort, bezeichnet man als *Lokalisation*.<sup>10</sup>

Hörereignisort und Schallquellenort müssen nicht immer übereinstimmen. Unter bestimmten Umständen kann es zu Richtungsirritationen, Lokalisationsfehlern und zur Wahrnehmung mehrerer Hörereignisse kommen. So kommt es z.B. bei schmalbandigen, sinusähnlichen oder „unnatürlich“ verzerrten Signalen häufig vor, dass das Hörereignis in einer bezüglich der Ohrachse symmetrisch gespiegelten Richtung wahrgenommen wird.<sup>11</sup>

---

<sup>10</sup> Vgl.: Blauert, Jens, *Räumliches Hören*, S. Hirzel Verlag, Stuttgart 1974, S. 30

<sup>11</sup> Vgl.: Ebd., S. 33 ff

### 4.1.1 Richtungshören

Die Richtungswahrnehmung in der Horizontalebene erfolgt durch Auswertung von Unterschieden der beiden Ohrsignale, die entstehen wenn eine Schallquelle aus der Medianebene ( $0^\circ$ ) auswandert. Hierbei treten interaurale Zeit- und Pegeldifferenzen, sowie spektrale Unterschiede, die durch die charakteristische Form von Ohrmuschel und äußerem Gehörgang zustande kommen, auf. Diese Signalunterschiede führen auch einzeln zu einer Lokalisation, beim natürlichen binauralen Hören treten sie jedoch immer gemeinsam auf. Eine getrennte Untersuchung der einzelnen Vorgänge ist folglich im freien Schallfeld nicht möglich. Man arbeitet stattdessen fast immer mit dichotischer Kopfhörerbeschallung.<sup>12</sup> Die so ermittelten Ergebnisse können zwar nicht ohne Einschränkung auf die natürliche Hörsituation übertragen werden, lassen aber Rückschlüsse auf die Richtungswahrnehmung im freien Schallfeld zu. Da bei Wiedergabe über Kopfhörer das Hörereignis zumeist im oder sehr nahe am Kopf entsteht, spricht man in diesem Fall nicht von Lokalisation, sondern bezeichnet die stattfindende Auslenkung als *Lateralisation*. Zur Beschreibung der Hörereignisorte verwendet man ein kopfbezogenes Koordinatensystem.

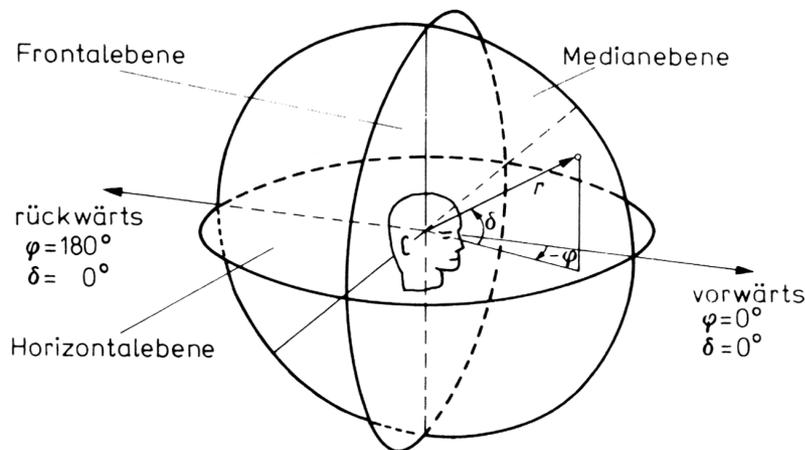


Abb. 2; kopfbezogenes Koordinatensystem<sup>13</sup>

<sup>12</sup> Bei dichotischer Kopfhörerbeschallung erhalten die beiden Kopfhörer unterschiedliche Signale, bei diotischer Beschallung erhalten beide Kopfhörer das gleiche Signal, bei monotischer Darbietung bekommt nur ein Kopfhörer ein Signal.

<sup>13</sup> Aus: Blauert, Jens, *Räumliches Hören*, S. Hirzel Verlag, Stuttgart 1974, S. 11

### *Interaurale Zeitdifferenzen*

Befindet sich eine Schallquelle außerhalb der Medianebene, treten Laufzeitunterschiede zwischen linken und rechtem Ohr auf. Dabei wird das Signal in Richtung desjenigen Ohres ausgelenkt, bei dem es früher eintrifft. Die interaurale Zeitdifferenz ist bei genau seitlichem Schalleinfall ( $90^\circ$ ) am größten. Messungen haben ergeben, dass bis zu einer Laufzeitdifferenz von  $630 \mu\text{s}$  die Auslenkung linear verläuft, danach flacht die Kurve ab, bis bei etwa  $1 \text{ ms}$  die maximale seitliche Auslenkung erreicht ist. Es hat sich gezeigt, dass die Bewertung der interauralen Zeitdifferenzen aufgrund gleicher Signalflanken geschehen muß. Versuche bei denen die Phasenlage um  $180^\circ$  verschoben war, führten zu einer unscharfen, verbreiterten Lateralisation oder zur Wahrnehmung zweier Hörereignisse.<sup>14</sup>

Eine Laufzeitdifferenz von  $\Delta t = 630 \mu\text{s}$  entspricht einem Schalleinfall von  $90^\circ$  genau in der Richtung der Hörachse, also der Verbindung beider Ohren. Der „wirksame Ohrabstand“<sup>15</sup>, d.h. der maximale Wegunterschied zwischen den Ohren, ergibt sich somit aus dem Produkt der interauralen Zeitdifferenz und der Schallgeschwindigkeit:

$$a = \Delta t \cdot c = 0,63 \text{ ms} \cdot 10^{-3} \cdot 344 \text{ m/s} = 0,217 \text{ m} \approx 21,7 \text{ cm}$$

Die Grenzfrequenz für die maximale Auslenkung ist somit:

$$f_{\text{grenz}} = \frac{c}{2\lambda} = \frac{344 \text{ m/s}}{0,434 \text{ m}} = 792,63 \text{ Hz} \approx 800 \text{ Hz}$$

Bei reinen Tönen verringert sich die maximale Auslenkung oberhalb der Grenzfrequenz  $f_{\text{grenz}} = 800 \text{ Hz}$ , bis das Gehör bei etwa  $1600 \text{ Hz}$  keinerlei Zeit- bzw. Phasendifferenzen mehr erkennen kann. Bei breitbandigen Signalen ist jedoch festzustellen, daß auch über  $1,6 \text{ kHz}$  noch unveränderte Auslenkung besteht. Es werden nun die Signalhüllkurven zur Auswertung herangezogen.

---

<sup>14</sup> Vgl.: Blauert, Jens, *Räumliches Hören*, S. Hirzel Verlag, Stuttgart 1974, S. 117 ff.

<sup>15</sup> Aus: Sengpiel, Eberhard, *Der Ohrabstand - Welcher?*,  
<http://www.sengpielaudio.com/DerOhrabstand-Welcher.pdf>, Seitenaufruf vom 10.03.2005

### *Interaurale Pegeldifferenzen*

Im freien Schallfeld treten bei seitlichem Schalleinfall neben Laufzeitunterschieden immer auch interaurale Pegeldifferenzen auf. Die Pegelunterschiede sind dabei immer frequenzabhängig. Unterhalb von etwa 300 Hz gibt es praktisch keine Pegeldifferenzen, da tiefe Frequenzen aufgrund ihrer großen Wellenlänge um den Kopf gebeugt werden. Mit steigender Frequenz stellt der Kopf zunehmend ein Hindernis dar, bei hohen Frequenzen sind die durch die Abschattung und Druckstau verursachten Pegelunterschiede am größten. Es treten folglich auch Klangfarbenunterschiede zwischen den Ohrsignalen auf, die zwar nicht bewußt wahrgenommen werden, aber bei der Richtungsbestimmung miteinfließen. Voraussetzung hierbei ist allerdings die Kenntnis über die Klangfarbe bei frontalem Schalleinfall.

Bei Versuchen mit Kopfhörerbeschallung mit frequenzunabhängigen Pegeldifferenzen wurden für die maximale seitliche Auslenkung Werte zwischen 10 und 15 dB ermittelt. Der Grenzwert für die Maximalauslenkung ist dabei schwer zu bestimmen, da ab einer Differenz von etwa 8 bis 10 dB, die Breite der Hörereignisses und somit die Lateralisationsunschärfe zunimmt. Hinzukommt, dass die für eine bestimmte seitliche Auslenkung erforderliche Pegeldifferenz wiederum frequenzabhängig ist. Im Bereich um 2 kHz ist sie am geringsten und steigt zu tieferen und in geringerem Maß auch zu höheren Frequenzen hin an.

Beim Zusammenwirken von interauralen Zeit- und Pegeldifferenzen, wie es beim „natürlichen“ Hören im freien Schallfeld der Fall ist, dominiert je nach Frequenz einer der beiden Lokalisationsmechanismen. Unterhalb von 1,6 kHz dienen hauptsächlich interaurale Zeitdifferenzen von Trägerschwingungen der Richtungsbestimmung, während oberhalb von 1600 Hz interaurale Pegeldifferenzen und Hüllkurvenverschiebungen zum Tragen kommen.

### *Spektraldifferenzen*

Außer Laufzeit- und Pegeldifferenzen werden vom Gehirn zusätzlich auch noch spektrale Veränderungen ausgewertet, um die Richtung einer Schallquelle zu bestimmen. Durch die Form der Ohrmuschel und des äußeren Gehörgangs wird ein Signal je nach Einfallswinkel unterschiedlich spektral verändert. Diese Änderungen nehmen wir beim normalen binauralem Hören nicht bewusst wahr, sie werden aber vom Gehirn zusätzlich ausgewertet, um den Einfallswinkel einer Schallquelle zu bestimmen. Deshalb lassen sich auch bei einohrigem Hören (monaural) Richtungseindrücke gewinnen.

Befindet sich eine Schallquelle in der Medianebene, treten praktisch keine interauralen Zeit- und Pegeldifferenzen auf. Die Entscheidung, ob ein Signal von vorne, hinten oder oben kommt, wird ausschließlich durch die spektralen Veränderungen erkannt. Die Lokalisation erfolgt hierbei über die *richtungsbestimmenden Bänder*.<sup>16</sup> Untersuchungen haben dabei ergeben, dass einzelne Frequenzbänder unabhängig von Einfallswinkel und Schalldruckpegel für verschiedene Richtungen dominant sind. Der Frequenzbereich zwischen 250 Hz und 500 Hz, sowie zwischen 2,5 kHz und 4 kHz wurde von den Probanden als Schalleinfall von vorne interpretiert, um 1 kHz und oberhalb von 12 kHz als Schalleinfall von hinten und der Bereich um 8 kHz als von oben.

Die Richtungsbestimmung in der Medianebene funktioniert nur bei breitbandigen Signalen, schmalbandige Signale können nicht lokalisiert werden. Zudem spielen Lerneffekte eine wichtige Rolle. Besitzt das Gehör Erfahrung über die spektrale Zusammensetzung des Signals, ist die Lokalisation wesentlich präziser als bei einem unbekanntem Signal. Während in der Horizontalebene von vorne eintreffende Schallquellen auf etwa 3° und seitlich einfallende auf etwa 4 bis 5° genau lokalisiert werden können, liegt die Lokalisationsunschärfe in der Medianebene bei bekannten Signalen bei 9°, bei unbekanntem bei etwa 17°.

---

<sup>16</sup> Vgl.: Blauert, Jens, *Räumliches Hören*, S. Hirzel Verlag, Stuttgart 1974, S. 90 ff.

### 4.1.2 Entfernungshören

Der Lautstärkepegel eines Schallereignisses nimmt bei Verdoppelung der Entfernung zur Schallquelle bei gleichmäßiger Abstrahlung um 6 dB ab. Die meisten Schallquellen – wie Musikinstrumente, die menschliche Stimme oder auch Lautsprecher – strahlen verschiedene Frequenzbereiche unterschiedlich ab. Tiefe Frequenzen werden dabei ungerichtet abgestrahlt, mit steigender Frequenz wird die Abstrahlung zunehmend gerichteter. Neben der Pegelabnahme kommt es mit wachsender Entfernung deshalb immer auch zu einer Klangfarbenänderung. Dieser Effekt wird noch verstärkt durch das Phänomen der *Kurven gleicher Lautstärke*. Das menschliche Gehör nimmt bei geringerer Lautstärke tief- und hochfrequente Signalanteile schwächer wahr. Hinzukommt noch, dass mit zunehmendem Abstand (>15 m) hohe Frequenzen durch Luftabsorption gedämpft werden (Dissipation).

Diese Kriterien funktionieren auch bei monauralem Hören und sind vor allem in reflexionsarmer Umgebung, also im Freien, für die Entfernungsbestimmung von Bedeutung. Entscheidend ist hierbei wieder, dass das Schallereignis dem Hörer vertraut ist, da es sonst zu Fehleinschätzungen kommt. Selbst bei bekannten Signalen tendiert der Mensch dazu, das Hörereignis näher als das tatsächliche Schallereignis einzuschätzen.

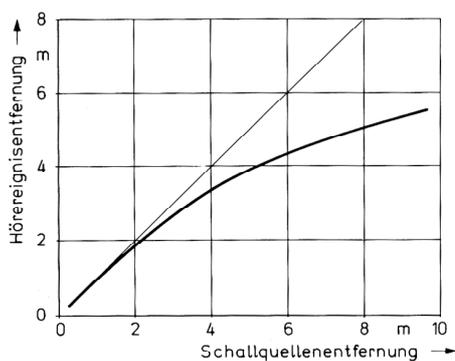


Abb. 3; Lokalisation zwischen Schallquellen- und Hörereignisentfernung<sup>17</sup>

<sup>17</sup> Aus: Blauert, Jens, *Räumliches Hören*, S. Hirzel Verlag, Stuttgart 1974, S.100

In geschlossenen Räumen stehen dem Gehör darüber hinaus noch andere Kriterien zur Verfügung. Am Hörort treffen nach dem Direktschall zuerst einzelne *frühe Reflexionen* (early reflections) ein, die sich in Folge rasch zum Nachhall verdichten. Die Zeit zwischen Direktschall und dem Auftreten der ersten Reflexionen bezeichnet man als *Predelay*. Je geringer der Abstand des Hörers von der Schallquelle ist, desto größer ist der Direktschall gegenüber frühen Reflexionen und dem Nachhall. Mit zunehmender Entfernung sinkt der direkte Schallanteil, während der Diffusschall unabhängig von der Schallquellenentfernung konstant bleibt.

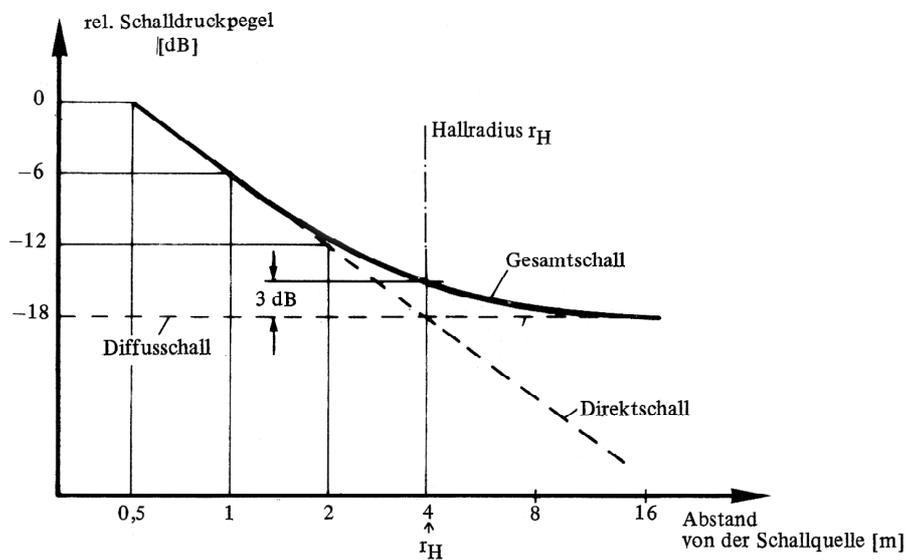


Abb. 4; Überlagerung von Direkt- und Diffusschall<sup>18</sup>

Im Zusammenhang damit steht auch das Phänomen, dass die Einschwingphase eines Signals (Attack) mit zunehmender Entfernung „weicher“ erscheint, da der Anhall - die Zeit, die nach Einsetzen der Schallereignisses vergeht, bis der Schalldruck im Diffusfeld auf seinen endgültigen Wert angestiegen ist - mit zunehmendem Diffusschallanteil auch deutlicher wahrgenommen wird.

<sup>18</sup> Aus: Dickreiter, Michael, *Handbuch der Tonstudiotchnik*, Band 1, K. G. Sauer Verlag, München 1997, S. 37

Bei größerem Abstand zur Schallquelle verringert sich zudem das Predelay, da der Weg des Schalls über die Wände im Verhältnis zum Weg des Direktschalls kleiner wird.

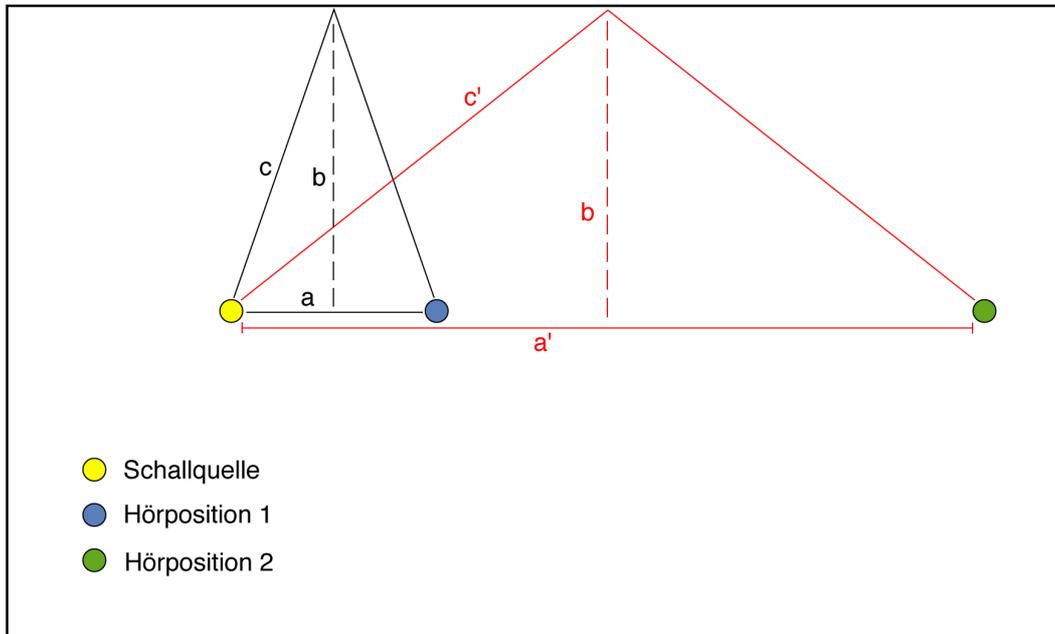


Abb. 5; Predelay bei unterschiedlichen Entfernungen

Beispiel:

Für  $a = 2 \text{ m}$       $a' = 8 \text{ m}$       $b = 3 \text{ m}$

ergibt sich nach Pythagoras:  $c = 3,16 \text{ m}$     und     $c' = 5 \text{ m}$

Somit erhält man für Hörposition 1 ein Predelay von:

$$t_{H1} = t_1 - t_2 = \frac{2 \cdot 3,16 \text{ m}}{344 \text{ m/s}} - \frac{2 \text{ m}}{344 \text{ m/s}} = 0,0126 \text{ s} = 12,5 \text{ ms}$$

und für Hörposition 2 ein Predelay von:

$$t_{H2} = t_1 - t_2 = \frac{2 \cdot 5 \text{ m}}{344 \text{ m/s}} - \frac{8 \text{ m}}{344 \text{ m/s}} = 0,0058 \text{ s} = 5,8 \text{ ms}$$

## 4.2 Lokalisation bei elektroakustischer Wiedergabe

Bei der elektroakustischen Wiedergabe über zwei oder mehr Lautsprecher erfolgt die Lokalisation anders als bei der Wahrnehmung „natürlicher“ Schallquellen. An den Ohren überlagern sich die Signale zweier oder mehrerer Ersatzschallquellen. Geben zwei Lautsprecher kohärente<sup>19</sup> Signale ab, unterscheidet man zwischen drei Fällen von Hörereignissen<sup>20</sup>:

1. Es entsteht ein einzelnes Hörereignis, dessen Ort von den abgestrahlten Signalen beider Lautsprecher abhängig ist. (siehe 4.2.1 Summenlokalisation)
2. Es entsteht ein einzelnes Hörereignis, dessen Ort nur von einem der beiden Lautsprecher abhängig ist, der andere spielt für den Hörereignisort keine Rolle. (siehe 4.2.2 Präzedenzeffekt)
3. Es entstehen zwei getrennte Hörereignisse, bei der die Signale der beiden Lautsprecher mehr oder weniger unabhängig voneinander für die Lokalisation ausgewertet werden.

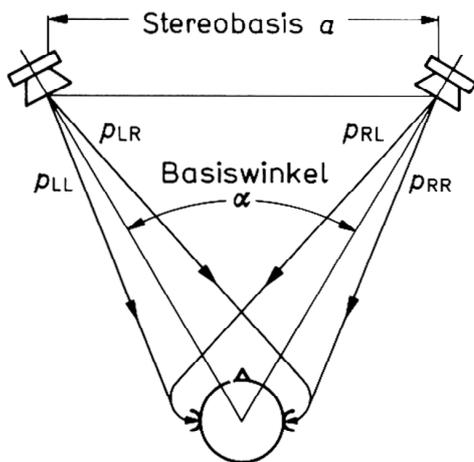


Abb. 6; Überlagerung der Lautsprechersignale bei Stereo-Standardaufstellung.<sup>21</sup>

<sup>19</sup> Der Begriff *kohärent* wird z.T. unterschiedlich verwendet. Im folgenden soll gelten: Zwei Signale sind kohärent, wenn sie identisch sind, oder den gleichen Kurvenverlauf, aber unterschiedliche Amplitude oder Phasenlaufzeit haben. Bei leicht abweichenden Kurvenverläufen (Kohärenzgrad  $< 1$ ), spricht man von Teilkohärenz. (vgl.: Blauert, Jens, *Räumliches Hören*, S. Hirzel Verlag, Stuttgart 1974, S. 191)

<sup>20</sup> Aus: Blauert, Jens, *Räumliches Hören*, S. Hirzel Verlag, Stuttgart 1974, S. 163

<sup>21</sup> Aus: Ebd., S. 161

### 4.2.1 Summenlokalisierung

Die Summenlokalisierungstheorie besagt, dass die Lautsprecher signale sich an den Ohren jeweils so addieren, dass die resultierenden Ohrsignale ein einzelnes Hörereignis hervorrufen. Das Gehör interpretiert die Signale, als kämen sie von einer einzigen, fiktiven Schallquelle, der sogenannten Phantomschallquelle.

Werden zwei identische Signale gleichzeitig und mit gleichem Pegel auf zwei Lautsprecher gegeben, die mit einem Basiswinkel von  $60^\circ$  frontal und symmetrisch zur Medianebene vor dem Hörer stehen, so wird die Phantomschallquelle in der Mitte der Lautsprecherbasis lokalisiert. Verändert man innerhalb bestimmter Grenzen kontinuierlich die Pegel- und/oder Laufzeitunterschiede der Signale, so wandert die Phantomschallquelle entlang der Stereobasis in Richtung des Lautsprechers, der das stärkere bzw. frühere Signal abstrahlt. Je nach verwendetem Signal (Knacke, Tonimpulse, Dauertöne) und Versuchsbedingung (frei beweglicher oder fixierter Kopf) unterscheiden sich die Ergebnisse verschiedener Untersuchungen, immer jedoch wandert das Hörereignis zum lauterem bzw. unverzögerten Lautsprecher.

Pegelunterschiede zwischen den Lautsprechern führen zur Auslenkung der Phantomschallquelle in Richtung des Lautsprechers mit dem höheren Pegel. Bei 15 bis 20 dB Pegeldifferenz wird die Schallquelle vollständig aus der Richtung des lauterem Lautsprechers wahrgenommen. Bis zu einem Winkel von  $\pm 20^\circ$  ist die Auslenkung ziemlich linear und liegt je nach Signal bei etwa  $2,1$  bis  $2,5^\circ/\text{dB}$ . Durch Pegeldifferenzen erzeugte Phantomschallquellen zeichnen sich durch hohe Lokalisationsschärfe und präzise Abbildung aus.

Bei Laufzeitunterschieden wandert die Phantomschallquelle auf der Stereobasis aus der Mitte heraus zum nichtverzögerten Lautsprecher. Ab einer Verzögerung von 1 bis 2 ms wird das Signal nur noch aus der Richtung des früher abstrahlenden Lautsprechers wahrgenommen. Bei breitbandigen Signalen bewirkt eine Differenz von  $\Delta t = 0,1$  ms bei einer Stereo-Lautsprecheranordnung mit  $60^\circ$  Basiswinkel eine Auslenkung von etwa  $3,8^\circ$  aus der Basismitte. Mit steigendem  $\Delta t$  nimmt die Ausdehnung der Phantomschallquelle zu, die Abbildungsschärfe sinkt, außerdem kommt es zu deutlichen Klangfarbenänderungen.

Treten Pegel- und Laufzeitdifferenzen gemeinsam auf, so ist die seitliche Auslenkung der Phantomschallquelle bei gleichsinnigem Zusammenwirken – d.h. das frühere Signal ist auch lauter – größer, bei gegensinnigem Zusammenwirken kleiner, als bei reinem Pegel- bzw. Laufzeitunterschied. Auf diese Weise ist es im Prinzip möglich eine pegelbedingte Auslenkung durch entgegengerichtete Laufzeitdifferenz auszugleichen und umgekehrt. Dies führt allerdings zu einem verwaschenen, z.T. mehrdeutigen Hörereignis mit geringer Lokalisationsschärfe und starker Klangfärbung, was in der Regel unerwünscht ist. Für die Praxis ist interessanter, welche Pegeldifferenz die gleiche seitliche Auslenkung bewirkt wie eine entsprechende Laufzeitdifferenz.

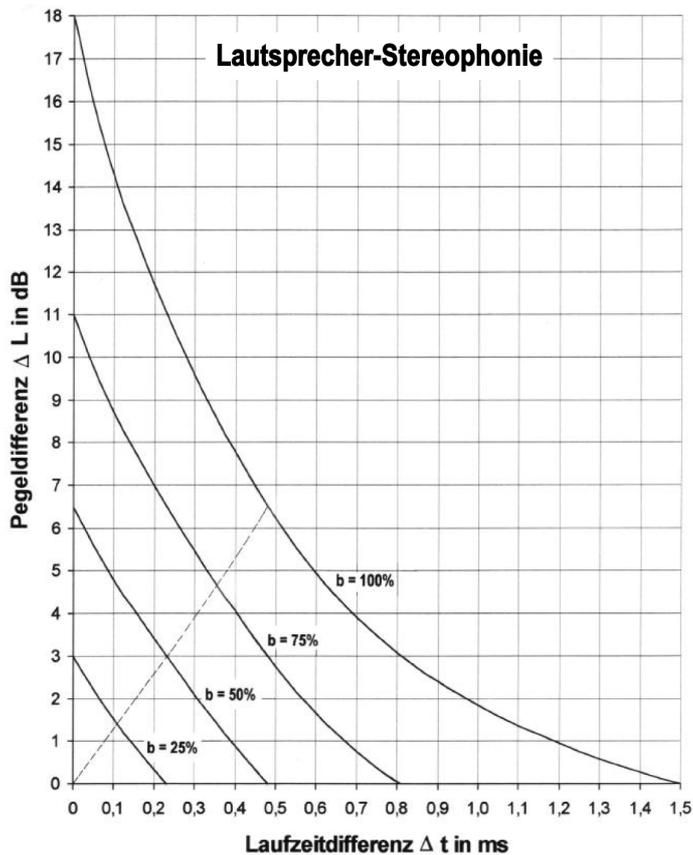


Abb. 7; Kurven gleicher Hörereignisrichtungen, Hörereignisrichtung  $b$  = prozentuale Auslenkung aus der Lautsprecherbasismitte<sup>22</sup>

<sup>22</sup> Aus: Sengpiel, Eberhard, *Kurven gleicher Hörereignisrichtungen*,  
<http://www.sengpielaudio.com/KurvenGleiHoereignis.pdf>, Seitenaufruf vom 10.03.2005

#### 4.2.2 Präzedenzeffekt

Der Präzedenzeffekt – auch bekannt als *Haas-Effekt* oder *Gesetz der ersten Wellenfront* – bezeichnet das Phänomen, dass früher eintreffende Signale bei der Lokalisation gegenüber nacheilenden bevorzugt werden. Dies ist z.B. bei Schallrückwürfen im Raum oder bei verzögerten Lautsprechersignalen der Fall. Ab einem Laufzeitunterschied von ca. 2 ms bleibt der Hörereignisort konstant im unverzögerten Lautsprecher. Selbst wenn der später abstrahlende einen um bis zu 10 dB höheren Pegel besitzt, bleibt der frühere allein für die Lokalisation entscheidend. Das verzögerte Signal kann jedoch Klangfarbenveränderungen, Lautstärkeerhöhungen sowie einen räumlichen Eindruck bewirken.

Erhöht man weiterhin die Laufzeitdifferenz, so werden schließlich zwei zeitlich und richtungsmäßig getrennte Signale wahrgenommen. Die Echoschwelle wird ab etwa 20 - 60 ms erreicht und ist stark von der Signalstruktur und von eventuellen Verdeckungseffekten durch andere Signale abhängig. Bei impulsartigen, perkussiven Signalen werden schon deutlich früher Echos wahrgenommen, als bei dauerhaften, stationären Klängen. Im Übergangsbereich zur Echoschwelle erscheint die Phantomschallquelle meist über die komplette Lautsprecherbasis ausgedehnt.

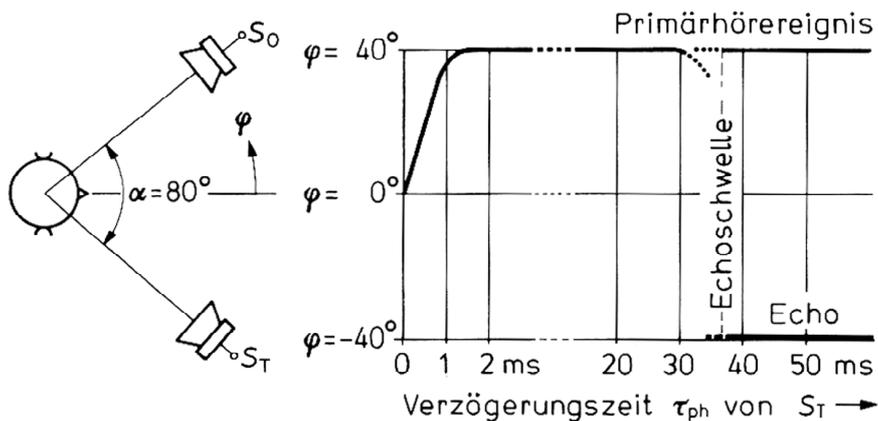


Abb. 8; Hörereignisrichtungen bei Sprache mittlerer Sprechgeschwindigkeit und gleichen Lautsprecherpegeln<sup>23</sup>

<sup>23</sup> Aus: Blauert, Jens, *Räumliches Hören*, S. Hirzel Verlag, Stuttgart 1974, S. 179

### 4.2.3 Assoziationsmodell von Theile

Die Wahrnehmung von Phantomschallquellen kommt im natürlichen Schallfeld nicht vor und tritt nur bei Wiedergabe über zwei oder mehr Lautsprecher auf. Zur Erklärung wird im allgemeinen die Summenlokalisierungstheorie verwendet. Einige Phänomene lassen sich damit allerdings nicht zufriedenstellend erklären.

#### *Der „spektrale Einwand“ zur Summenlokalisierung*

Da sich selbst bei idealer Hörposition unmöglich beide Ohren genau im Schnittpunkt der Lautsprecherachsen befinden können (s. Abb. 6), kommt es folglich auch bei identischen Lautsprechersignalen an jedem Ohr zu geringen Laufzeitunterschieden. Diese bewirken einen an beiden Ohren identischen kammfilterartigen Frequenzgang. Trotz dieser ganz offensichtlich vorhandenen Frequenzauslöschungen und -anhebungen, die nicht zu den spektralen Merkmalen führen, die eine am Ort der Phantomschallquelle befindliche Realschallquelle liefern würde, werden keine oder kaum Klangfärbungen oder eine Änderung des Entfernungseindrucks wahrgenommen.

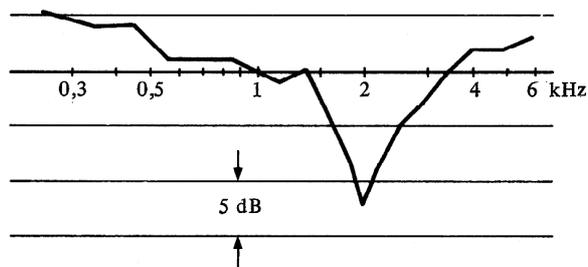


Abb. 9; Frequenzgang am Ohr bei einer Phantomschallquelle in der Mitte der Lautsprecherbasis<sup>24</sup>

---

<sup>24</sup> Aus: Dickreiter, Michael, *Handbuch der Tonstudiotchnik*, Band 1, K. G. Sauer Verlag, München 1997, S. 136

Desweiteren ist es möglich durch äquivalente Laufzeit- und Pegeldifferenzen Phantomschallquellen am gleichen Hörereignisort zu erzeugen. Beides führt zu einem nahezu identischen, unverfärbten Hörereignis, obwohl die Signale sehr unterschiedliche Spektren aufweisen. Beide Phänomene sind mit der Summenlokalisationstheorie nicht zu erklären.

### *Assoziationsmodell*

Das Assoziationsmodell von Günther Theile liefert eine Erklärung für diese wie auch andere Phänomene räumlichen Hörens. Die Theorie geht davon aus, dass die Wahrnehmung eines Hörereignisses stets durch die assoziative Verknüpfung eines eintreffenden akustischen Musters mit einem im Gehirn gespeicherten Muster erfolgt, auch wenn nur Teile des gespeicherten Musters im eintreffenden vorkommen.

Dabei wird ein akustischer Reiz nach einer zuvor erfolgten spektralen Analyse zwei unterschiedlichen, nacheinander folgenden Verarbeitungsstufen zugeführt. Entsprechend der zwei grundlegenden, voneinander unabhängigen, aber stets gemeinsam auftretenden Schallquelleneigenschaften *Ort* und *Gestalt*, durchlaufen die Ohrsignale zunächst eine ortsbestimmende und anschließend eine gestaltbestimmende Assoziationsstufe. Beide Stufen bestimmen stets gemeinsam die Höreigenschaften.

Die Ortsassoziationsstufe vergleicht ankommende Ohrsignale mit Reizmustern, die infolge der Erfahrung bestimmten Hörereignisorten zugeordnet sind. Nur wenn sie bezüglich Zeit und Spektrum mit gespeicherten Mustern vereinbar sind, können sie als Lokalisationsreize interpretiert werden. Dazu werden die ankommenden Signale nach erfolgter Ortsbestimmung zunächst durch inverse Filterung von den richtungsabhängigen Klangfärbungen befreit, was erklärt, warum sie bei der Bildung von Phantomschallquellen nicht wahrgenommen werden; anschließend werden das reine Sendesignal sowie die gewonnene Richtungs- und Entfernungsinformation getrennt an die gestaltbestimmende Assoziationsstufe weitergeleitet.

In der Gestaltassoziationsstufe erfolgt unabhängig vom Hörereignisort die inhaltliche Auswertung der Signale. Dabei wird die Gestalt der Signale ebenfalls mit durch Hörerfahrung gespeicherten Mustern verglichen.

Die Funktion des akustischen Assoziationssystems wird außerdem noch von anderen Systemen beeinflusst. Direkt von einem unspezifischen System, das eine Koordination der Erregungsverarbeitung, sowie eine bewußtseins- und aufmerksamkeitsabhängige Erregungsselektion bewirkt, und indirekt von spezifischen Systemen, die visuelle und/oder taktile Einflüsse hinzusteuern.

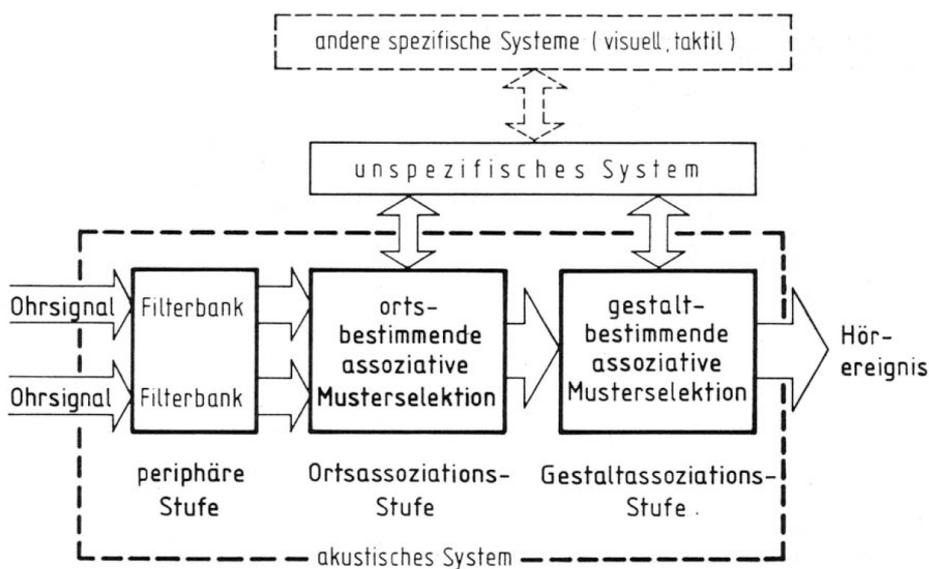


Abb. 10; Funktionsprinzip des Assoziationsmodells von Theile<sup>25</sup>

Für die Lokalisation im überlagerten Schallfeld ergeben sich folgende Erkenntnisse:

Die beiden Lautsprechersignale rufen zunächst zwei unterschiedliche Ortsassoziationen hervor. Sind die Signale in ihrer Gestalt ausreichend ähnlich, verschmelzen sie zu einem einzigen Hörereignis und damit zu einem gemeinsamen Hörereignisort, dem Ort der Phantomschallquelle.

<sup>25</sup> Aus: Theile, Günther, *Über die Lokalisation im überlagerten Schallfeld* (Diss. 1980), [http://www.irt.de/wittek/hauptmikrofon/theile/UEBER\\_DIE\\_LOKALISATION\\_deutsch.pdf](http://www.irt.de/wittek/hauptmikrofon/theile/UEBER_DIE_LOKALISATION_deutsch.pdf), S. 35  
Seitenaufruf vom 21.03.2005

Werden zwei Signale am gleichen Ort lokalisiert, besitzen aber unterschiedliche Gestalt, so werden sie auch getrennt voneinander ausgewertet.

Entsprechend läßt sich der Präzedenzeffekt als *Gesetz des ersten Lokalisationsreizes* verstehen. Lautsprecher-signale mit Laufzeitdifferenzen  $> 2$  ms führen dazu, dass zwei Lokalisationsreize zeitversetzt eintreffen und nur der erste Lokalisationsreiz ausgewertet wird.

### **4.3 Erzeugung von Richtung und Entfernung in der 2.0-Stereofonie**

Eine überzeugende Abbildung von Schallquellen läßt sich erreichen, indem man die Erkenntnisse über die Lokalisation im natürlichen und überlagerten Schallfeld berücksichtigt und deren Gegebenheiten nachbildet.

#### **4.3.1 Richtungsabbildung**

Für die Richtungsabbildung ergeben sich folgende Möglichkeiten:

1. durch Pegelunterschiede:

Diese Methode findet am häufigsten Anwendung und kommt in allen gängigen Mischpulten und softwarebasierten Audio-Workstations zum Einsatz. Hierbei wird das Signal über ein Panorama-Potentiometer (Panpot) auf eine Stereosammelschiene geroutet. In der Mittelstellung erhalten linker und rechter Kanal den gleichen Pegel, bei Links-/Rechtsanschlag wird das Signal nur auf den jeweiligen Kanal gegeben und in den Zwischenstellungen erzeugt das Panpot Pegeldifferenzen entsprechend der verschiedenen Positionen der Phantom-schallquelle auf der Lautsprecherbasis. Um einen gleichbleibenden Lautstärke-eindruck zu gewährleisten müssen die Signalpegel auf dem Stereobus entsprechend der Panpot-Stellung reduziert werden, da durch die Überlagerung von linkem und rechtem Signalanteil Pegelerhöhungen entstehen. In der Mittelstellung beträgt die Dämpfung in jedem Kanal 3 dB und nimmt zu den Seiten kontinuierlich ab.

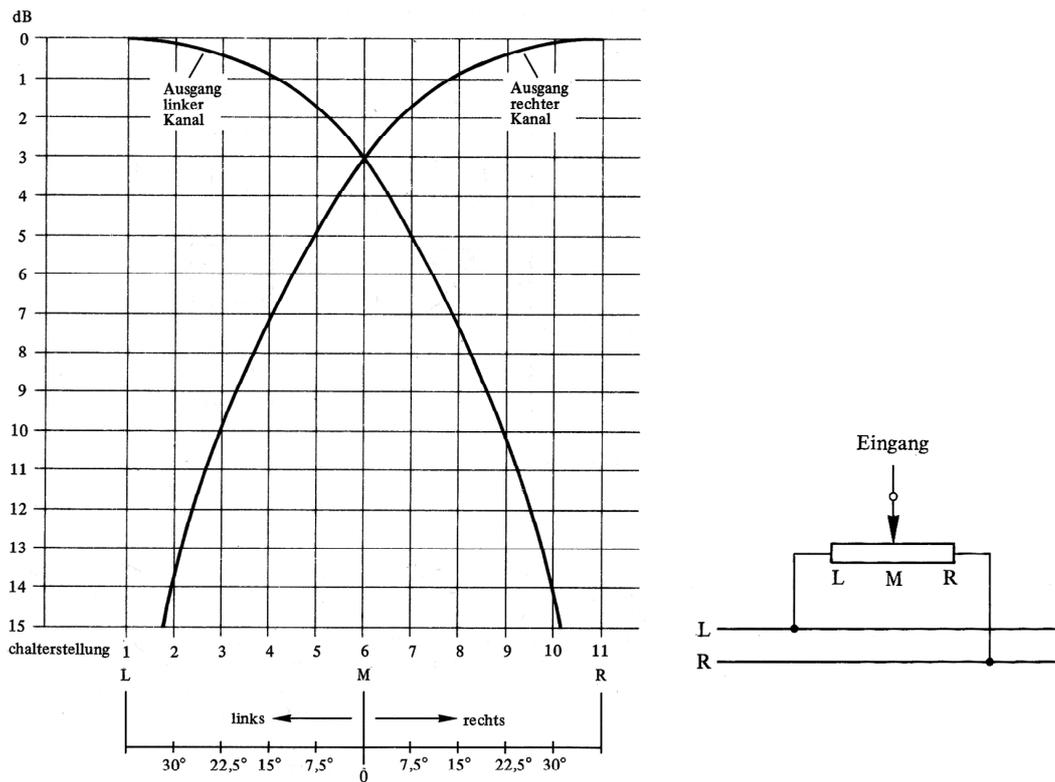


Abb. 11 u. 12; Dämpfungsverlauf und Schaltungsprinzip eines Panpots<sup>26</sup>

Möchte man ein Stereosignal mittels Panpots seitlich positionieren oder die Abbildungsbreite verringern, so ist dies nur bei pegelstereofonen Aufnahmen problemlos möglich. Bei Laufzeit- und Äquivalenzstereofonie geht durch eine nachträgliche Bearbeitung mit Panpots die gute räumliche Abbildung verloren, außerdem entstehen meist unschöne Kammfiltereffekte. Um dennoch das Hörereignis aus der Mitte zu verschieben, empfiehlt es sich in solchen Fällen die Panpots auf Links-/Rechtsanschlag zu lassen und die Kanalfader unterschiedlich einzustellen.

<sup>26</sup> Aus: Dickreiter, Michael, *Handbuch der Tonstudioteknik*, Band 1, K. G. Sauer Verlag, München 1997, S. 374 f.

### 2. durch Laufzeitunterschiede:

Das Signal wird auf beide Kanäle des Stereobusses gegeben, einer davon wird zeitlich verzögert. Das Hörereignis wandert auf der Stereobasis in Richtung des unverzögerten Kanals. Bei monofonen Signalen ist diese Methode weniger zu empfehlen; die Abbildungsschärfe ist geringer als bei Intensitätspannung, außerdem kommt es aufgrund des zeitlichen Versatzes zu Phasenauslöschungen, die sich vor allem bei Monowiedergabe sehr störend bemerkbar machen.

Bei Stereosignalen lassen sich mit diesem Verfahren hingegen wesentlich natürlichere Resultate erzielen als mit Panpots. Das Klangbild bleibt trotz seitlicher Verschiebung räumlich und verliert seine natürliche Abbildung nicht. Versucht man ein stereofon aufgenommenes Signal mit Pegelunterschieden auf der Seite zu positionieren, so wird es zunehmend zu einer Monoschallquelle, da der andere Lautsprecher nicht mehr aktiv ist.

### 3. durch spektrale Unterschiede:

Spektraldifferenzen, wie sie beim natürlichen Hören durch Abschattung und Beugung der Schallwellen an Kopf und Ohrmuschel entstehen, lassen sich aufnahmeseitig mit Trennkörper-Stereomikrofonen (Kugelfläche, OSS, Clara, SASS) oder dem Kunstkopf erzielen. Die Spektraldifferenzen sind jedoch nach Ansicht vieler Tonmeister für die Richtungsabbildung bei Lautsprecherwiedergabe ungeeignet, da sie von jedem Lautsprecher unter erneuter Beeinflussung durch die Ohrmuscheln auf jeweils beide Ohren gelangen, was Klangfärbungen und z.T. unscharfe Phantomschallquellen zur Folge hat. Bei Kopfhörerwiedergabe hingegen können diese sogenannten kopfbezogenen Stereofonieverfahren besonders Kunstkopfaufnahmen eine sehr realistische Abbildung ohne *Im-Kopf-Lokalisation* erzeugen.

Es gibt aber auch gegensätzliche Meinungen bezüglich der Verwendung von Trennkörpermikrofonen für die Lautsprecherwiedergabe, hauptsächlich von den Herstellern solcher Mikrofone.

Mittlerweile gibt es auch Computerprogramme, die diese kopfbezogenen Übertragungsfunktionen, sogenannte HRTF (Head Related Transfer Function) nachbilden und z.T. beeindruckende Räumlichkeitseffekte erzielen. So lassen sich Hörereignisse u.a. auch hinter dem Hörer positionieren. Dies funktioniert allerdings nur bei Kopfhörerwiedergabe überzeugend, da die Signale dabei nicht erneut durch die Ohrmuscheln beeinträchtigt werden.

Erzeugt man bei einem Monosignal mittels Equalizer Frequenzunterschiede zwischen linkem und rechten Kanal, so führt dies zu keiner brauchbaren Richtungsabbildung. Es entstehen allenfalls Pseudostereofonie-Effekte. Je nach Frequenzzusammensetzung bzw. gespielter Tonlage wird das Hörereignis stärker links bzw. rechts lokalisiert.

### **4.3.2 Entfernungsabbildung**

Für den Entfernungseindruck sind vor allem die ersten Reflexionen und wann sie in Bezug auf das direkte Schallsignal eintreffen von Bedeutung. Um den Eindruck von Nähe zu erreichen, muß der Pegel der ersten Reflexionen und der sich anschließende Nachhall gering und mit relativ großem Predelay hinzuge-mischt werden. Möchte man dagegen eine weit entfernte Schallquelle simulieren, so ist der Pegel der Reflexionen und des Nachhalls größer und das Predelay kleiner zu wählen.

Zusätzlich kann man das Signal noch im Bass- und Höhenbereich absenken, um die Dämpfung der Höhen durch Luftabsorption bei zunehmender Entfernung und den Effekt der *Kurven gleicher Lautstärke* nachzubilden. Bei leiseren, also entfernteren Signalen, werden Bässe und Höhen im Verhältnis zum mittleren Frequenzbereich leiser wahrgenommen. Zur perfekten Illusion komprimiert man noch den Attack des Signals, um die Einschwingphase weicher und somit die Schallquelle weiter entfernt wirken zu lassen.

Die Länge der Nachhallzeit ist für den Entfernungseindruck unerheblich und gibt nur Aufschluß über die Größe und Beschaffenheit des Raumes. Man kann problemlos eine Schallquelle trotz langem Nachhall in einer Mischung vorne plazieren. Soll der Hall deutlich zu hören sein, das Signal aber trotzdem nicht an Nähe und Direktheit verlieren, so erhöht man das Predelay, um den Diffusanteil gewissermaßen vom Direktsignal zu entkoppeln. Nach dieser Methode wird bei Popballaden sehr häufig mit dem Lead-Gesang verfahren. Predelay-Zeiten von 60 ms und mehr sind dabei nicht ungewöhnlich.

Eberhard Sengpiel weist noch auf eine andere Möglichkeit zur Erzeugung von Nähe und Distanz hin.<sup>27</sup> Unter Berücksichtigung der *richtungsbestimmenden Bänder* läßt sich ein Signal präsenter oder diffuser abbilden.

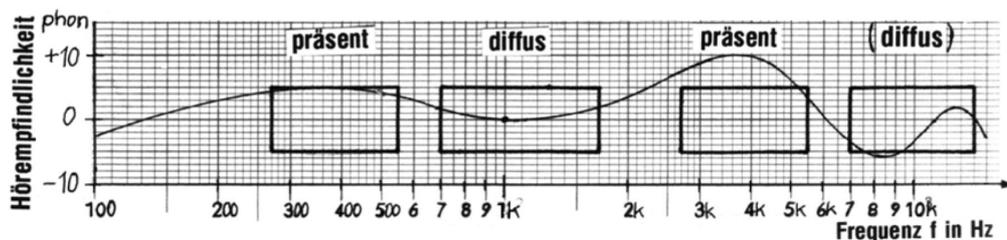


Abb. 13; Die Präsent- und Diffus-Bänder in der 80 dB-Hörfähigkeitskurve<sup>28</sup>

Die *vorne*-Bänder in den Frequenzbereichen um 350 Hz und 4 kHz entsprechen dabei dem Präsenzbereich, die *hinten*-Bänder im Bereich um 1,2 kHz und 12 kHz und das *oben*-Band bei 8 kHz dem Diffusbereich. Ein Anheben bzw. Absenken des Pegels in den entsprechenden Frequenzbereichen, kann einen präsenteren und damit näheren, bzw. einen diffuseren und damit distanzierteren Eindruck bewirken.

---

<sup>27</sup> Aus: Sengpiel, Eberhard, *Die Bedeutung der Blauertschen Bänder*, <http://www.sengpielaudio.com/DieBedeutungDerBlauertschenBaender.pdf>, Seitenaufruf vom 10.03.2005

<sup>28</sup> Aus: Ebd.

## 5. Theoretische Überlegungen zur 3/2-Stereofonie

Die 3/2-Stereofonie stellt ebenso wie die 2.0-Stereofonie eine Transformation dar, bei der ein natürliches Schallereignis auf die Lautsprecherwiedergabe übertragen wird. Dennoch kann man mit einer guten Surround-Mischung einen bedeutend realistischeren Höreindruck erzielen als mit nur zwei Kanälen.

In der 3/2-Stereofonie bleibt einiges genau gleich wie in der Zweikanal-Stereofonie. Nach wie vor können Signale als Phantomschallquellen auf der vorderen Lautsprecherbasis oder im „virtuellen“ Raum dahinter platziert werden. Durch die zusätzlichen Lautsprecher in der Fünfkanal-Stereofonie ergeben sich weitere Lautsprecherbasen, so dass sich theoretisch Schallquellen auf dem gesamten Lautsprecherkreis und dem virtuellen Raum dahinter abbilden lassen.

### 5.1 Unterschiede zur 2.0-Stereofonie

Der Hauptunterschied zur Zweikanal-Stereofonie besteht darin, dass in der Fünfkanal-Stereofonie nun wirklich eine zweite Dimension existiert. Bei Zweikanal-Stereo gibt es de facto nur eine Dimension, die räumliche Tiefe läßt sich nur mittels den in Kapitel 4.3.2 (Entfernungsabbildung) beschriebenen Methoden simulieren. Im Gegensatz dazu ist in der Fünfkanal-Stereofonie ein zweidimensionales Feld innerhalb des Lautsprecherkreises vorhanden, in dem sich Schallquellen theoretisch ohne das zusätzliche Hinzufügen von Reflexionen platzieren lassen sollten.

Desweiteren besteht die Möglichkeit Signale seitlich und hinter dem Hörer zu positionieren und so den Eindruck von Umhüllung oder gezielte Effekte zu schaffen. Gerade in der Popmusik eröffnen sich dadurch beim Mischen ganz neue, kreative Gestaltungsmöglichkeiten.

Auch kann man davon ausgehen, dass die Informationsdichte in der Fünfkanal-Stereofonie höher sein kann als in der Zweikanal-Stereofonie. Es lassen sich dichtere Arrangements realisieren, die bei zweikanaliger Wiedergabe überladen klingen würden.

Bildet man ein Signal als Phantomschallquelle in der Mitte zwischen linkem und rechtem Lautsprecher ab, wandert es bei seitlicher Hörposition mit. Mit Hilfe des Center-Lautsprechers lassen sich jetzt Schallquellen stabil vorne in der Mitte plazieren. Zudem ist diese im Vergleich zu einer Phantomschallquelle präziser lokalisiert und klingt weniger verwaschen.

Neben der reinen Mittenabbildung bietet der Center-Kanal aber noch weitere Funktionen. Möchte man eine Schallquelle z.B. bei 25% seitlicher Auslenkung (entspricht einem Winkel von  $-7,5^\circ$  bei 3/2-Stereo Standard und optimaler Hörposition) plazieren, so kann man dies durch Bildung einer Phantomschallquelle aus linkem und rechten Kanal wie in der Zweikanal-Stereofonie erreichen oder durch Kombination von linkem und Center-Kanal oder sogar allen drei vorderen Lautsprechern. Diese auf verschiedene Weise erzeugten Phantomschallquellen weisen z.T. große Unterschiede bezüglich Abbildungsschärfe, Ausdehnung, Entfernungseindruck und Klangfarbe auf.<sup>29</sup>

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	$\Delta L(L/R) = 1,5 \text{ dB}$ ; $\Delta t(L/R) = -0,1 \text{ ms}$ ; $\Delta L(L/C) = 9 \text{ dB}$ ; $\Delta t(L/C) = -1,0 \text{ ms}$ ;
...	...	...

Abb. 14; verschiedene Möglichkeiten der Phantomschallquellenbildung bei 25% seitl. Auslenkung<sup>30</sup>

<sup>29</sup> Vgl.: Wittek, Helmut, *Untersuchungen zur Richtungsabbildung mit L-C-R Hauptmikrofonen*, (Dipl. 2000), [http://www.irt.de/wittek/hauptmikrofon/DA\\_Helmut\\_Wittek.pdf](http://www.irt.de/wittek/hauptmikrofon/DA_Helmut_Wittek.pdf), S. 15 ff., Seitenaufwurf vom 21.3.2005

<sup>30</sup> Aus: Ebd: S. 16

### *Die „zweite Dimension“*

Wie bereits erwähnt, existiert in der Fünfkanal-Stereofonie innerhalb des Lautsprecherkreises ein zweidimensionales Feld, in dem Schallquellen plaziert werden können. Auch hier sind wieder zahlreiche Lautsprecherkombinationen für die Phantomschallquellenbildung möglich, je nachdem wo die Schallquelle abgebildet werden soll. Es stellt sich die Frage, inwiefern dieses Feld ausgenutzt werden kann und wie sich Phantomschallquellen darin verhalten. So sollte es theoretisch möglich sein, eine echte Tiefenstaffelung zu erzielen, also eine räumlich nach hinten gestaffelte Abbildung von Schallquellen nur innerhalb des Lautsprecherkreises ohne den „virtuellen“ Raum hinter den Lautsprechern zu verwenden.

### **5.2 Probleme bei der Positionierung von Schallquellen**

Die Abbildung von Schallquellen vor und hinter dem Hörer funktioniert gut. Zwar wandern die Phantomschallquellen bei seitlichen Veränderungen der Hörposition mit, bei kleineren Bewegungen bleibt das Hörereignis aber relativ fest an seinem Platz und läßt sich präzise lokalisieren.

Ganz anders verhält es sich dagegen bei seitlich platzierten Schallquellen. Diese sind in besonderem Maße instabil. Schon kleinste Kopfbewegungen lassen das Hörereignis seitlich zwischen vorderem und hinterem Lautsprecher springen.<sup>31</sup> Mitunter werden je nach Frequenz und Tonlage des Signals auch zwei räumlich getrennte Hörereignisse an den jeweiligen Lautsprechern wahrgenommen.<sup>32</sup>

---

<sup>31</sup> Vgl: Sengpiel, Eberhard, *Seitliche Phantomschallquellen gibt es nicht...*, <http://www.sengpielaudio.com/SeitlichePhantomschallquellen.pdf>, Seitenaufruf vom 10.03.2005

<sup>32</sup> Vgl.: Blauert, Jens, *Räumliches Hören*, S. Hirzel Verlag, Stuttgart 1974, S. 196

Das Assoziationsmodell von Theile kann eine Erklärung dazu liefern:<sup>33</sup>

Bei einer seitlich platzierten Schallquelle, bei der ein monofones Signal gleichzeitig und mit gleichem Pegel auf den vorderen und hinteren seitlichen Lautsprecher (L und LS bzw. R und RS) gegeben wird, kommt es aufgrund einer zu kleinen Laufzeitdifferenz der sich überlagernden Lautsprechersignale zu einer Störung der Lokalisationsreizselektion. An jedem Ohr befinden sich beide Maxima an denselben Stellen, so dass die Lokalisationsreize in der Ortsassoziationsstufe nicht ausreichend unterschieden werden können.

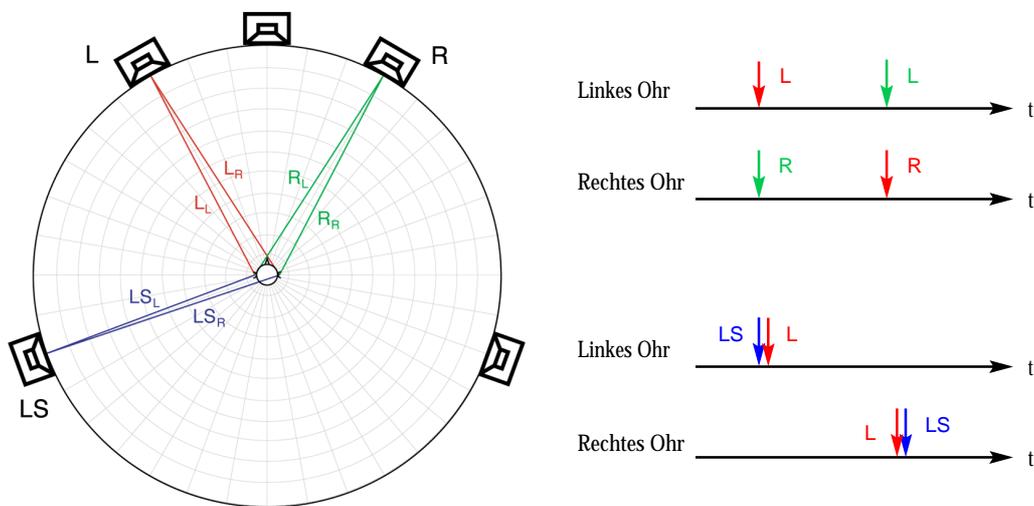


Abb. 15; interaurale Laufzeitmerkmale bei frontaler und seitlicher Beschallung<sup>34</sup>

Wird ein monofones Signal irgendwo auf der Stereobasis zwischen den vorderen Lautsprechern platziert, so rufen die zwei an den Ohren eintreffenden Signale zwar unterschiedliche Ortsassoziationen hervor, diese sind aber in ihrer Gestalt so ähnlich, dass sie im Gehirn als ein einziges Hörereignis interpretiert werden und damit einen gemeinsamen Hörereignisort hervorrufen.

<sup>33</sup> Vgl.: Theile, Günther, *Über die Lokalisation im überlagerten Schallfeld* (Diss. 1980), [http://www.irt.de/wittek/hauptmikrofon/theile/UEBER\\_DIE\\_LOKALISATION\\_deutsch.pdf](http://www.irt.de/wittek/hauptmikrofon/theile/UEBER_DIE_LOKALISATION_deutsch.pdf), S.48 f. Seitenaufruf vom 21.03.2005

<sup>34</sup> Nach.: Ebd., S. 48

Bei seitlichen Phantomschallquellen funktioniert dies vermutlich deshalb nicht mehr, weil die Lokalisation zwischen seitlichem vorderen und hinteren Lautsprecher nicht mehr aufgrund von Laufzeit- und Pegelunterschieden erfolgt, sondern hauptsächlich durch die spektrale Filterung, die durch die Ohrmuscheln und den äußeren Gehörgang verursacht werden.

Eine Phantommitte zwischen L und LS bzw. R und RS ist demnach instabil, weil das von hinten kommende Signal eine ganz andere Filterung durch die Ohrmuschel erfährt, als das von vorne kommende Signal. Die eintreffenden Signale haben bei der Interpretation im Gehirn nicht mehr genug Ähnlichkeit, um eine gemeinsame Gestaltassoziation hervorzurufen und damit als ein einzelnes, klar lokalisierbares Hörereignis interpretiert zu werden.

Ein wesentlicher Anreiz Surround-Mischungen zu gestalten, wäre die Möglichkeit, die bisherige Zweikanal-Stereobasis zu verbreitern. Dies scheint nach den Überlegungen zur Seitenabbildung aber nicht so ohne weiteres möglich. Gerade diese fehlende Seitenabbildung ist ein oft zitiertes Problem und wird bei so mancher bereits existierenden Surround-Produktion bemängelt. Häufig wird kritisiert, dass die vorne positionierten Instrumente und der hinten erzeugte Raum keine Einheit bilden, das Gesamtklangbild scheint zu zerfallen. Einzelne hinten platzierte Hörereignisse wirken ohne seitliche Verbindung irritierend und werden als aufgesetzt empfunden.

Es gilt zu herauszufinden, inwiefern es in gewissen Grenzen doch möglich ist Schallquellen an den Seiten abzubilden und somit die Zweikanal-Stereobreite zu erweitern und wie sich ein Zerfallen des Gesamtklangbildes vermeiden läßt.

Eine wirklich stabile Abbildung seitlich platzierter Schallquellen läßt sich mit der 3/2-Lautsprecheranordnung vermutlich nicht erreichen. Eine andere Boxenaufstellung wäre für Mehrkanal-Audio daher eigentlich sinnvoller. Da vordere Phantomschallquellen ohne Probleme lokalisiert werden können, und der LFE-Kanal bei Audioproduktionen prinzipiell auch nicht notwendig ist, könnte man diese beiden Kanäle auch zur Verbesserung der Seitenabbildung verwenden.

Entsprechende Vorschläge<sup>35</sup> wurden in den letzten Jahren auch auf verschiedenen Tonmeistertagungen diskutiert.

Ein anderer Ansatz unter Beibehaltung der 3/2-Kanalaufteilung wurde von Wolfgang Zieglmeier und Günther Theile verfolgt.<sup>36</sup> Zur Verbesserung der Seitenabbildung wurde der Einsatz von zwei Stützlautsprechern bei 70° untersucht. Dabei wurde das Stützsiegel mittels Dolby-Surround-Pro-Logic-Decoder generiert. Auch hierbei wurde eine Verbesserung der Seitenabbildung festgestellt.

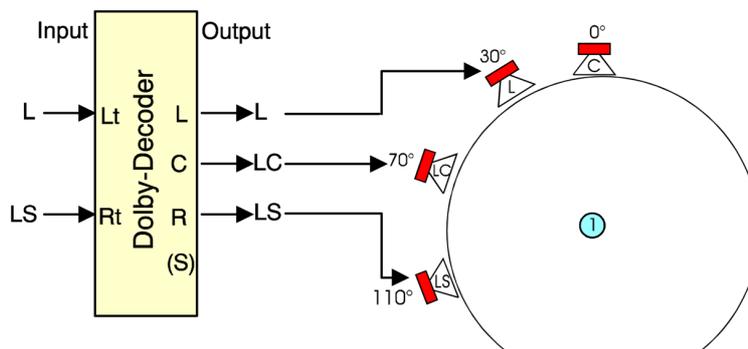


Abb. 16; Einsatz des Dolby-Decoders zur Erzeugung eines Stützsiegels (LC)<sup>37</sup>

Es stellt sich allerdings die Frage was all diese Verbesserungsvorschläge in der Praxis bringen. Mehrkanal-Audio wird nur dann eine Chance haben sich auf dem Markt durchzusetzen, wenn die bestehende 3/2-Lautsprecheranordnung der Heimkino-Surroundanlagen übernommen wird. Die Mehrheit der Konsumenten wird kaum davon zu überzeugen sein, dass sie zum Musikhören eine andere Boxenaufstellung benötigen als zum Fernsehen, geschweige denn sich noch weitere Lautsprecher ins Wohnzimmer stellen.

---

<sup>35</sup> Vgl.: 6.0-Format (2+2+2) von W. Dabringhaus oder David Chesky

<sup>36</sup> Vgl.: Zieglmeier, W./Theile, G., *Darstellung seitlicher Schallquellen bei Anwendung des 3/2-Stereo Formates*, [http://www.irt.de/wittek/hauptmikrofon/theile/Seitl.Lokalisation\\_TMT1996.pdf](http://www.irt.de/wittek/hauptmikrofon/theile/Seitl.Lokalisation_TMT1996.pdf), Seitenaufwurf vom 21.03.2005

<sup>37</sup> Aus: Ebd., S. 9

Ein Hauptziel bei Surround-Mischungen ist es, eine möglichst stabile, platzunabhängige Lokalisation zu erzielen, d.h. eine Mischung so zu gestalten, dass auch Hörer außerhalb der optimalen Hörposition eine akzeptable Abbildung der Schallquellen erfahren. Wenn sich ein Hörer innerhalb des Lautsprecherkreises bewegt, sollten sich die Hörereignisse nicht zu stark verschieben, vor allem nicht in andere Richtungen kippen. Ein Hörereignis, das bei optimaler Hörposition vorne lokalisiert wird und vom Tonmeister dort plaziert wurde, sollte bei nicht optimaler Hörposition nicht plötzlich hinten wahrgenommen werden.

Vermutlich wird hier vermehrt mit Laufzeitunterschieden und Filterungen, als mit Pegeldifferenzen gearbeitet werden müssen, da bei Richtungserzeugung durch Pegeldifferenzen eine wesentlich größere Abhängigkeit von der Hörposition besteht. Ein Pegelunterschied wird schon mit einer kleinen Veränderung der Hörposition ausgeglichen. Auch bei ungünstigeren Hörpositionen sollten vorne plazierte Schallquellen unbedingt nach wie vor vorne lokalisiert werden, ansonsten wird das ausbleibende Surround-Hörerlebnis beim Konsumenten schnell zum Frust. In vielen Wohnzimmern werden die hinteren Lautsprecher wahrscheinlich nie oder zumindest selten optimal aufgestellt werden können, und sich im Verhältnis zu den vorderen Lautsprechern oft zu nahe am Zuhörer befinden. Hier besteht die große Gefahr, dass das Klanggeschehen hauptsächlich hinten wahrgenommen wird. Zwar gibt es in den meisten Surround-Verstärkern die Möglichkeit die einzelnen Lautsprecher mit Delays dem eigenen Wohnraum anzupassen, aber man sollte nicht unbedingt davon ausgehen, dass der Durchschnittsanwender die richtigen Korrekturen vornimmt.

Es ist anzunehmen, dass eine stabile *vorne*-Lokalisation unter Berücksichtigung des Präzedenzeffektes erreicht werden kann. Signale, die vorne lokalisiert werden sollen, müssen zuerst von vorne beim Hörer eintreffen und folglich von den hinteren Lautsprechern verzögert und eventuell zusätzlich gefiltert abgestrahlt werden.

Bei laufzeitbasierter Richtungserzeugung ergibt sich ein anderes Problem. Die Lokalisation ist hier abhängig von der Entfernung der Lautsprecher zueinander. Ist der Lautsprecherkreis beim Hörer deutlich größer oder kleiner als der im Tonstudio, in welchem die Mischung entstanden ist, so können sich unterschiedliche Richtungsabbildungen ergeben. Nur an der optimalen Hörposition stimmt die Lokalisation immer, da (theoretisch) alle Lautsprecher die gleiche Entfernung zum Hörer haben.

Wahrscheinlich liefert eine Kombination aus Laufzeit- und Pegelunterschieden und passenden Filterungen die überzeugendsten Misch-Resultate. Wie Signale bezüglich Pegeln, Filterungen und Verzögerungen auf die einzelnen Lautsprecher gegeben werden müssen, um eine gewünschte, möglichst Hörplatz-unabhängige Abbildung zu erreichen, soll in praktischen Versuchen herausgefunden werden.

## 6. Misch-Experimente

Auf Grundlage der Lokalisationstheorien und der Überlegungen zur Fünfkanal-Stereofonie wurden verschiedene Misch-Experimente und Hörversuche durchgeführt.

### 6.1 Centerlautsprecher

Über Sinn und Verwendungsmöglichkeiten des Centerlautsprechers wurde schon viel diskutiert. Die Vorteile beim Kinoton liegen klar auf der Hand, dagegen ist die Verwendung bei Audioproduktionen weniger ersichtlich. Es stellt sich die Frage, ob der Center überflüssig ist, da in der Zweikanal-Stereofonie die Phantommitte bisher auch gut funktioniert.

Versuch 1a:

Ein Schlagzeug wurde auf L-R gemischt, einzig die Bassdrum wurde auf den Center gelegt. Zum Vergleich wurde sie auch als Phantomschallquelle aus L und R erzeugt, außerdem eine kombinierte Version, bei der die Bassdrum auf alle drei Frontkanäle mit gleichem Pegel gegeben wurde. Die Lautstärken wurden jeweils angeglichen.

Versuch 1b:

Wieder wurde ein Schlagzeug auf L-R gemischt, diesmal wurde die Snare auf den Center geroutet. Ebenso wurde wieder eine Version mit Phantommitte und eine Kombination aus L-C-R erstellt.

Versuch 1c:

Eine Sängerin wurde mit einem Brauner Velvet aufgenommen. Verglichen wurde wieder Signal auf Centerlautsprecher, Phantommitte aus L und R und die Kombination. Außerdem wurde ein Vergleich zwischen Center und Phantommitte in Verbindung mit einer Band, die auf L-R gemischt wurde, gemacht.

### Versuch 1d:

Ein Flügel wurde mit drei AKG C414 B-TL II aufgenommen. Die Mikrofone wurden in einem Abstand von ca. 5 cm in einer erweiterten AB-Aufstellung (AB+C) – ähnlich dem von Herrmann und Henkels entwickelten INA-3-Verfahren<sup>38</sup> – auf den Flügel gerichtet. Die Mikrofone waren jeweils ca. 30 cm auseinander und auf Nierencharakteristik gestellt. Es wurden folgende Mischungen erzeugt: zwei Monovarianten, Centersignal auf Center bzw. Centersignal auf L-R, AB klassisch auf L-R und eine 3.0-Variante mit zusätzlichem Centersignal. Außerdem wurde eine Version erstellt, bei der das Centermikrofon zur AB-Stereofonie auf L-R hinzugemischt wurde und umgekehrt eine 3.0-Version, bei der das Centersignal durch Summierung von A und B gewonnen wurde. Die Lautstärkerverhältnisse wurden jeweils wieder angeglichen.

### Auswertung Versuch 1a:

Wie nicht anders zu erwarten, ist die Abbildung der Bassdrum im Center gegenüber der Phantommitte und der kombinierten Version klarer und präziser lokalisiert. Auch bei seitlichen Hörpositionen behält sie ihren Platz fest in der Mitte. Dies wirkt aber zusammen mit dem restlichen Schlagzeug, welches seitlich mitwandert, eher unangenehm. Hier zeigen sich die anderen beiden Mischungen homogener. Bezüglich der Klangfarbe klingt die Bassdrum im Center wiederum am neutralsten, allerdings ergeben sich bei der Phantommitteneubildung Kammfiltereffekte, die den Klang der Bassdrum in diesem Fall positiv beeinflussen. Hier muß allerdings von Aufnahme zu Aufnahme mit Unterschieden gerechnet werden, je nach Instrument, Mikrofonawahl und -positionierung und Equalizer-Bearbeitung.

Die L-C-R-Version bietet eine ausgewogene und stabile Abbildung, klanglich fällt sie vor allem gegenüber der Phantommitteneubildung ab.

---

<sup>38</sup> Herrmann, U./Henkels, V., *Vergleich fünf verschiedener Surround-Hauptmikrofonverfahren*, Diplomarbeit, FH Düsseldorf, März 1997

### Auswertung Versuch 1b:

Der Versuch mit der Snare liefert ähnliche Ergebnisse. Die Snare im Center wirkt sehr präzise, aber auch hier stört bei seitlicher Hörposition, dass sich das restliche Klanggeschehen verschiebt, die Snare jedoch nicht. Die positiven Frequenzauslöschungen bei der Phantommitte bleiben in diesem Versuch aus, dennoch wirkt die Snare gerade durch die verwaschenere, breitere Abbildung natürlicher. Sie bekommt eine realistischer klingende, räumliche Ausdehnung im Gegensatz zu der punktuell und eng wirkenden Snare im Center. Die kombinierte Version zeigt wieder eine gute Abbildung, klanglich ist sie aber etwas schwächer als die anderen beiden.

### Auswertung Versuch 1c:

Bei der Stimme machen sich die Klangfärbungen beim Zusammenwirken mehrerer Lautsprecher deutlich negativer bemerkbar. Über den Center wird die Stimme wesentlich neutraler und natürlicher wiedergegeben. Dennoch wirkt die räumliche Abbildung bei den anderen beiden Varianten durch die Unschärfe fast angenehmer. Im Zusammenhang mit der auf L-R gemischten Band stört wieder das die Stimme im Gegensatz zu den restlichen Instrumenten nicht mitwandert. Bei idealer Hörposition hebt sich die Stimme im Center besser vom Bandsound ab, fällt aber dadurch auch stärker aus dem Gesamtmix heraus.

### Auswertung Versuch 1d:

Da der Flügel eine räumlich ausgedehnte Klangquelle darstellt und man auch von Aufnahmen eine breite Abbildung gewohnt ist, überzeugt bei den Monoversionen die Phantomschallquelle. Sie wirkt nicht ganz so extrem eng, wie allein aus dem Centerlautsprecher. Bei den stereofonen Mischungen klingt die dreikanalige Version, bei der jedem Lautsprecher ein Mikrofon zugeordnet wurde, in Bezug auf Abbildung, Klangfarbe und Stabilität am überzeugendsten. Die AB-Mischung ist klanglich auch sehr schön. Die Abbildung ist jedoch schon sehr breit, was in Popmusik vielleicht noch nicht stört, bei diesem klassischen Stück (Chopin) aber eher grenzwertig ist.

Die beiden Varianten, bei denen das Centersignal aus der Summierung von L und R gewonnen bzw. das Centermikrofon auf L-R hinzugemischt wurde, sind klanglich aufgrund von Phasenauslöschungen schlechter. Die Stabilität ist mit dem L-R-generierten Centersignal besser als ohne, allerdings geht der räumliche Eindruck etwas verloren, so als würde man das Stereosignal mit Panpots in der Breite verringern.

### Zusammenfassung Versuch 1:

Der Centerlautsprecher ist definitiv ein Zugewinn, als reiner Ersatz der Phantommitte, stellt er aber keine befriedigende Lösung dar. In Bezug auf die Stabilität liefert die Kombination aus Center und Phantommitte bei monofonen Signalen die besten Ergebnisse, allerdings häufig mit klanglichen Einbußen. Hier wird man mit verschiedenen Pegelverhältnissen zwischen Center und L-R und eventuell mit Verzögerungen experimentieren müssen, um das jeweils beste Resultat bezüglich Abbildung, Stabilität und Klangfarbe zu erzielen. Es ist daher beim Mischen zu empfehlen, sich erst über die Positionierung des Instruments klar zu werden, bevor man klangliche Bearbeitungen vornimmt, da sich durch die Kammfiltereffekte zusätzliche Klangänderungen ergeben. Diese können aber – gerade in der Popmusik – durchaus eine positive Wirkung haben und klangbildend eingesetzt werden.

Bei stereofonen Signalen kann der Center eine deutliche Bereicherung sein. Man sollte jedoch bei der Aufnahme darauf achten, dass die Mikrofone untereinander einen ausreichend großen Abstand haben, da sonst durch das Übersprechen ebenfalls Phasenauslöschungen auftreten. Der Abstand zweier benachbarter Mikrofone sollte dabei mindestens dreimal so groß wie der Abstand zur Schallquelle sein.<sup>39</sup> Hat man nur eine zweikanalige Aufzeichnung und möchte aus Gründen der Stabilität dennoch ein Centersignal erzeugen, sollte man auch hier ausprobieren bis das günstigste Pegel- und Laufzeitverhältnis gefunden wurde. Allgemeingültige Werte können hier vermutlich nicht gefunden werden, da es von Aufnahme zu Aufnahme zu große Unterschiede gibt.

---

<sup>39</sup> Vgl.: 3:1-Regel: Görne, Thomas, *Mikrofone in Theorie und Praxis*, Elektor-Verlag GmbH, Aachen 1996, S. 229

## 6.2 Seitliche Abbildung

Nach den Überlegungen in Kapitel 5.2 sind seitliche Phantomschallquellen instabil. Der Versuch, eine Schallquelle mittels Surround-Panpot in einer fließenden Bewegung vom vorderen zum hinteren Lautsprecher einer Seite zu bewegen, bestätigt dies. Das Hörereignis scheint sich zunächst kaum vom vorderen Lautsprecher zu lösen und springt dann plötzlich und wird im hinteren Lautsprecher lokalisiert. Dennoch soll untersucht werden, in welchen Grenzen und unter welchen Bedingungen es möglich ist, Schallquellen an den Seiten zu plazieren. In den folgenden Versuchen wurde nur die linke Seite untersucht, in der Annahme, dass die Ergebnisse aufgrund der Ohrsymmetrie genauso für die rechte Seite gelten.

Versuch 2a:

Eine Sängerin, aufgenommen mit einem Brauner Valvet, wurde gleichzeitig und mit gleichem Pegel auf L und LS gegeben. In 1 dB-Schritten wurde der Pegel im hinteren Lautsprecher abgesenkt, wobei der Pegel des vorderen konstant blieb. Entsprechend wurde auch mit dem umgekehrten Fall (LS konstant, L leiser) verfahren.

Versuch 2b:

Wieder wurde die Stimme auf L und LS geroutet, diesmal wurde der hintere Lautsprecher in 0,1 ms-Abstufungen verzögert, ebenso wurde der umgekehrte Fall (LS konstant, L verzögert) untersucht.

Versuch 3a:

Ein Flügel wurde monofon mit einem AKG C414B-TL II aufgenommen, das Signal wurde mit gleichem Pegel auf L und LS gegeben. Es wurde wie in Versuch 2a verfahren.

Versuch 3b:

Gleiche Bedingungen wie in Versuch 3a, der monofon aufgenommene Flügel wurde wie in Versuch 2b auf Laufzeitdifferenzen untersucht.

Versuch 3c:

Bei diesem Versuch wurde der Flügel stereofon mit zwei AKG C414B-TL II in AB-Anordnung (Mikrofonbasis 60 cm) aufgezeichnet. Das linke Signal wurde auf LS gelegt, das rechte auf L. Wieder wurde in 1 dB-Schritten der hintere Lautsprecher abgesenkt.

Versuch 3d:

Gleiche Versuchsanordnung wie in 3c. Es wurden wieder Laufzeitdifferenzen in 0,1 ms-Abstufungen erzeugt.

Versuch 3e:

Der stereofone Flügel wurde wie in Versuch 3c auf L und LS verteilt. Diesmal wurde die Abbildung mit dem Surround-Panner der Software Pro Tools verschmälert. Das L-Signal wurde zu 25% auf LS gegeben und das LS entsprechend zu 25% auf L.

Versuch 4a:

Eine Akustikgitarre wurde monofon mit einem Neumann KM140 in einem Abstand von ca. 70 cm aufgenommen. Das Signal wurde auf L und LS geroutet, wieder wurden verschiedene Pegeldifferenzen erzeugt.

Versuch 4b:

Gleiche Bedingungen wie in 4a, es wurden wieder Laufzeitdifferenzen in 0,1 ms-Schritten erzeugt.

### Versuch 4c:

Eine Akustikgitarre wurde mit zwei KM 140 im XY-Verfahren stereofon aufgezeichnet. Ein Signal wurde auf L, das andere auf LS gegeben. Wieder wurde in 1 dB-Schritten der hintere Lautsprecher abgesenkt.

### Versuch 4d:

Gleiche Bedingungen wie in Versuch 4c, diesmal mit verschiedenen Laufzeitdifferenzen.

### Versuch 4e und f:

Gleiche Vorgehensweise wie in den Versuchen 4c und d. Die Gitarre wurde hierbei mit zwei Neumann KM 130 im AB-Verfahren (Mikrofonbasis 25 cm) aufgenommen.

### Auswertung Versuch 2a:

Bei gleichem Pegel zwischen L und LS, optimaler Hörposition und unbewegtem Kopf läßt sich die Stimme bei etwa  $70^\circ \pm 10^\circ$  lokalisieren. Wie nicht anders zu erwarten, kippt die Phantomschallquelle schon bei kleinsten Kopfbewegungen und -drehungen sofort in die entsprechende Richtung und zwar wesentlich stärker als bei Phantomschallquellen zwischen L und R. Auch wenn man den Kopf absolut still hält, ist die Abbildung nicht wirklich stabil. Je nach Stimmlage und gesungenen Tönen scheinen Stimmanteile aus verschiedenen Richtungen zu kommen. So wurden höhere Töne eher von vorne, tiefe eher von hinten lokalisiert. Ebenso fallen Zischlaute im hinteren Lautsprecher störend auf.

Auffällig ist, dass die Stimme näher wahrgenommen wird, als wenn sie nur aus einem der beiden Lautsprecher kommt. Anders als bei vorderen Phantomschallquellen wird das Hörereignis nicht zwischen den beiden Lautsprechern wahrgenommen, sondern verläßt die Stereobasis und wird innerhalb des Lautsprecherkreises lokalisiert.

Das Klangbild hat Ähnlichkeit mit einem Pseudostereofonie-Effekt, bei der die beiden Kanäle unterschiedlich gefiltert wurden und wirkt unangenehm beengend. Verringert man den Pegel des hinteren Lautsprechers, so wandert die Stimme in Richtung des vorderen. Mit zunehmender Differenz wird das Hörereignis stabiler und näher am Lautsprecherkreis wahrgenommen. Bei ca. 6 dB weniger Pegel auf dem hinteren Lautsprecher erscheint die Stimme bei ca. 45°, wobei Kopfbewegungen noch kein allzu starkes Springen des Hörereignisses verursachen. Ab ca. 10 dB Pegelunterschied wird die Stimme nur noch aus der Richtung des vorderen Lautsprechers wahrgenommen. Verringert man umgekehrt den Pegel des vorderen Lautsprechers, so wird die Stimme bei Differenzen zwischen 7 und 9 dB bei ca. 90°, also in Höhe der Ohrachse lokalisiert.

Auswertung Versuch 2b:

Bei Laufzeitdifferenzen ergibt sich ein ähnliches Bild. Durch Verzögern des hinteren Lautsprechers wandert die Stimme nach vorne. Bei 0,5 ms Differenz wird sie bei etwa 45°, ab ca. 1 ms nur noch aus Richtung des vorderen Lautsprechers lokalisiert. Die Abbildung wirkt stabiler als mit Pegelunterschieden, allerdings auch etwas unschärfer. Hinzukommt, dass sich je nach Verzögerungszeit Kammfiltereffekte störend bemerkbar machen. Außerdem wird hauptsächlich bei Lauten mit hochfrequenten Signalanteilen (Zischlaute wie s, sch, f, u.ä.) der hintere Lautsprecher störend wahrgenommen. Mit steigender Verzögerung macht sich dieser Effekt zunehmend bemerkbar. Man kann hier weniger von Echo sprechen, vielmehr scheint das Klangbild in zwei einzelne Hörereignisse zu zerfallen. Der umgekehrte Fall, bei dem der vordere Lautsprecher verzögert wird, zeigt bei entsprechenden Laufzeitdifferenzen ähnliche, entgegengesetzte Resultate.

Auswertung Versuch 3a und b:

Der Versuch mit dem monofonen Flügel liefert ähnliche Ergebnisse wie bei den Versuchen 2a und b. Der Flügel bekommt eine räumliche Ausdehnung, allerdings wirkt die Abbildung noch instabiler als bei der Stimme.

Die Lokalisation ist stärker frequenzabhängig. Je nachdem in welcher Lage gespielt wird, springt das Hörereignis. Auch der unangenehm beengende Höreindruck und der Pseudostereofonie-Effekt fallen hier stärker auf. Je größer der Pegelunterschied oder die Verzögerung, desto mehr wandert das Hörereignis zum lauterem bzw. früheren Lautsprecher und desto klarer wird es auf der Basis zwischen L und LS lokalisiert.

Auswertung Versuch 3c:

Anders als man aus den bisherigen Versuchen erwarten könnte, entsteht bei der stereofonen Aufnahme ein vollkommen anderer Höreindruck. Der Flügel ist nun von der Entfernung richtig auf dem Lautsprecherkreis plaziert, die subjektiv unangenehm empfundene Nähe ist nicht mehr vorhanden. Die Abbildung ist jedoch sehr breit und das Klangbild zerfällt etwas. Reduziert man den Pegel des hinteren Lautsprechers so läßt sich die Breite des Flügels verringern, dabei verschiebt sich natürlich das Hörereignis etwas in Richtung des vorderen Lautsprechers. Bei Pegeldifferenzen zwischen 4 und 8 dB wird der Flügel zwischen 55° und 30° ziemlich stabil und mit einer natürlichen Ausdehnung lokalisiert.

Auswertung Versuch 3d:

Der Versuch mit Laufzeitdifferenzen liefert ähnliche Resultate wie Versuch 3c. Auch hier läßt sich die räumliche Ausdehnung des Flügels durch Verzögern des hinteren Lautsprechers eingrenzen. Das Hörereignis verschiebt sich wieder in Richtung des vorderen Lautsprechers. Je nach Laufzeitdifferenz entstehen z.T. störend wirkende Frequenzauslöschungen, so dass hier probiert werden muß, welche Verzögerung unproblematisch ist.

Auswertung Versuch 3e:

Zwar läßt sich auch in diesem Versuch die Ausdehnung des Flügels verringern, allerdings fällt hierbei auf, dass der beengende Eindruck wieder auftritt. Ebenso geht die Räumlichkeit der Abbildung verloren. Die Lokalisation wird wieder instabiler.

Auswertung Versuch 4a und b:

Die Versuche mit der monofonen Gitarre liefern weitgehend die gleichen Ergebnisse wie bei der Stimme und dem monofon aufgenommenen Flügel. Die Gitarre scheint sich dabei etwas besser lokalisieren zu lassen als der Flügel, das Klangbild zerfällt weniger stark.

Auswertung Versuch 4c und d:

Die Versuche mit dem XY-Verfahren liefern trotz Experimenten mit verschiedenen Öffnungswinkeln keine so überzeugenden Resultate wie die AB-Stereofonie. Der Höreindruck ist weniger eindrucksvoll, wieder etwas beengend und unterscheidet sich nicht so stark von dem der monofonen Gitarre. Mit Pegel- und Laufzeitdifferenzen lassen sich aber auch hier brauchbare Abbildungen erzielen.

Auswertung Versuch 4e und f:

Wie auch schon bei der stereofonen Flügelaufnahme kann hier mit entsprechenden Pegel- oder Laufzeitdifferenzen eine stabile Abbildung zwischen  $55^\circ$  und  $30^\circ$  erreicht werden. Zusätzlich wurden unterschiedliche Mikrofonbasen ausprobiert. Bei zu kleiner Basis tendiert der Höreindruck in Richtung der monofonen Aufnahme, bei zu großer ist die Abbildung zu breit und das Klangbild zerfällt ebenso.

Zusammenfassung Versuch 2 - 4:

Monosignale lassen sich nicht zufriedenstellend auf der Seite abbilden. Die Abbildung ist extrem instabil, springt schon bei sehr geringen Bewegungen des Hörers in den jeweils näheren Lautsprecher und zerfällt z.T. in einzelne Hörereignisse bei den Lautsprechern. Oftmals erscheinen die Signale räumlich ausgedehnt und werden innerhalb des Lautsprecherkreises lokalisiert. Zusätzlich entsteht in den meisten Fällen ein unangenehm, beengender Eindruck.

Dass ein Signal gleichzeitig und mit gleichem Pegel sowohl aus einer vorderen als auch hinteren Einfallsrichtung auf unser Ohr trifft, kommt in der Natur nicht vor und stellt für unser Gehör einen unbekanntem Reiz dar und wird deshalb vermutlich als unangenehm empfunden. Die Pseudostereofonie-Wirkung kommt wahrscheinlich dadurch zustande, dass das von vorne eintreffende Lautsprecher-Signal durch die Ohrmuschel spektral anders gefiltert wird als das von hinten eintreffende.

Durch Pegel- und Laufzeitunterschiede läßt sich ein Signal in Richtung des vorderen oder hinteren Lautsprechers verschieben. Die Abbildung wird zunehmend stabiler, das Hörereignis wird näher am Lautsprecherkreis wahrgenommen und der negative Klangeindruck wird reduziert. Mit etwa 6 bis 8 dB Pegeldifferenz oder 0,4 bis 0,8 ms Verzögerung läßt sich ein Monosignal bei etwa 45° bzw. 90° relativ stabil seitlich plazieren. Bei Laufzeitdifferenz fällt jedoch oft der spätere Lautsprecher vor allem bei hochfrequenten Signalanteilen störend auf. Hier wird man für brauchbare Resultate wohl zusätzlich mit einer Höhenabsenkung arbeiten müssen. Desweiteren fällt auf, dass bei Intensitätspannung die vom Display des Surround-Panpots angezeigte Schalleinfallsrichtung oftmals nicht mit der wirklichen Hörereignisrichtung übereinstimmt.

Stereosignale, besonders laufzeitbasierende, lassen sich ziemlich gut an der Seite plazieren. Die Abbildung wirkt jedoch für die meisten Anwendungen zu breit. Versucht man die Breite mit dem Surround-Panpot zu verringern, hat dies sofort wieder den unangenehm, beengenden Klangeindruck zur Folge, weil dabei wieder identische Signalanteile von vorne und hinten eintreffen.

Durch Pegelabsenkung und z.T. auch durch Verzögerung eines der beiden Lautsprecher-Signale kann die Abbildungsbreite ohne Klangeinbußen verringert werden. Dabei verschiebt sich jedoch das Hörereignis in Richtung des lautereren bzw. früheren Lautsprechers.

Pegelstereofone Signale lassen sich weniger gut auf der Seite abbilden. Offenbar spielt der Kohärenzgrad zwischen L- und LS-Signal hierbei eine große Rolle. Bei Pegeldifferenz sind die Mikrofon-Signale weniger dekorreliert als bei Laufzeitstereofonie, haben also einen hohen Kohärenzgrad.

Dies ist anscheinend für die seitliche Abbildung nachteilig. Ähnliches läßt sich auch beim AB-Verfahren mit verschiedenen Mikrofonbasen beobachten. Ist die Basis zu klein, so sind die Signale einander zu ähnlich, zu kohärent. Bei Vergrößerung der Mikrofonbasis sinkt der Kohärenzgrad. Ist die Basis zu groß, der Kohärenzgrad zu klein, zerfällt das Klangbild in einzelne Hörereignisse.<sup>40</sup>

Interessant wäre noch, wie sich äquivalenzstereofone Aufnahmen (ORTF, NOS) bei seitlicher Abbildung verhalten. Vermutlich ähnlich wie die aufgenommenen AB-Stereofonie-Beispiele, bei denen ja auch Pegeldifferenzen vorhanden sind. Bei der Flügelaufnahme waren die Mikrofone auf Nierencharakteristik eingestellt, und bei der Gitarrenaufnahme entstehen aufgrund des relativ geringen Abstands der Druckempfänger zum Instrument frequenzabhängige Pegeldifferenzen.

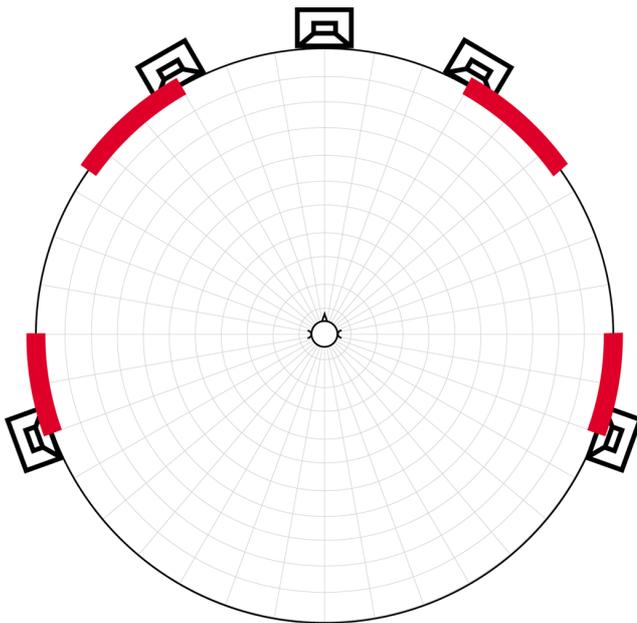


Abb. 17; Bereiche, in denen eine zufriedenstellende seitliche Abbildung möglich ist.

---

<sup>40</sup> Vgl.: Theile, Günther, *Multichannel Natural Music Recording Based on Psychoacoustic Principles*, [http://www.irt.de/wittek/hauptmikrofon/theile/Multich\\_Recording.pdf](http://www.irt.de/wittek/hauptmikrofon/theile/Multich_Recording.pdf), S. 32 f.  
Seitenaufwurf vom 21.03.2005

### 6.3 Das Surroundfeld

Nicht nur zwischen vorderen und seitlichen Lautsprechern können Phantom-schallquellen gebildet werden, sondern auch zwischen vorderen und hinteren Lautsprechern. Es gilt zu untersuchen, wie sich ein monofones Signal verhält, wenn es auf mehrere Boxen gegeben wird, also innerhalb des Lautsprecherkreises plaziert wird und wo es tatsächlich lokalisiert wird. Desweiteren ist von Interesse, wie sich eine möglichst stabile Abbildung erreichen läßt und ob die *richtungsbestimmenden Bänder* eventuell dazu beitragen können.

Versuch 5a:

Eine Sängerin wurde mittels Surround-Panpot genau in der Mitte des Surroundfeldes plaziert. Das Signal wird dabei auf den Center und die hinteren beiden Lautsprecher geroutet, wobei LS und RS je 3 dB weniger Pegel erhalten. Wie bei der Phantommitte zwischen L und R, wird auch hier eine 3 dB-Mittendämpfung erzeugt, um die Pegelzunahme bei der Summierung auszugleichen. In 1 dB-Schritten wurde der Pegel der hinteren Lautsprecher abgesenkt.

Versuch 5b:

Wieder wurde die Stimme auf C, LS und RS gegeben, diesmal wurde das Signal in den hinteren Lautsprechern in den Frequenzbereichen um 1,2 kHz und 8 kHz sowie oberhalb von 12 kHz relativ breitbandig (Q-Faktor 2) abgesenkt.

Versuch 5c:

Bei diesem Versuch wurde die Stimme ebenfalls auf C, LS und RS geroutet, die hinteren Lautsprecher wurden dabei in 1 ms-Abstufungen verzögert.

Versuch 5d:

Gleiche Vorgehensweise wie in Versuch 5c, zusätzlich wurden die hinteren Lautsprecher noch bei 1,2 kHz, 8 kHz und oberhalb von 12 kHz abgesenkt.

### Versuch 6a:

Eine Snaredrum wurde genau in der Mitte des Surroundfeldes plaziert. Dabei wurde das Signal auf alle fünf Lautsprecher gegeben. Der Pegel von L und R wurde in 1 dB-Schritten abgesenkt, der Pegel der hinteren Lautsprecher um jeweils 2 dB.

### Versuch 6b:

Gleiche Voraussetzungen wie in 6a, diesmal wurden L und R in 1 ms-Schritten verzögert, die hinteren Lautsprecher um jeweils 2 ms.

### Auswertung Versuch 5a:

Entgegen den Vermutungen wird die Stimme bei Platzierung in der Mitte des Lautsprecherkreises nicht im oder über dem Kopf wahrgenommen. Zwar scheint die Stimme nahe am Kopf zu sein, sie ist aber nicht mehr klar zu lokalisieren, wirkt räumlich ausgedehnt und wird vor allem bei Zischlauten aus den verschiedenen Lautsprechern wahrgenommen. Auch ist wieder ein unangenehmer, beengender Klangeindruck festzustellen. Desweiteren ist die Lokalisation extrem von der Hörposition abhängig. Befindet man sich etwas näher an den hinteren Lautsprechern, so kippt die Hörereignisrichtung nach hinten. Auch bei seitlichen Bewegungen wird die Stimme sofort im linken bzw. rechten hinteren Lautsprecher lokalisiert.

Senkt man den Pegel der hinteren Lautsprecher etwas ab, so verbessert sich die Lokalisation, sofern man sich nicht allzu weit aus der optimalen Hörposition bewegt. Die Stimme erscheint näher am Hörer und wirkt zugleich räumlich ausgedehnt. Der beengende Klangeindruck ist zwar immer noch vorhanden, allerdings nicht mehr so stark ausgeprägt. Bei weiterer Pegelreduzierung (ab ca. 8 dB Differenz) geht der räumliche Höreindruck deutlich zurück und die Stimme scheint nur noch aus dem Center zu kommen.

### Auswertung Versuch 5b:

Durch die Absenkung der Blauertschen *diffus*-Bänder wird ebenfalls eine Verbesserung der Lokalisation erzielt.<sup>41</sup> Bei Werten zwischen 4 und 8 dB läßt sich eine akzeptable Stabilität bei bleibendem räumlich Höreindruck erreichen. Höhenreiche Zischlaute werden zwar nicht mehr von hinten wahrgenommen, der unangenehm beengende Klang bleibt allerdings bestehen.

### Auswertung Versuch 5c:

Schon bei einer Verzögerungszeit von 2 ms ist eine klare *vorne*-Lokalisation der Stimme wahrzunehmen. Allerdings ist vor allem bei Zischlauten wieder eine Art Echo aus den hinteren Lautsprechern festzustellen. Bei größeren Laufzeitdifferenzen wird die Stimme noch stabiler vorne lokalisiert, die Echos fallen aber zunehmend stärker auf. Zwar tritt hier der beengende Klangeindruck weniger in Erscheinung, dafür werden je nach Verzögerungszeit Kammfiltereffekte störend wahrgenommen.

### Auswertung Versuch 5d:

Durch die zusätzliche Filterung der hinteren Lautsprecher ergibt sich eine sehr stabile Richtungslokalisierung ohne störende Echos. Bei Verzögerungswerten bis zu 10 ms lassen sich gute räumliche Höreindrücke erzielen. Selbst wenn man sich deutlich aus der optimalen Hörposition entfernt, bleibt die Vorzugsrichtung *vorne* bestehen. Dennoch ist auch hier der Klangeindruck nie völlig zufriedenstellend.

---

<sup>41</sup> Anmerkung: Innerhalb dieser Versuchsreihe wurde auch mit einer generellen Höhenabsenkung bei einer Grenzfrequenz von 3,5 kHz experimentiert. Hierbei ließen sich im Vergleich mit den *richtungsbestimmenden Bändern* kaum Unterschiede wahrnehmen.

### Auswertung Versuch 6a:

Die Ergebnisse dieses Versuchs entsprechen weitgehend denen aus Versuch 5a. Bei Platzierung der Snare genau in der Mitte ist die Lokalisation extrem instabil, auch hier werden die hochfrequenten Signalanteile (Snareteppich) störend aus einzelnen Lautsprechern wahrgenommen. Die Snare wirkt durch die hinzugekommenen Lautsprecher L und R noch ausgedehnter und auch etwas stabiler, als wenn sie vorne nur aus dem Center kommt. Auffallend ist, dass im Gegensatz zur Stimme der unangenehm beengende Klangeindruck kaum wahrzunehmen ist.

### Auswertung Versuch 6b:

Durch die Laufzeitdifferenzen auch zwischen Center und L-R kommt es zu einer starken räumlichen Aufweitung des Höreindrucks. Ganz anders als bei den Versuchen mit Pegelunterschieden entsteht hier ein wirklicher Raumeindruck. Die Snare ist immer vorne zu lokalisieren, allerdings stören auch hier wieder hochfrequente Signalanteile aus den hinteren Lautsprechern. Diese lassen sich aber nach den bisherigen Erkenntnissen durch entsprechende Filterung beseitigen. Je nach Verzögerungszeit treten auch wieder negative Kammfiltereffekte auf, der beengende Klangeindruck ist jedoch nicht festzustellen.

### Zusammenfassung Versuch 5 & 6:

Je nach Schallereignis und Panningmethode ergeben sich sehr unterschiedliche Höreindrücke. Die konkrete Abbildung einer Schallquelle im Bereich des sweet spot kann nicht erreicht werden. Die Phantomschallquelle läßt sich nicht klar lokalisieren, wirkt räumlich ausgedehnt, ist zugleich aber instabil, und vor allem hochfrequente Signalanteile werden störend aus den hinteren Lautsprechern wahrgenommen. Je weiter man das Signal in Richtung der vorderen oder hinteren Lautsprecherfront verschiebt, desto stabiler und präziser wird die Abbildung.

Bei Verzögerung der hinteren Lautsprecher kommt der Präzedenzeffekt zum Tragen. Das Hörereignis wird selbst bei ungünstigeren Hörpositionen immer vorne lokalisiert. Erst wenn man sich sehr nahe an den hinteren Lautsprechern befindet, kippt die Hörereignisrichtung. Die bei den Laufzeitdifferenzen auftretenden Echos, die störende Wahrnehmung der hinteren Lautsprecher, läßt sich durch eine Höhenabsenkung beseitigen. Allerdings entstehen durch die unterschiedlichen Phasenlagen Frequenzauslöschungen und -anhebungen, die den Klang negativ beeinflussen können.

Wie auch schon bei der seitlichen Abbildung monofoner Signale aufgefallen ist, entsteht ein unangenehmer, beengender Klangeindruck, wenn von vorne und hinten identische Signale an den Ohren eintreffen. Interessanterweise ist dies bei den Versuchen mit der Snare vor allem bei Laufzeitdifferenzen fast gar nicht festzustellen. Vermutlich liegt dies an der sich rasch ändernden zeitlichen Struktur perkussiver Signale. Dadurch treten die Schallenergien nicht mehr gleichzeitig am Hörplatz auf.

Während die Stimme bei Wiedergabe über vordere und hintere Lautsprecher immer etwas befremdlich klingt, lassen sich bei perkussiven Signalen interessante Sounds kreieren. Besonders Laufzeitdifferenzen erzeugen eindrucksvolle, sehr räumlich klingende Hörereignisse, bei gleichzeitig stabiler Lokalisation. Vor allem in Kombination mit Höhenabsenkungen und Pegeldifferenzen lassen sich überzeugende Resultate erzielen.

Bei Pegeldifferenzen entsteht ein breite, ausgedehnte Abbildung der Snare, die innerhalb des Lautsprecherkreises lokalisiert wird. Mit Laufzeitunterschieden läßt sich ein gänzlich anderer Eindruck erzeugen. Die Snare wird klar vorne wahrgenommen, um sie herum entsteht aber ein sehr natürlich klingender Raumeindruck. Die Verzögerungen wirken dabei wie Schallrückwürfe von reflektierenden Flächen. Man hat tatsächlich den Eindruck, dass man sich mit in dem Raum befindet, in dem die Snare gespielt wird. Ein solches Hörerlebnis läßt sich in der Zweikanal-Stereofonie nicht erzeugen.

## 6.4 Umhüllung

Umhüllung ist die Wahrnehmung eines nicht lokalisierbaren Hörereignisses, dass den Hörer vollständig umgibt. In diesen Versuchen soll herausgefunden werden mit welchen Methoden sich eine bestmögliche Umhüllung erzeugen läßt, so dass der Hörer das Gefühl hat, das Instrument erklingt rund um ihn herum und er befindet sich in der Mitte davon.

Versuch 7a:

Ein monofones Streicher-Pad wurde zeitgleich und mit gleichen Pegeln auf alle fünf Lautsprecher geroutet.

Versuch 7b:

Das Streicher-Pad wurde wieder mit gleichen Pegeln auf alle fünf Lautsprecher gegeben, zusätzlich wurden die Kanäle L und R um jeweils 3 ms, die hinteren Lautsprecher um je 6 ms verzögert.

Versuch 7c:

Bei diesem Versuch wurde das Streicher-Pad im Pro Tools fünfmal kopiert. Bei jeder Kopie wurde mit einem Pitch Shifter die Tonhöhe jeweils unterschiedlich minimal um 3 bis 6 Cent verändert. Jede Kopie wurde auf einen anderen Lautsprecher gegeben.

Versuch 8a:

Eine E-Gitarre mit verzerrten Gitarrensound wurde monofon aufgenommen. Das Signal wurde wieder zeitgleich und mit gleichen Pegeln auf alle fünf Lautsprecher geroutet.

### Versuch 8b:

Gleiches Vorgehen wie in Versuch 7c. Verzerrte E-Gitarre fünfmal unterschiedlich gepitcht.

### Versuch 8c:

Die E-Gitarre wurde fünfmal möglichst ähnlich eingespielt. Jede Spur wurde auf einen anderen Lautsprecher geroutet.

### Versuch 8d:

Diesmal wurde die E-Gitarre mit leichten Spielvariationen und jeweils leicht veränderter Klangeinstellung am Verstärker fünfmal eingespielt. Wieder wurde jede Spur wurde auf einen anderen Lautsprecher gegeben.

### Auswertung Versuch 7a:

Wie auch schon in Versuch 6a mit der Snare, ist das Hörereignis nicht klar zu lokalisieren, wird z.T. in Kopfnähe wahrgenommen und kippt bei leichten Bewegungen in die jeweilige Richtung. Einzelne Lautsprecher fallen weniger auf als beim Snare-Versuch, was mit Sicherheit an dem stationären Klang des Streicher-Pads liegt. Ein überzeugendes Umhüllungsgefühl stellt sich aber nicht ein. Auch der unangenehme, beengende Eindruck ist festzustellen.

### Auswertung Versuch 7b:

Die Laufzeitunterschiede bewirken wie in Versuch 6b eine deutliche räumliche Aufweitung. Das Klanggeschehen rückt vom Hörer ab, drängt auf allen Seiten nach außen, ist nicht mehr beengend, sondern sehr offen. Es entsteht der Eindruck vom Streicherklang umhüllt zu sein.

### Auswertung Versuch 7c:

Bei diesem Versuch ist der Umhüllungseindruck noch stärker. Die Streicher bilden rundum einen sehr homogenen Klangteppich, der sich leicht um den Hörer zu bewegen scheint. Phasing-Effekte sind deutlich wahrnehmbar.

### Auswertung Versuch 8a:

Die Ergebnisse entsprechen denen aus Versuch 7a. Es wird kein Umhüllungseindruck erzielt.

### Auswertung Versuch 8b:

Auch hier stellt sich wie in Versuch 7c ein sehr gutes Umhüllunggefühl ein. Allerdings klingt der Gitarrensound unnatürlich, künstlich, synthetisch. Ein leichtes Wandern des Hörereindrucks ist wahrzunehmen, ebenso das Phasing.

### Auswertung Versuch 8c:

Das Umhüllungseindruck ist in diesem Versuch noch eindrucksvoller, eine regelrechte *Wall Of Sound* entsteht um den Hörer herum. Der Höreindruck ist sehr homogen ohne klangliche Einbußen und Phasing-Effekte.

### Auswertung Versuch 8d:

Auch diesmal ist der Umhüllungseindruck sehr gut, der Klang ist angenehm ohne Phasing-Effekte. Allerdings ist der Gesamteindruck nicht so homogen wie in Versuch 8c, ab und zu werden einzelne Richtungen wahrgenommen, was an den leichten Spielnuancen und den etwas unterschiedlichen Klangeinstellungen am Verstärker liegt. Dies fällt aber nicht negativ auf, sondern läßt das Klangbild lebendig erscheinen.

Zusammenfassung Versuch 7 & 8:

Ähnlich wie sich in der Zweikanal-Stereofonie durch doppelten der Signale, eine extrem breite, über die gesamte Stereobasis ausgedehnte Fläche erzeugen läßt, so kann in der Fünfkanaal-Stereofonie durch die gleichen Methoden eine eindrucksvolle Umhüllung geschaffen werden. Die durch Verzögerung oder Pitch Shifting erzeugten Pseudosurround-Effekte sind dabei nicht ganz so spektakulär und klanglich überzeugend, wie wenn das Signal fünfmal eingespielt wurde. Dennoch läßt sich auch damit ohne große Probleme eine überzeugende Umhüllung erzielen.

### 7. Surround-Produktion

Die aus den Misch-Experimenten gewonnenen Erkenntnisse sollten auch in einer Surround-Produktion umgesetzt werden. Das Ziel hierbei war, ein räumlich eindrucksvolles, umhüllendes Klangbild zu schaffen, bei dem die Abbildung stabil bleibt und die Instrumente auch an weniger optimalen Hörpositionen noch aus der gewünschten Richtung lokalisiert werden. Dabei sollten die gestalterischen Möglichkeiten der Fünfkanal-Stereofonie ausgeschöpft werden und dem Zuhörer ein neuartiges, gegenüber der Zweikanal-Stereofonie deutlich gesteigertes Hörerlebnis bescheren.

Für die Surround-Produktion konnte die Stuttgarter Band *Schnute* gewonnen werden. Die Besetzung der Gruppe ist Andi Zbik (Schlagzeug, E-Drums), Michael Deak (Kontrabaß, E-Bass), Fabian Wendt (Stick, E-Bass), Carsten Netz (Klarinette, Saxofon, Querflöte) und Axel Hemprich (Keyboards, Sampler, Synthesizer). Die Musik wird häufig als NuJazz bezeichnet und setzt sich aus Einflüssen aus den Bereichen Jazz, Drum&Bass, Ambient und TripHop zusammen. Sie bietet sich außerdem hervorragend für Surround-Experimente an.

#### 7.1 Aufnahme

Die Aufnahme einer Musikgruppe – egal ob es sich dabei um ein klassisches Ensemble, eine Jazzformation oder eine Rockband handelt – stellt eine sehr anspruchsvolle, oftmals unterschätzte Aufgabe dar. Neben der technischen und klanggestalterischen Seite, ist vor allem ein entspannter Umgang mit den Künstlern und eine harmonische Arbeitsatmosphäre von elementarer Bedeutung. Das Ziel einer jeden Produktion sollte sein, nicht nur eine spieltechnisch perfekte, sondern auch emotional packende Aufnahme zu erreichen. Es ist letztendlich die Musik, die den Hörer fesselt und nicht technische Perfektion.

Die Aufnahmen erfolgten im Tonstudio des Fachbereichs Electronic Media der Hochschule der Medien. Auf ausdrücklichen Wunsch der Band wurde *live* eingespielt, d.h. die Band spielte den Song gemeinsam, alle Instrumente wurden – was im Popbereich mittlerweile eher unüblich ist – gleichzeitig aufgenommen. Einzig ein paar elektronische Sounds aus dem Sampler wurden nachträglich als Overdubs hinzugefügt. Das hatte jedoch ausschließlich technische Gründe. Es standen nur drei achtkanalige Interfaces, somit nur 24 Eingänge zur Verfügung und diese waren komplett belegt. Aus diesem Grund wurden die Sampler-Signale bei der Live-Einspielung nur mono aufgezeichnet und hinterher – sofern überhaupt zweikanalig vorhanden – durch die Stereoversion ersetzt.

Um ein Übersprechen der Instrumente zu vermeiden und so möglichst saubere Signale für die Nachbearbeitung zu erhalten, wurden die Musiker auf verschiedene Räume verteilt. Das Schlagzeug wurde im großen Aufnahmerraum aufgestellt. Ebenso konnten die elektronischen Klangerzeuger (Keyboard, Sampler, Synthesizer) und der Chapman-Stick<sup>42</sup> bedenkenlos dort aufgebaut werden, da sie keinen natürlichen Schall erzeugen. Der Kontrabaß samt Verstärker wurde in der B-Regie, die Klarinette im kleinen Aufnahmerraum (Sprecherkabine) plaziert. Dank der großen Sichtfenster zwischen den Räumen war auch der Blickkontakt zwischen den Musikern gegeben. Es wurden vier getrennte Kopfhörer-Mischungen erstellt. Dabei erhielten Schlagzeuger, Bassist und Klarinetttist je einen eigenen Mix, die anderen beiden mußten sich eine Kopfhörer-Mischung teilen, was aber keine Probleme bereitete.

Es wurden mehrere Song-Durchläufe, sogenannte Takes aufgenommen. Die endgültige Fassung wurde in Abstimmung mit den Musikern aus zwei verschiedenen Takes zusammen geschnitten. Aufnahme und Schnitt erfolgten mit der Software Pro Tools der Firma Digidesign.

---

<sup>42</sup> Beim sogenannten *Chapman-Stick* handelt es sich um ein acht- bis zwölfsaitiges Instrument (benannt nach seinem Erfinder Chapman), bei dem die Klangerzeugung auf die gleiche Weise wie bei einer E-Gitarre erfolgt. Das Instrument besitzt keinen Resonanzkorpus (daher die Bezeichnung Stick) und wird mit der sog. Tapping-Technik gespielt. Dabei werden die Saiten allein durch Fingerdruck zum Schwingen gebracht, d.h. die Saiten werden nicht wie bei der Gitarre gezupft.

### *Mikrofonauswahl und Positionierung*

Am Schlagzeug kamen insgesamt neun Mikrofone zum Einsatz. Die Bassdrum wurde mit drei Mikrofonen abgenommen, ein Sennheiser MD 421 wurde im Kessel relativ mittig in ca. 10 cm Abstand auf das Schlagfell gerichtet, um viel Anschlagsgeräusch zu bekommen. Hinzu kam ein Beyerdynamic KM 88, das etwas mehr seitlich platziert wurde, um etwas mehr „Ton“ einzufangen und zuletzt wurde noch ein Electrovoice RE-20 außen am Resonanzfell platziert, was die Natürlichkeit der Bassdrum verbessert und den Gesamtklang abrundet.

Die Snaredrum wurde ebenfalls dreifach mikrofoniert. Ein Shure SM 57 wurde in einem Winkel von ca. 35° schräg von oben etwa auf die Mitte zwischen Mittelpunkt und Rand der Snare gerichtet. Daneben wurde ein Sennheiser MD 421 mit etwas flacherem Winkel platziert und ein weiteres SM 57 wurde von unten auf das Resonanzfell gerichtet, um den Klang des Snareteppichs zu verstärken.

An der Hi-Hat wurde ein AKG C 414 B-TL II verwendet. Die Richtcharakteristik wurde dabei auf Hypernieren gestellt, da sich so das Übersprechen der Snare am besten minimieren ließ. Als Overhead-Mikrofone kamen zwei AKG C 451 mit Nierenkapsel (CK1) zum Einsatz, die jedoch weniger als Stereomikrofon (AB) platziert, sondern relativ gezielt auf die beiden Becken ausgerichtet wurden.

Der Grund dafür bestand darin, dass der Gesamtsound des Schlagzeugs durch zwei zusätzlich aufgestellte Ambience-Mikrofonpaare eingefangen wurde. Ein Stereopärchen wurde dabei in Groß-AB (Abstand ca. 2 m) in etwa 3,5 m Höhe und ähnlicher Entfernung zum Schlagzeug angebracht. Hierfür wurden zwei Neumann TLM 170 verwendet.

Bei früheren Schlagzeugaufnahmen im Tonstudio der Hochschule hatte ich gute Erfahrungen mit einem im Vorraum aufgestellten Stereomikrofon gemacht, so dass auch bei diesen Aufnahmen ein vor dem Aufnahmeraum platziertes Stereomikrofon in AB-Technik (Abstand ca. 60 cm) mit zwei Neumann KM 130 zum Einsatz kam. Die Verbindungstüren zum Aufnahmeraum blieben verständlicherweise offen.

Relativ ungewöhnlich ist, dass der Schlagzeuger keine Toms spielt. Stattdessen benutzt er mehrere Pads, mit denen er ein Drum-Modul antriggern und so zusätzlich elektronische Klänge erzeugen kann. Diese wurden über DI-Boxen aufs Mischpult geroutet. Ebenso wurde mit dem Stick und den elektronischen Klangerzeugern verfahren.

Der Kontrabaß wurde mit einem Brauner Großmembran-Röhrenmikrofon vom Typ Valvet abgenommen. Das Mikrofon wurde in etwa 30 cm Abstand oberhalb des rechten F-Lochs seitlich des Halses plaziert. Am Steg des Basses war außerdem ein Tonabnehmer angebracht, so dass ein zweites Signal am Direct-Out des Bassverstärkers abgegriffen werden konnte. Der Lautsprecher des Verstärkers war nicht in Betrieb.

Der Klarinettist der Band spielt sein Instrument nicht auf traditionelle Art und Weise. Vielmehr schickt er das Instrumentensignal über ein Mikrofon in ein kleines Mischpult und fügt meist verschiedene Effekte hinzu. Es kommen u.a. Delay, Phaser, Equalizer und ein Ringmodulator zum Einsatz. Für die Effekte wurde ein Neumann TLM 170 R verwendet. Zusätzlich wurde das Instrument noch mit einem Brauner Valvet abgenommen, so dass bei der Mischung auch der natürliche Klang der Klarinette zur Verfügung stand.

Die Auswahl und Positionierung der Mikrofone erfolgte nach subjektiven, ästhetischen Kriterien und wurde durch persönliche Erfahrungswerte und z.T. durch Ausprobieren ermittelt.

### 7.2 Mischung und Mastering

Die künstlerische Gestaltung der Mischung – wie auch schon die Aufnahme – unterliegt immer den subjektiven, klangästhetischen Vorstellungen des jeweiligen Tonmeisters. Bei dieser Produktion sollten zwei in der technischen Herangehensweise unterschiedliche, bezüglich der Abbildung jedoch prinzipiell ähnliche Surround-Mischungen erstellt werden. Der Mix A erfolgte unter Berücksichtigung der aus den Misch-Experimenten gewonnenen Erkenntnisse mit Pegel- und Laufzeitunterschieden als auch Filterungen. Der Mix B wurde nur unter Verwendung des Surround-Panpots von Pro Tools erstellt, bei dem die Richtungsabbildung allein durch Pegeldifferenzen erzeugt wird.

Es ist klar, dass der Mix A aufgrund der zusätzlichen Laufzeitunterschiede und Filterungen einen wesentlichen größeren Aufwand erfordert. Zuerst einmal mußte in Pro Tools eine entsprechende Session angelegt werden, mit der eine solche Mischung überhaupt möglich ist. Dazu wurden in (fast) jedem Kanalzug fünf AUX-Wege angelegt, die jeweils auf einen eigenen Bus geroutet wurden, welche wiederum jeweils auf einen der fünf Lautsprecher gingen. Nur so können für jedes Instrument getrennt, unterschiedliche Laufzeiten und Frequenzänderungen zwischen den fünf Kanälen eingestellt werden. Aufgrund begrenzter Systemressourcen – bei 24 Spuren sind in diesem Fall allein für das Panning 120 AUX-Wege notwendig – ließ sich dies nicht vollständig umsetzen, so dass Kompromisse gemacht werden mußten (z.B. mehrere Instrumente auf einem gemeinsamen Bus mit derselben Verzögerungszeit).

Das Schlagzeug sollte vorne lokalisiert, dabei gut räumlich abgebildet werden, so dass der Hörer den Eindruck bekommt, er befinde sich in einem Abstand von ein paar Metern zum Schlagzeug im selben Raum. Dazu wurden die einzelnen Spuren auf vordere und hintere Lautsprecher geroutet, wobei die Surroundkanäle im Pegel meist um 4 - 6 dB reduziert wurden. Die Kanäle L und R wurden gegenüber dem Center um 2 - 4 ms, LS und RS um 4 - 8 ms verzögert. Bassdrum und Snare wurden mittig zwischen linker und rechter Seite positioniert. Die verschiedenen Mikrofonsignale wurden z.T. unterschiedlichen Lautsprechern zugeordnet.

So wurde z.B. das Sennheiser MD 421 der Bassdrum nur auf den Center, das Beyerdynamic KM 88 auf L-R und das RE-20 auf alle drei vorderen und zusätzlich noch dezent auf die hinteren Lautsprecher geroutet. Bassdrum und Snare wurden etwas komprimiert, Noisegates wurden keine verwendet, auch softwareseitig wurden die Spuren nicht „gesäubert“.

Die Hi-Hat wurde leicht nach links und die Overheads ca. 60% aus der Mitte nach links und rechts gepannt. Die seitlichen Auslenkungen wurden nur über Pegelunterschiede erzeugt.

Die Ambience-Spuren wurden zeitlich nach vorne verschoben, da beim Hinzumischen – aufgrund der durch die Entfernung zum Schlagzeug auftretenden natürlichen Laufzeitdifferenzen – Echos wahrzunehmen waren. Anschließend wurde der jeweils linke Kanal der Stereospur auf L und LS, der jeweils rechte Kanal entsprechend auf R und RS gegeben. Die hinteren Lautsprecher wurden verzögert, zusätzlich wurden die Höhen abgesenkt. Auch bei der Snare und den Overheadspuren mußten die hinteren Lautsprecher gefiltert werden.

Der Kontrabaß wurde auch vorne plaziert und um ca. 5-10% aus der Mitte nach links verschoben. Das DI-Signal, welches vor allem Bässe und untere Mitten lieferte, wurde mittels Pegeldifferenzen leicht aus der vorderen Lautsprecherfront geholt. Das Mikrofonsignal wurde auf alle Lautsprecher gegeben und über Laufzeit- und Pegelunterschiede ähnlich dem DI-Signal abgebildet. Durch die Verwendung von Laufzeitdifferenzen wird der Kontrabaß sehr natürlich räumlich wahrgenommen, so als wäre er vor dem Schlagzeug plaziert und im selben Raum aufgenommen worden. Beide Spuren wurden etwas komprimiert.

Der Stick wurde zwischen L und LS bei ca. 50° abgebildet. Es kamen Laufzeit- und Pegelunterschiede sowie eine Höhenabsenkung im hinteren Lautsprecher zum Einsatz. Ähnlich wurde mit der funky WahWah-Gitarre, welche aus dem Sampler kam, auf der rechten Seite verfahren.

Das Streicher-Pad wurde fünfmal kopiert, jeweils leicht in der Tonhöhe verändert und auf je einen anderen Lautsprecher verteilt, um so ein eindrucksvolles Umhüllungsgefühl zu schaffen.

Der Synthesizer-Sound wurde hauptsächlich auf LS und R gegeben. Dadurch ist er nicht klar zu lokalisieren, wird z.T. in Kopfnähe wahrgenommen und kippt je nach Hörposition. Dies war jedoch beabsichtigt.

Die Klarinette wurde ebenfalls vorne, aber innerhalb des Lautsprecherkreises platziert und ist näher am Hörer als Schlagzeug oder Kontrabaß. Auch hier kamen wieder Laufzeitdifferenzen zum Einsatz. Die Klarinette wurde künstlich verhallt, Hallanteile wurden auf alle Lautsprecher außer dem Center (und dem LFE) gegeben. Einzig das DI-Signal des Kontrabaß und das RE-20 der Bassdrum wurden etwas auf den LFE-Kanal gegeben.

Die meisten Spuren wurden etwas mit dem Equalizer bearbeitet. Dies geschah nach klangästhetischen Gesichtspunkten bzw. um das jeweilige Instrument besser in den Mix zu integrieren. Beim Mastering wurden nur geringfügige Klangkorrekturen vorgenommen. Die Summe wurde leicht komprimiert und etwas limitiert.

Bei Mix B wurde versucht, die Richtungen der Instrumente möglichst ähnlichen im Mix A, aber allein durch Pegeldifferenzen zu erzeugen. Während die seitlichen Instrumente Stick und WahWah-Gitarre, wie auch die Streicher und die Klarinette noch relativ ähnlich platziert werden konnten, gelang es nicht das Schlagzeug und den Kontrabaß ähnlich räumlich abzubilden. Hierzu könnte man noch Versuche mit künstlichen Raumsimulationen anstellen.

Zusammenfassend kann festgestellt werden, dass der Mix A, räumlicher, weiter, offener, der natürlichen Hörsituation ähnlicher und subjektiv eindrucksvoller und besser klingt. Der Mix B wirkt dagegen wesentlich enger, kompakter. Die Surround-Möglichkeiten konnten in Mix A eindeutig besser genutzt werden.

Insgesamt fällt auf, dass im Gegensatz zu zweikanaligen Mischungen, die einzelnen Spuren wesentlich weniger nachbearbeitet werden müssen, um eine gut gestaffelte, differenzierte Mischung zu erreichen.

### 8. Hörversuche

Die Misch-Experimente wie auch die Surround-Produktion haben viele interessante Erkenntnisse geliefert. Um sich jedoch nicht ausschließlich auf die in den Selbstversuchen gewonnenen Urteile zu verlassen, wurde ein Hörversuch mit 20 Probanden durchgeführt.

#### 8.1 Vorüberlegungen und Konzeption

Hörversuche unterliegen den subjektiven Empfindungen der Probanden und sind somit in den meisten Fällen fehlerbehaftet. Zwar ist laut ITU-Empfehlung<sup>43</sup> bei Hörversuchen ab 20 Probanden Signifikanz erreicht, so dass statistisch relevante Aussagen gemacht werden können, dennoch stellt diese Zahl natürlich nur einen winzigen Bruchteil aller Musikkonsumenten dar. Es lassen sich anhand der Versuchsergebnisse daher höchstens Tendenzen erkennen. Hörversuche sind jedoch letztlich die einzige Möglichkeit, Aufschluß über die Wahrnehmungen der Hörer zu bekommen.

In Vorversuchen zeigte sich, dass eine einseitige Fragestellung, bei der Richtung, Entfernung und räumliche Ausdehnung der Hörereignisse in eine Draufsicht des Abhörtraums eingezeichnet werden sollte, schwer auszuwertende Ergebnisse liefern und somit keine eindeutigen Aussagen zulassen würde. Aus diesem Grund bestand der Hörversuch zum größten Teil aus A/B-Vergleichen. Hauptsächlich ging es dabei um die subjektive Bewertung der räumlichen Abbildung und der Klangfarbe, um Stabilität und Umhüllung. Bei zwei Versuchsreihen wurden die Probanden außerdem gefragt, ob sie die hinteren Lautsprecher als störend wahrnehmen und ob sie den Eindruck haben, dass das Gesamtklangbild in einzelne Hörereignisse an den einzelnen Lautsprechern zerfällt.

---

<sup>43</sup> ITU-R BS. 1116-1, *Methods for the subjective assessment of small impairments in audio systems including multi-channel sound systems*, Rec., International Telecommunications Union, Geneva, Switzerland (1994-1997)

Laut ITU-Empfehlung<sup>44</sup> sollte die Versuchsdauer 30 Minuten nicht überschreiten, Herrmann und Henkels<sup>45</sup> sprechen von einer Stunde. Aufgrund der großen Anzahl der Hörbeispiele wurde die Dauer des Hörversuchs auf 50 Minuten festgelegt. Dennoch mußte eine Auswahl aus den zahlreichen Hörbeispielen vorgenommen werden. Dabei wurden soweit möglich solche Beispiele ausgewählt, bei denen relativ deutliche Unterschiede wahrzunehmen sind.

Für die A/B-Vergleiche wurde eine zur Mitte symmetrische, siebenstufige Skala gewählt, um eine möglichst differenzierte Beurteilung zu erhalten.

A				B		
viel besser	besser	etwas besser	gleich gut	etwas besser	besser	viel besser

Da man nicht davon ausgehen kann, dass jeder Proband unter „etwas besser“, „besser“ bzw. „viel besser“ das gleiche versteht, kommt nur eine *Ordinalskala*<sup>46</sup> in Betracht. Bei allen A/B-Vergleichen wurde jeweils ein Beispiel wiederholt und zusätzlich ein Vergleich mit identischen Hörbeispielen durchgeführt, um die *Reliabilität* und die *Validität*<sup>47</sup> zu prüfen.

Da die Abhörlautstärke bekanntlich großen Einfluß auf die Wahrnehmung hat (lautere Beispiele werden meist als „besser“ beurteilt), wurden die Hörbeispiele aneinander angeglichen. Dabei ist festzustellen, dass trotz Schalldruckpegelmeßgerät und zusätzlicher Gehöranpassung teilweise leicht unterschiedliche Lautstärkeempfindungen aufgrund eines etwas anderen Richtungs- oder Entfernungseindrucks auftreten.

---

<sup>44</sup> ITU-R BS. 1116-1, *Methods for the subjective assessment of small impairments in audio systems including multi-channel sound systems*, Rec., International Telecommunications Union, Geneva, Switzerland (1994-1997).

<sup>45</sup> Herrmann, U./Henkels, V., *Vergleich fünf verschiedener Surround-Hauptmikrofonverfahren*, Diplomarbeit, FH Düsseldorf, März 1997

<sup>46</sup> Bei der *Ordinalskala* gibt es keine definierte Maßeinheit, es bleibt undefiniert, wie groß die Unterschiede zwischen den einzelnen Antwortklassen sind. Es wird nur eine Rangreihenfolge gebildet.

<sup>47</sup> Die *Reliabilität* zeigt die Zuverlässigkeit eines Verfahrens an. Reliabilität ist gegeben, wenn bei Versuchswiederholung in kurzem zeitlichen Abstand mindestens 70% der Probanden die gleiche oder nur eine geringfügig abweichende Antwort geben. Die *Validität* gibt an, ob die Ergebnisse Gültigkeit besitzen. Dabei müssen bei Vergleich identischer Beispiele 60-70% der Antworten in der Antwortklasse „gleich“ auftreten. (Zöfel, P., *Statistik in der Praxis*, Fischer Verlag, Stuttgart, 1985)

### 8.2 Ablauf der Hörversuche

Die Versuche fanden in der Regie A im Tonstudio der Hochschule der Medien statt. Die dortige 3/2-Stereo-Abhöranlage wie auch der Regieraum entsprechen den ITU-Richtlinien. Bei den Lautsprechern handelt es sich um aktive 2-Wege-Nahfeldmonitore des Herstellers A.D.A.M., vom Typ S2 (A).

Um die Testpersonen möglichst wenig visuell abzulenken, wurden Computermonitore, Pegelanzeigen, etc. ausgeschaltet bzw. abgedeckt. Die Hörversuchssession wurde in Pro Tools so angelegt, dass die Probanden keine Anzeichen auf die Hörbeispiele erhielten, aus denen man Rückschlüsse auf das Gehörte hätte ziehen können.

Jeder Proband wurde vor Beginn der Hörversuche explizit darauf hingewiesen, dass es in diesen Versuchen nicht um die Beurteilung irgendwelcher Hörfähigkeiten geht, sondern um subjektive Höreindrücke. Es gibt dabei keine richtige oder falsche Antwort. Die Versuchspersonen sollten deshalb ehrlich antworten und angeben was sie tatsächlich hören, und bei Unsicherheiten besser nichts ankreuzen. Bei jedem Versuch bestand zudem die Möglichkeit schriftliche Anmerkungen zu machen.

Die Versuchspersonen wurden im sweet spot plaziert. Im Gegensatz zu vielen anderen Hörversuchen war es ihnen gestattet sich zu bewegen. Sie wurden bei einigen Fragen sogar ausdrücklich gebeten sich vor und zurück zu beugen und den Kopf zu drehen, um die Hörbeispiele auch in anderen Positionen und unter nicht ganz idealen Abhörbedingungen zu beurteilen. Bedenkt man, dass es kaum Musikkonsumenten gibt, die unter den üblichen Versuchsbedingungen (ideale Hörposition, stillgehaltener Kopf) Musik hören, und es in diesem Hörversuch nicht um exakt messbare Werte, sondern um die subjektiven Eindrücke der Probanden ging, so ist ein solches Vorgehen eher aussagekräftig.

Es wurde empfohlen, sich den ersten Durchgang eines A/B-Vergleichs am optimalen Abhörplatz anzuhören und bei Wiederholung andere Positionen auszuprobieren. Jeder Vergleich konnte auf Wunsch der Probanden mehrfach wiederholt werden.

Die Teilnehmer erhielten zu jeder Frage des Hörversuchs ausführliche mündliche Erklärungen und wurden gebeten bei Unklarheiten über die Fragestellung oder einzelner Begriffe nachzufragen. Zum Teil wurden den Probanden zusätzlich Referenzbeispiele vorgespielt.

### *Fragen zum Hörversuch*

In den Hörbeispielen zu Frage 1 wurden verschiedene Instrumente (Bassdrum, Snare, Stimme, Flügel) in unterschiedlichen Variationen auf der vorderen Lautsprecherfront plaziert. Dabei wurde der Center mit der Phantommitte aus L-R und der Kombination beider verglichen. (Versuch 1, Kapitel 6.1) Die Frage war, welches Beispiel dem Hörer bezüglich Abbildung und Klangfarbe besser gefällt.

In Frage 2 ging es um die seitliche Abbildung eines Flügels. Verschiedene Mono- und Stereoverionen sollten bezüglich der räumlichen Abbildung und der Klangfarbe beurteilt werden. (Versuch 3, Kapitel 6.2)

In den Versuchen zu Frage 3 wurden eine Singstimme bzw. eine Snare in verschiedenen Variationen auf vordere und hintere Lautsprecher gegeben und im vorderen Bereich des Surroundfeldes plaziert. (Versuch 5 & 6, Kapitel 6.3) Wieder sollte die räumliche Abbildung und die Klangfarbe beurteilt werden. Es wurde angemerkt, auf die Stabilität des Abbildung zu achten. Zusätzlich wurde gefragt, ob die hinteren Lautsprecher störend wahrgenommen werden und ob der Hörer den Eindruck hat, dass das Gesamtbild in Klangwolken bei den einzelnen Lautsprechern zerfällt.

In Frage 4 sollten die Probanden ankreuzen, bei welchem Beispiel sie sich besser vom Klang umhüllt fühlen. (Versuch 7 & 8, Kapitel 6.4)

Im letzten Hörvergleich (Frage 5) wurde den Teilnehmern ein komplettes Musikstück (Kapitel 7) in zwei verschiedenen Mischungen vorgespielt. Die Probanden sollten angeben, welche Mischung ihnen besser gefällt und wenn möglich begründen warum.

Weiter war noch von Interesse ob die Probanden in irgendeiner Form im Audio-bereich tätig sind, Hörerfahrung haben oder ein Instrument spielen. Außerdem wurden sie gefragt was sie für Musik und wie sie in der Regel Musik hören (bewußt, nebenher, vor der Stereoanlage sitzend, im Auto, beim Kochen,...) Ebenso sollten sie angeben, ob sie zu Hause eine Surround-Anlage haben, wie diese – falls vorhanden – aufgestellt ist und ob sie Surround-Produktionen und wenn in welchen Formaten (DVD-Video, DVD-Audio, SACD, DTS-CD) besitzen. Ein Abdruck des gesamten Hörversuchsbogens findet sich im Anhang.

### **8.3 Auswertung und Ergebnisse**

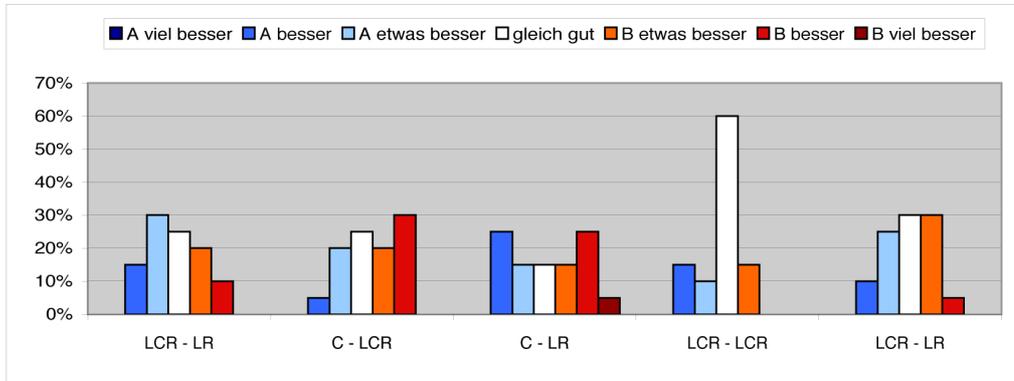
Die Auswertung der Fragebögen erfolgte mit Hilfe der Software Excel von Microsoft. Es wurden die absoluten Häufigkeiten in den Antwortklassen gezählt. Anschließend wurden die prozentualen Häufigkeiten, der Median, sowie das 1. und 3. Quartil berechnet.

Der Median ist derjenige Punkt der Skala, unterhalb und oberhalb dessen jeweils die Hälfte der Antworten liegt. Er zeigt ähnlich dem arithmetischen Mittel eine Tendenz der Antworten an und ist gegenüber einzelnen extremen Antworten relativ unempfindlich. Das 1. und 3. Quartil sind die Punkte der Skala unterhalb dessen 25% (1. Quartil) bzw. 75% (3. Quartil) der Antworten liegen. Liegen die beiden Quartile nah um den Median, so ist die Mehrzahl der Antworten sehr ähnlich, bei großen Abständen dagegen sind die Antworten stark gestreut. Die Ergebnisse wurden zusätzlich noch grafisch ausgewertet.

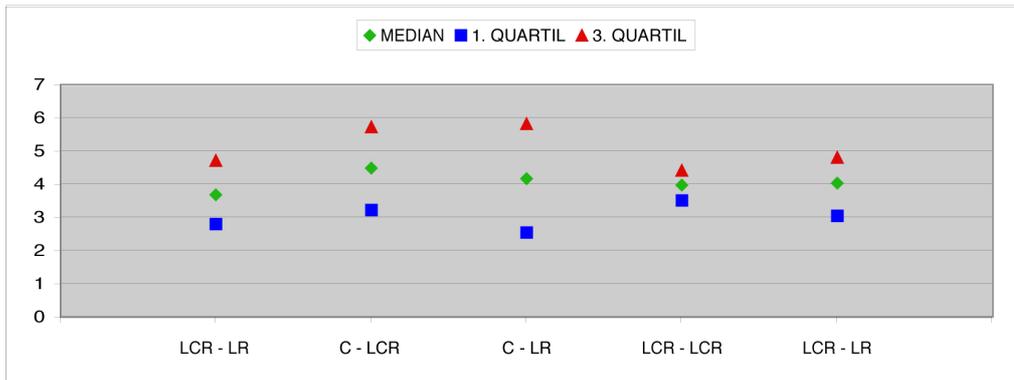
Beim überwiegenden Anteil der 20 Probanden handelte es sich um musikinteressierte Teilnehmer. 18 davon spielen ein oder mehrere Instrumente, singen oder sind sonst in irgendeiner Weise musikalisch tätig. Zwölf Teilnehmer sind Studenten der Hochschule der Medien. Insgesamt waren drei Probanden mit großer, elf mit leichter tontechnischer Hörerfahrung und sechs tontechnische Laien an den Hörversuchen beteiligt.

Frage 1.1 – Abbildung vorne, *Bassdrum*

„Welches Beispiel gefällt Ihnen bezüglich Abbildung und Klangfarbe besser?“



Prozentuale Häufigkeiten Frage 1.1, Bassdrum

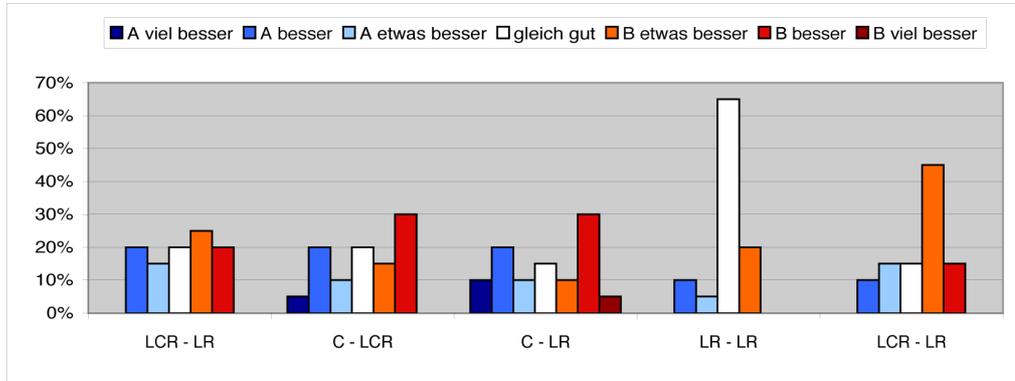


Median- und Quartilwerte Frage 1.1, Bassdrum

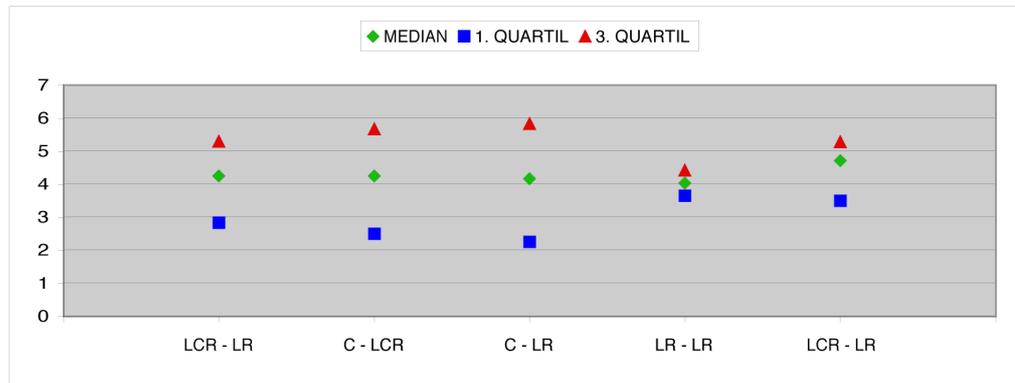
Bei den Vergleichen mit der Bassdrum liegen alle Mediane im Bereich der Antwortklasse 4 („gleich gut“). Die Antworten sind relativ stark gestreut, ein deutlicher Favorit ist nicht auszumachen. Will man eine Reihenfolge aufstellen, so liegen die Phantommitte aus L-R und die L-C-R-Kombination etwas vor der reinen Center-Version. Zwischen den beiden gibt es eine leichte Tendenz zugunsten der 3-Kanal-Variante, allerdings ist bei der Wiederholung des direkten Vergleichs kein Vorsprung mehr zu erkennen (Median exakt 4). Zwar sind einige Abweichungen zum ersten Vergleich vorhanden, die Reliabilität ist aber noch gegeben. Bei dem identischen Vergleich liegen 60% der Antworten in Klasse 4, somit ist auch noch Validität erreicht.

Frage 1.2 – Abbildung vorne, *Snare*

„Welches Beispiel gefällt Ihnen bezüglich Abbildung und Klangfarbe besser?“



Prozentuale Häufigkeiten Frage 1.2, Snare

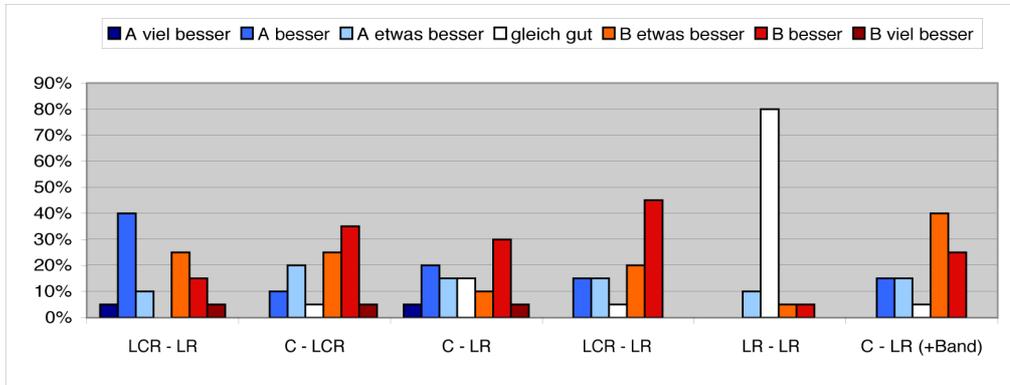


Median- und Quartilwerte Frage 1.2, Snare

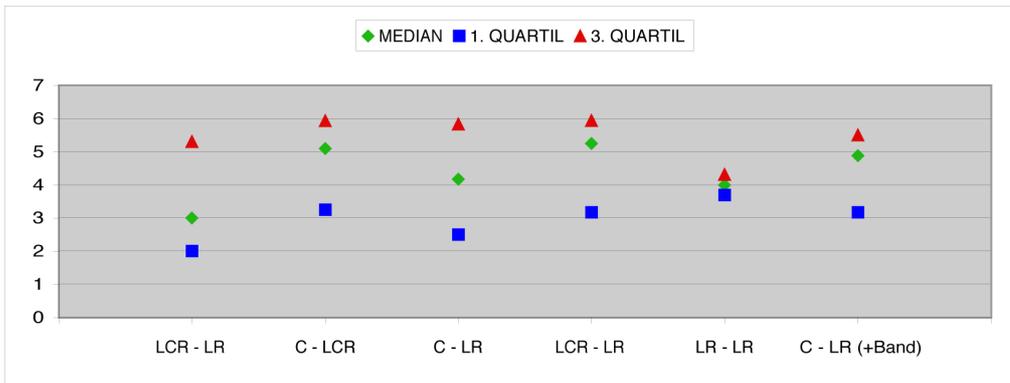
Auch bei den Snare-Vergleichen liegen fast alle Mediane im Bereich der Antwortklasse 4. Die Antworten sind ebenfalls ziemlich gestreut, vor allem beim Vergleich zwischen Center und Phantommitte gehen die Meinungen stark auseinander. Das reine Centersignal wurde wiederum am schwächsten bewertet, die Phantommitte – vor allem im Wiederholungsversuch zu erkennen – am meisten bevorzugt. Die Reliabilität ist nur eingeschränkt vorhanden, rund 40% der Probanden wichen von ihrer ersten Meinung ab. Die Validität ist gegeben, beim identischen Vergleich antworteten 65% mit „gleich gut“.

Frage 1.3 – Abbildung vorne, *Stimme*

„Welches Beispiel gefällt Ihnen bezüglich Abbildung und Klangfarbe besser?“



Prozentuale Häufigkeiten Frage 1.3, Stimme

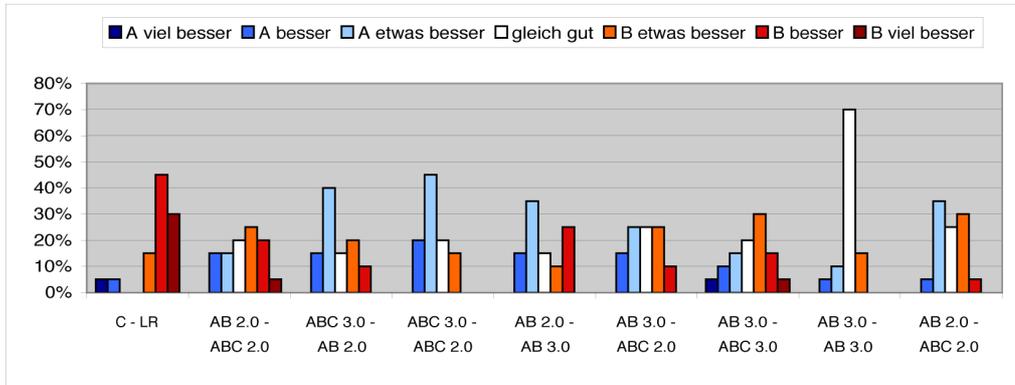


Median- und Quartilwerte Frage 1.3, Stimme

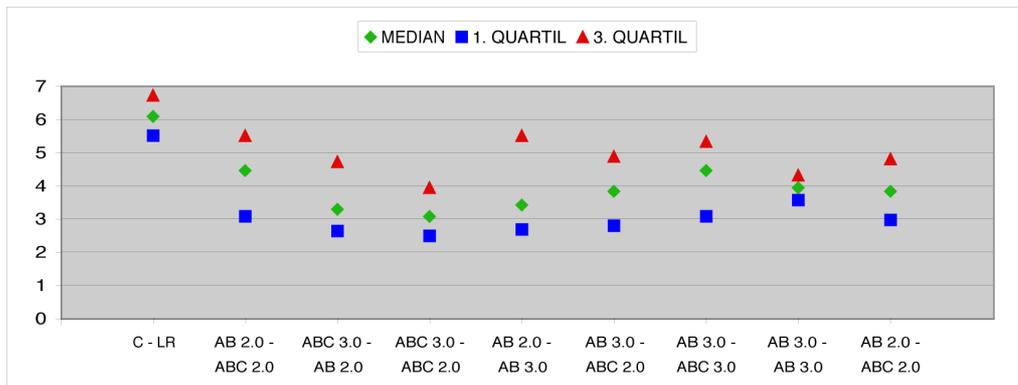
Bei den Stimmbeispielen liegen die Mediane erstmals deutlich in verschiedenen Antwortklassen. Auch hier gehen die Meinungen beim Vergleich zwischen Center und Phantommitte am stärksten auseinander. Der Center fällt aber insgesamt gegenüber den anderen beiden, vor allem zur 3-Kanal-Version ab. Zwischen Phantommitte und L-C-R-Kombination lassen die Antworten keine klare Präferenz erkennen. Im Wiederholungsversuch änderten über 50% der Versuchsteilnehmer ihre Meinung, 25% wechselten sogar die Seite, Reliabilität ist somit nicht gegeben. Beim identischen Vergleich liegen 80% der Antworten in AK 4, damit wird Validität erreicht. Bei dem zusätzlichen Vergleich mit auf L-R gemischter Band wurde die Phantommitte „etwas besser“ (AK 5) beurteilt.

Frage 1.4 – Abbildung vorne, *Piano*

„Welches Beispiel gefällt Ihnen bezüglich Abbildung und Klangfarbe besser?“



Prozentuale Häufigkeiten Frage 1.4, Piano

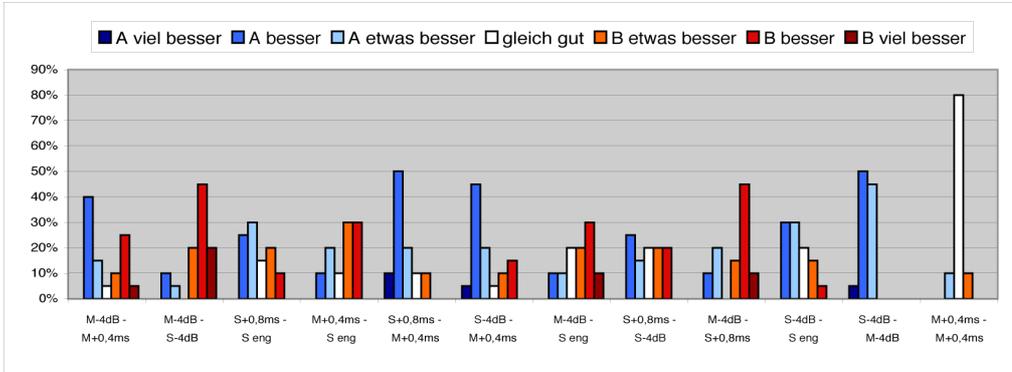


Median- und Quartilwerte Frage 1.4, Piano

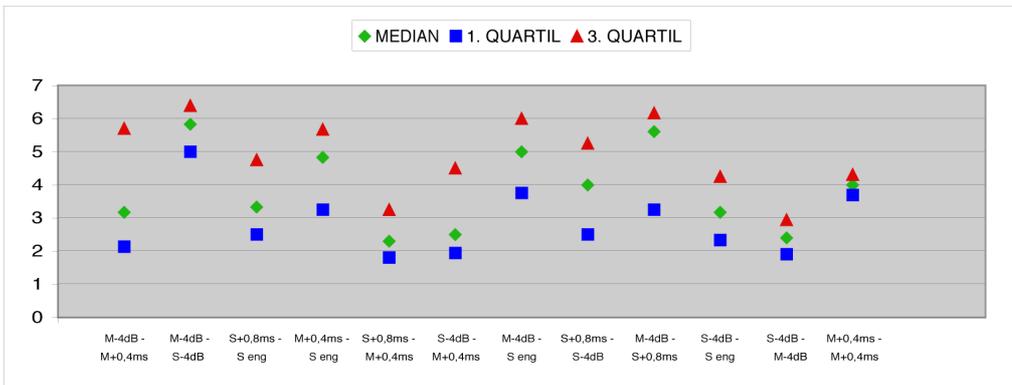
Der Monovergleich geht klar zugunsten der Phantommitte aus. Bei den Stereoverversionen wurde die dreikanalige Version (ABC 3.0), bei der jedes Mikrofon auf einen anderen Lautsprecher gegeben wurde, immer als „etwas besser“ beurteilt. Unter den anderen Versionen ist die Reihenfolge nicht eindeutig. Interessanterweise wurden die Versionen ABC 2.0 (Centermikrofon wurde auf L-R hinzuge-mischt) und AB 3.0 (Centersignal aus Summierung von L und R generiert), bei denen durch Kammfiltereffekte ein gefärbtes, verwascheneres Klangbild ent-steht, nicht eindeutig schlechter als die reine AB 2.0-Stereofassung bewertet. Die Reliabilität ist eingeschränkt (10% wechseln die Seite), Validität wird aber erreicht.

Frage 2.1 – Seitliche Abbildung zwischen L und LS, *Piano*

„Welches Beispiel gefällt Ihnen besser in Bezug auf die räumliche Abbildung?“



Prozentuale Häufigkeiten Frage 2.1, räumliche Abbildung

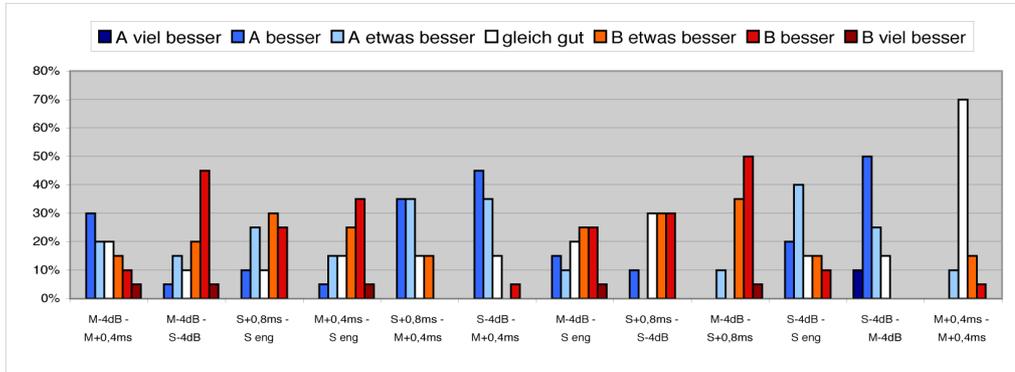


Median- und Quartilwerte Frage 2.1, räumliche Abbildung

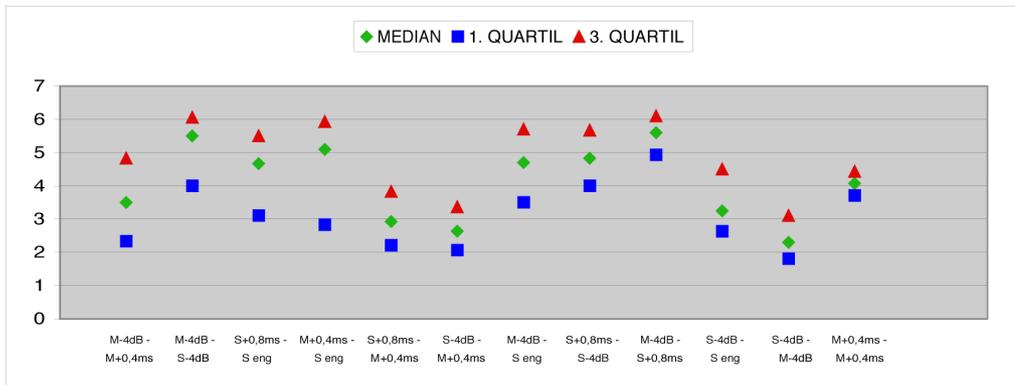
Wie nicht anders zu erwarten fallen die monofonen Beispiele deutlich gegenüber den Stereoverionen ab. Die Pegelvariante (M -4 dB) wurde dabei der Laufzeitversion (M +0,4 ms) etwas vorgezogen. Bei den Stereobeispielen sind Pegel- (S -4 dB) und Laufzeitversion (S +0,8 ms) bei der Frage nach der räumlichen Abbildung fast gleich auf, mit einer Nuance Vorsprung der Laufzeit. Die mittels Panpots in der Abbildung eingengte Version (S eng) überzeugte weniger. Hier tritt der unangenehm, beengende Höreindruck wieder auf. Im Wiederholungsvergleich wechselten 40% die Seite, Reliabilität ist somit nicht gegeben. Die Validität wird mit 80% der Antworten in AK 4 klar erreicht.

Frage 2.2 – Seitliche Abbildung zwischen L und LS, *Piano*

„Welches Beispiel gefällt Ihnen besser in Bezug auf die Klangfarbe?“



Prozentuale Häufigkeiten Frage 2.2, Klangfarbe

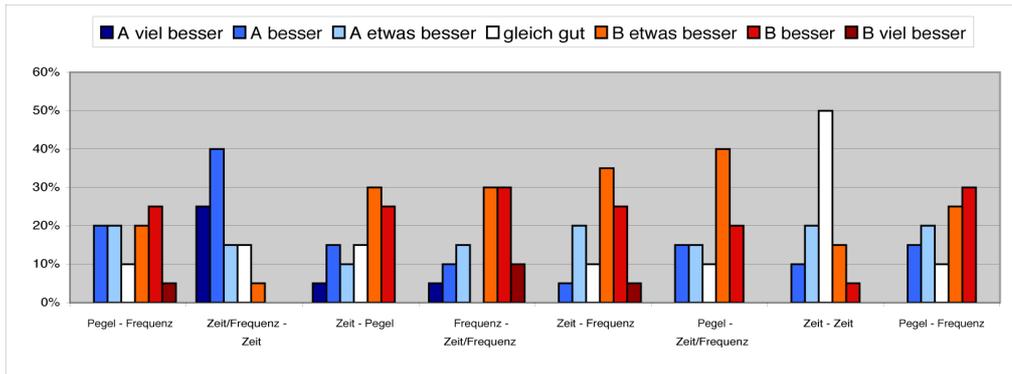


Median- und Quartilwerte Frage 2.2, Klangfarbe

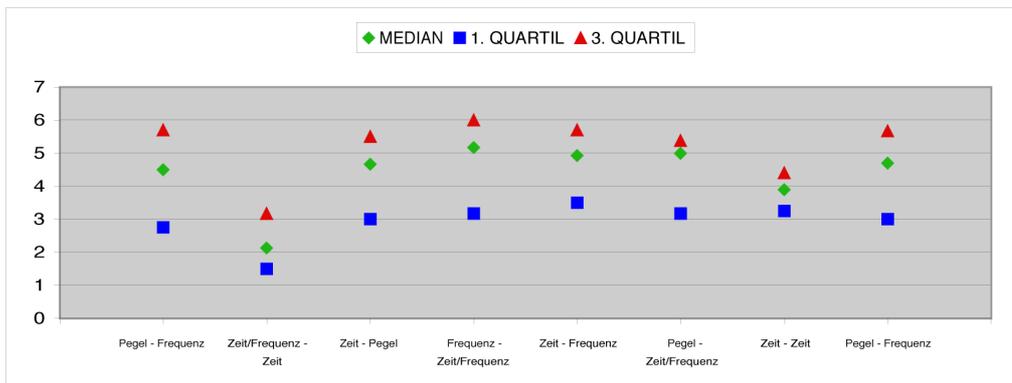
Bei der Bewertung der Klangfarbe der Hörbeispiele fallen die Ergebnisse relativ ähnlich aus. Allerdings wurden sowohl die Mono- als auch die Stereoverion der Pegelvariante bezüglich der Klangfarbe gegenüber den Laufzeitbeispielen etwas bevorzugt. Die Reliabilität ist leicht eingeschränkt, 20% der Probanden wechselten die Seite, Validität wird mit 70% der Antworten in AK 4 erreicht.

Frage 3.1 – Vorne-Hinten-Lokalisation, *Stimme*

„Bei den folgenden Hörbeispielen wurde versucht das Instrument aus der vorderen Lautsprecherbasis in den Raum reinzuholen ohne jedoch die Vorzugsrichtung "vorne mittig" aufzugeben. Bei welchem Beispiel scheint Ihnen dies besser gelungen?“



Prozentuale Häufigkeiten Frage 3.1, Stimme, Abbildung

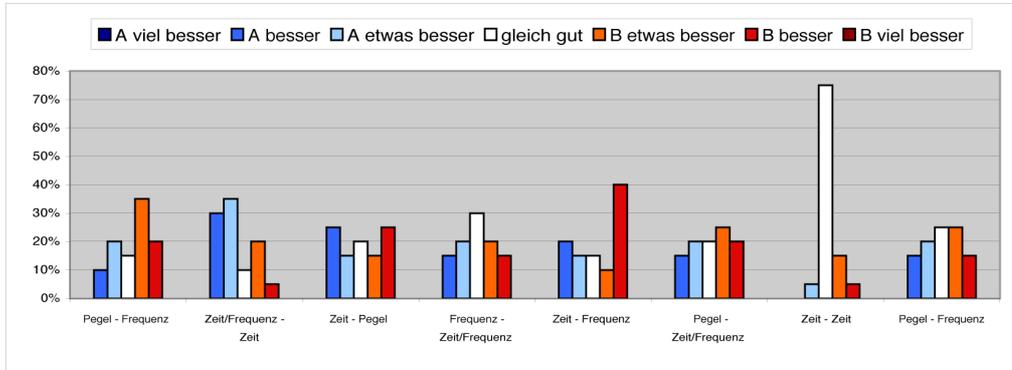


Median- und Quartilwerte Frage 3.1, Stimme, Abbildung

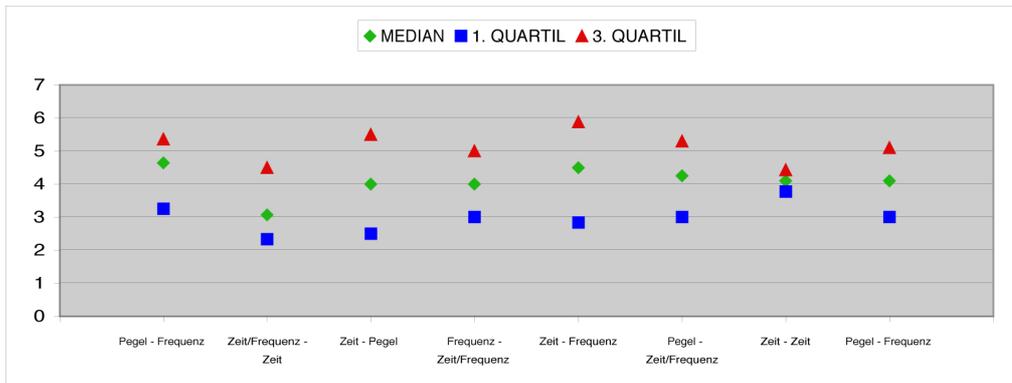
In Bezug auf die räumliche Abbildung kann eine ziemlich klare Reihenfolge festgestellt werden. Die Kombination aus Verzögerung (+6 ms) und Filterung (-6 dB bei 1,2 kHz, 8 kHz und oberhalb 12 kHz) der hinteren Lautsprecher liegt deutlich vor der reinen Filter-Variante, diese knapp vor der Pegel-Version (hintere Lautsprecher -6 dB) und den Schluß macht das Beispiel mit reinen Laufzeitunterschieden. Im Wiederholungsversuch änderten nur 20% ihre Meinung, Reliabilität ist somit gegeben. Validität wird jedoch nicht erreicht (50% in AK 4).

Frage 3.2 – Vorne-Hinten-Lokalisation, *Stimme*

„Welches Beispiel gefällt Ihnen klanglich besser?“



Prozentuale Häufigkeiten Frage 3.2, Stimme, Klangfarbe

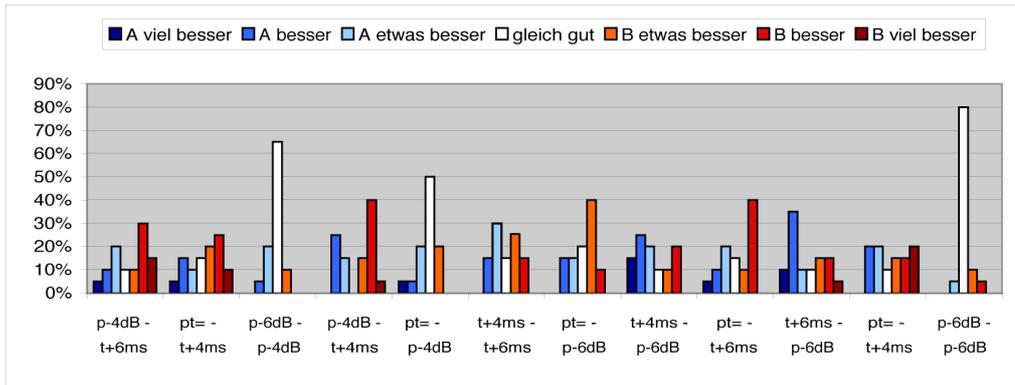


Median- und Quartilwerte Frage 3.2, Stimme, Klangfarbe

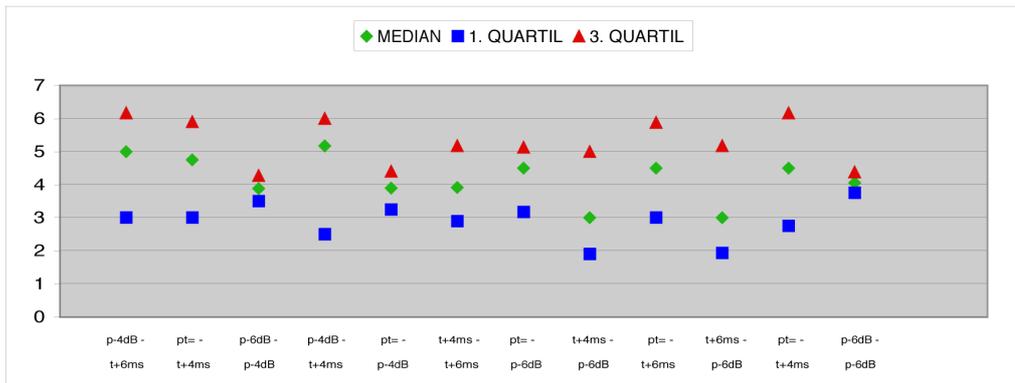
Auch bei der Beurteilung der Klangfarbe läßt sich eine Reihenfolge erkennen. Hierbei wurden die Beispiele mit der Frequenzabsenkung auf den hinteren Lautsprechern bevorzugt, ein klarer Favorit zwischen den beiden ist aber nicht auszumachen. Die reine Laufzeitversion schneidet wiederum am schlechtesten ab. Die Reliabilität ist leicht eingeschränkt, da doch einige Probanden ihre Meinung änderten, es gibt aber kaum Seitenwechsel. Da bei dem identischen Vergleich 75% der Antworten in Klasse 4 liegen, ist Validität gegeben.

Frage 3.3 – Vorne-Hinten-Lokalisation, *Snare*

„Bei den folgenden Hörbeispielen wurde versucht das Instrument aus der vorderen Lautsprecherbasis in den Raum reinzuholen ohne jedoch die Vorzugsrichtung "vorne mittig" aufzugeben. Bei welchem Beispiel scheint Ihnen dies besser gelungen?“



Prozentuale Häufigkeiten Frage 3.3, Snare, Abbildung

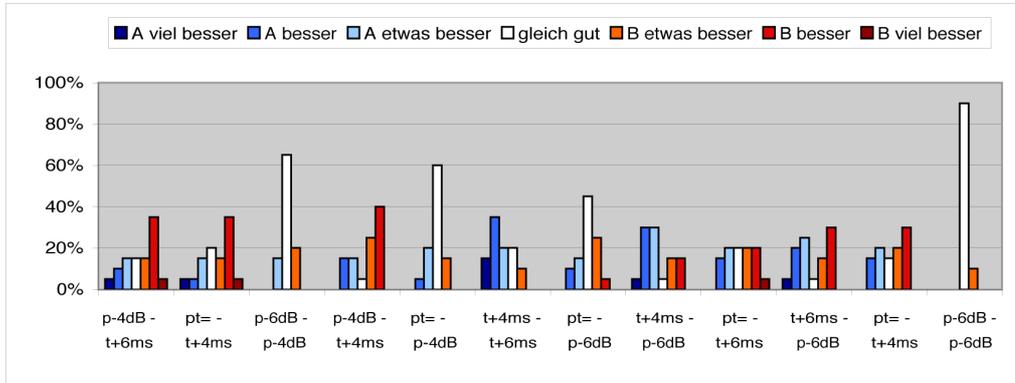


Median- und Quartilwerte Frage 3.3, Snare, Abbildung

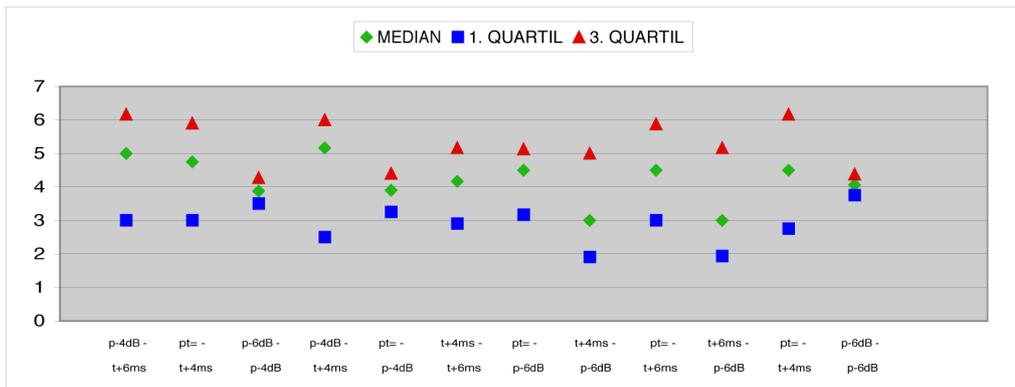
Bei den Versuchen die Snare innerhalb des Lautsprecherkreises abzubilden, wurden beide Laufzeitbeispiele (t +4 ms/t +6 ms) gegenüber den Pegelvarianten bevorzugt. Zwischen Pegel- und Laufzeitbeispielen sind jedoch große Streuungen zu erkennen. Die Verzögerung von L-R um 2 ms, und von LS-RS um 4 ms, wurde am besten beurteilt. Unter den Pegeldifferenzen (p -4 dB/p -6 dB) sind die Unterschiede wesentlich geringer, auch gegenüber der Variante mit gleichen Pegeln (pt=). Die Reliabilität ist eingeschränkt, Validität wird erreicht.

Frage 3.4 – Vorne-Hinten-Lokalisation, *Snare*

„Welches Beispiel gefällt Ihnen klanglich besser?“



Prozentuale Häufigkeiten Frage 3.4, Snare, Klangfarbe

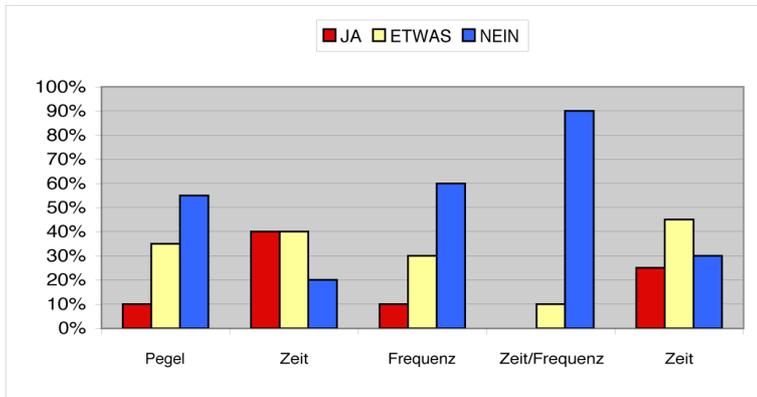


Median- und Quartilwerte Frage 3.4, Snare, Klangfarbe

Die Beurteilung der Klangfarbe weicht nicht wesentlich von der Ergebnissen zur räumlichen Abbildung ab. Hier ist jedoch die Laufzeitvariante mit 4 ms Verzögerung eindeutiger Favorit. Die Pegelbeispiele wurden – wie nicht anders zu erwarten – alle relativ ähnlich bewertet. Auch hier sind die Meinungen z.T. stark gestreut, die Unterschiede insgesamt aber geringer. Beim Wiederholungsversuch gab es einige Meinungsänderungen, die Reliabilität ist deshalb auch hier eingeschränkt, Validität ist aber mit 90% klar vorhanden.

Frage 3.5 – Vorne-Hinten-Lokalisation, *Stimme*

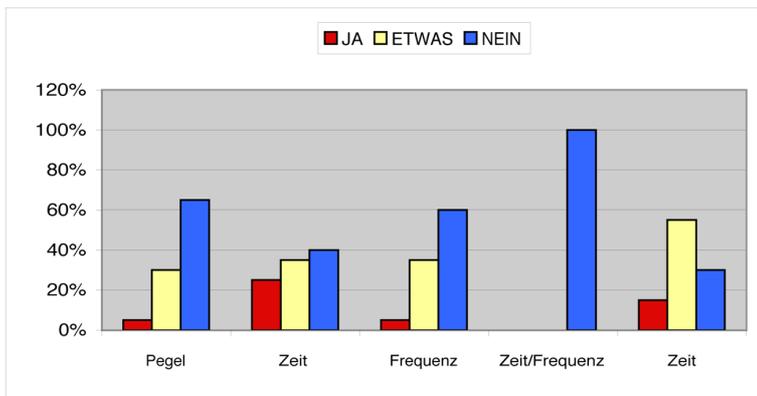
„Nehmen Sie die Surround-Lautsprecher als "störend" wahr?“



Prozentuale Häufigkeiten Frage 3.5, Stimme, störende Surround-Lautsprecher

Frage 3.6 – Vorne-Hinten-Lokalisation, *Stimme*

„Haben Sie den Eindruck, dass das Gesamtklangbild in Klangwolken bei den einzelnen Lautsprechern zerfällt?“

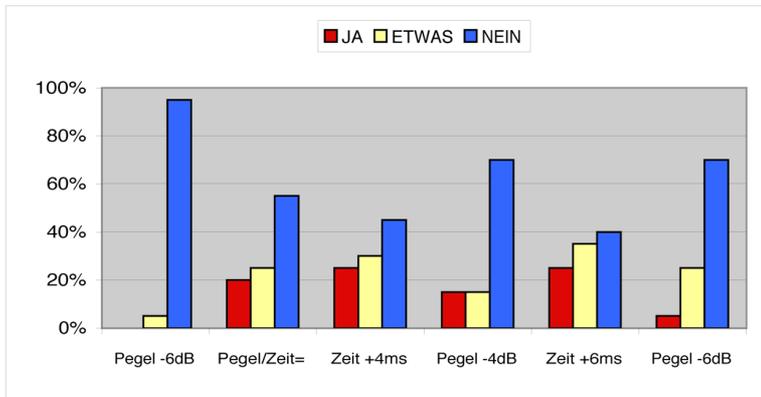


Prozentuale Häufigkeiten Frage 3.6, Stimme, Gesamtklangbild zerfällt

Bei reinen Laufzeitunterschieden wurden die Surround-Lautsprecher am häufigsten störend wahrgenommen, ebenso ein Zerfallen des Klangbildes. Bei den Beispielen mit Pegelabsenkung und Filterung fiel dies nur leicht und weniger störend auf. Die Kombination aus Laufzeitdifferenz und zusätzlicher Filterung der hinteren Lautsprecher wurde am positivsten beurteilt.

Frage 3.7 – Vorne-Hinten-Lokalisation, *Snare*

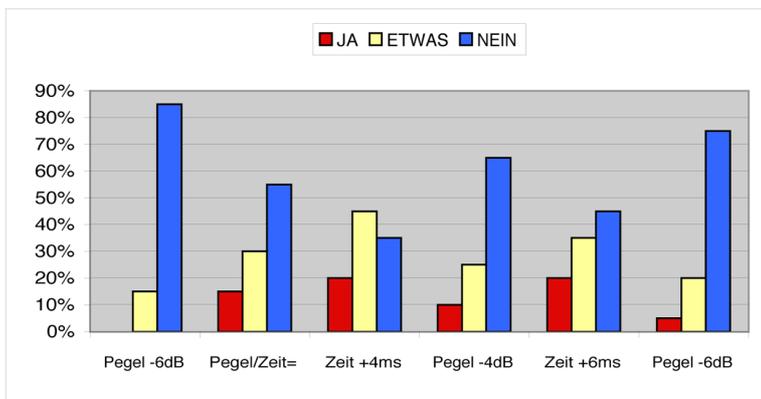
„Nehmen Sie die Surround-Lautsprecher als "störend" wahr?“



Prozentuale Häufigkeiten Frage 3.7, Snare, störende Surround-Lautsprecher

Frage 3.8 – Vorne-Hinten-Lokalisation, *Snare*

„Haben Sie den Eindruck, dass das Gesamtklangbild in Klangwolken bei den einzelnen Lautsprechern zerfällt?“

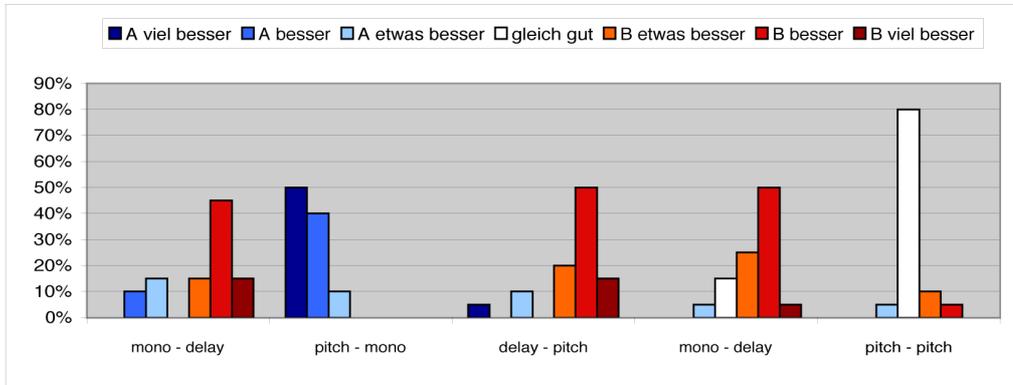


Prozentuale Häufigkeiten Frage 3.8, Snare, Gesamtklangbild zerfällt

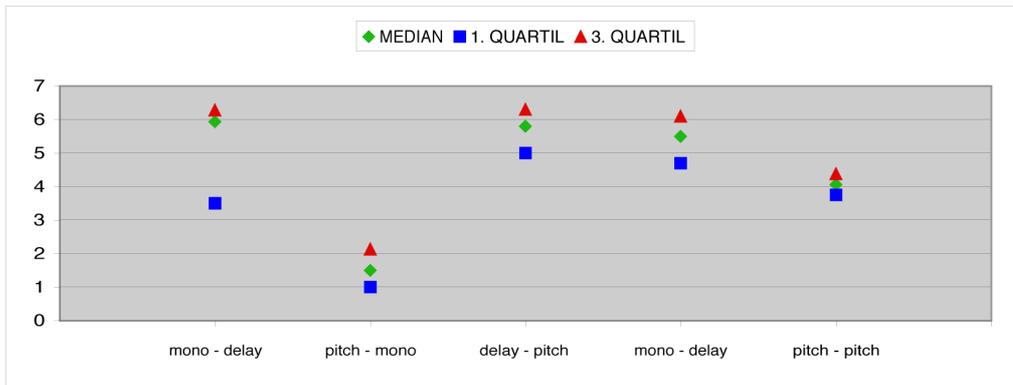
Hier zeigt sich ebenfalls, dass bei Laufzeitunterschieden ohne zusätzliche Filterung die negativen Aspekte am häufigsten wahrgenommen wurden. Bei Pegelreduktion um 6 dB fielen die hinteren Lautsprecher fast gar nicht mehr störend auf, auch ein Zerfallen des Klangbildes wurde nicht wahrgenommen.

Frage 4.1 – Umhüllung, *Streicher-Pad*

„Bei welchem Beispiel fühlen Sie sich besser vom Klang umhüllt?“



Prozentuale Häufigkeiten Frage 4.1, Streicher-Pad, Umhüllung

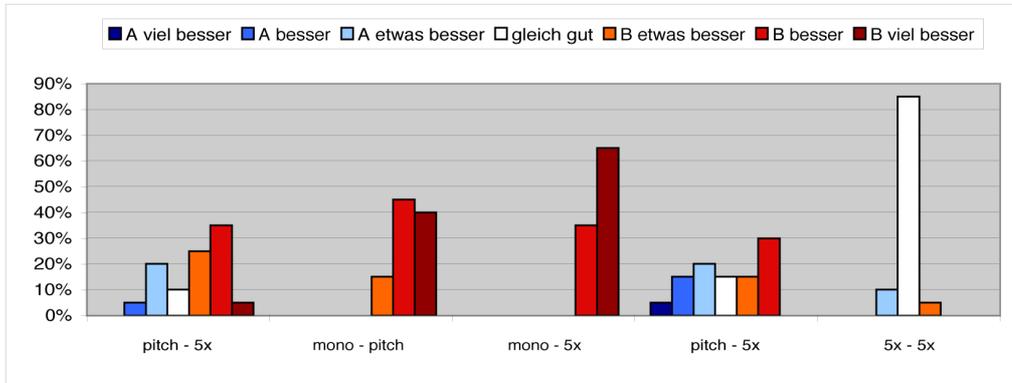


Median- und Quartilwerte Frage 4.1, Streicher-Pad, Umhüllung

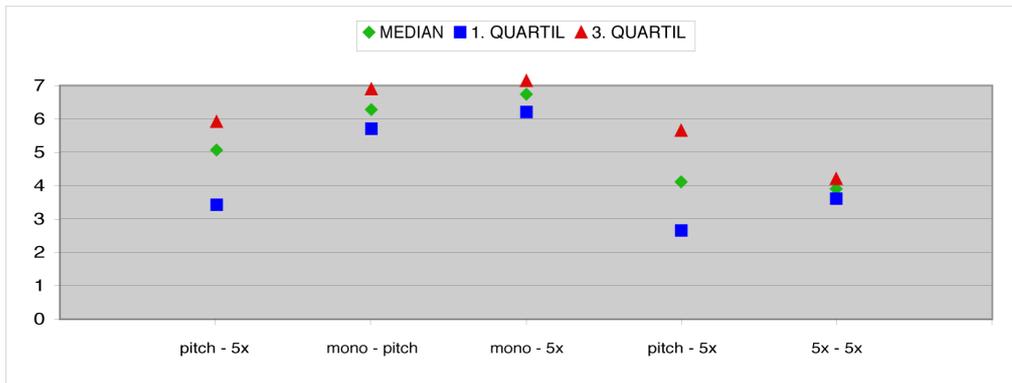
Bei der Frage nach der besten Umhüllung, wurde das Beispiel mit den leicht unterschiedlich gepitchten Spuren (Tonhöhe  $\pm 5$  Cent) eindeutig bevorzugt. Die Monovariante erzielte den geringsten Umhüllungseindruck. Die Reliabilität ist in diesem Versuch stark eingeschränkt, unterschiedliche Beurteilungen und 20% Seitenwechsel. Die Validität ist wiederum gegeben.

Frage 4.2 – Umhüllung, E-Gitarre

„Bei welchem Beispiel fühlen Sie sich besser vom Klang umhüllt?“



Prozentuale Häufigkeiten Frage 4.2, E-Gitarre, Umhüllung



Median- und Quartilwerte Frage 4.2, E-Gitarre, Umhüllung

Auch in diesem Versuch überzeugte die monofone Version am wenigsten. Die Pseudosurround-Variante mit leicht unterschiedlich gepitchter Tonhöhe wurde ähnlich gut wie das Beispiel mit fünfmal eingespielter Gitarre bewertet. Beim Vergleich der fünfmal mit etwas variiertes Spielweise und dezenter Klangänderungen eingespielten Gitarre, mit der Pitchvariante fällt eine relative große Streuung der Antworten auf. Die Reliabilität ist auch hier eingeschränkt, die Validität ist mit 85% der Antworten in AK 4 gegeben.

Bei dem Vergleich der beiden unterschiedlichen Musik-Mischungen favorisierten 13 der 20 Probanden (65%) den Mix A, der neben Pegelunterschieden noch mit zusätzliche Laufzeitdifferenzen und Filterungen erstellt wurde. Dazu ein paar Begründungen und Kommentare von Versuchsteilnehmern:

*„Mix A besser, weil Instrumente räumlicher abgebildet“, „klingt verspielter“, „es wird Platz im Raum geschaffen, Schallquellen wirken nicht so gedrängt“, „Vorteile von Surround besser genutzt“, „etwas luftiger“, „umschließt einen perfekt“, „differenzierter, klarer, druckvoller“, „Mix B zentraler, nicht so offen“, Schlagzeug klingt viel fetter“.*

Auch einer der Musiker der Band nahm an den Hörversuchen teil. Interessanterweise bevorzugte dieser – wenn auch nur geringfügig – den Mix B, mit der Begründung, dass er für diese Musikrichtung typischer sei. Mix A höre sich zwar eher so an, wie die Band in natura klingt, als Studioaufnahme aber gewöhnungsbedürftiger. Auch hierzu ein paar Bemerkungen von Probanden:

*„Mix B runder, weicher, nicht so differenziert, dadurch dichtere und vollere Wirkung“, „Mix B besser, weil direkter“, „Streicher fülliger“, „kompakter“, „Bass kommt besser raus“.*

Bei den restlichen Fragen zeigte sich, dass bei den Musikvorlieben so ziemlich alle Stilistiken von Rock, Pop, Funk, Soul, Metal, Elektro, TripHop und HipHop, über Jazz bis hin zu klassischer Musik vertreten waren. Ebenso breit gestreut waren die Hörgewohnheiten, von bewußtem Zuhören vor der Stereoanlage, über im Auto und beim Baden, über Kopfhörer oder PC-Lautsprecher.

Nur zwei der Versuchsteilnehmer gaben an eine Surroundanlage zu besitzen, diese aber nur für Filme oder Konzertmitschnitte (mit Bild) zu nutzen. Bei der Frage nach Surround-Produktionen verhält es sich nicht anders. Zwar haben immerhin neun der Probanden DVD-Videos, DVD-Audio oder SACD allerdings nur zwei Personen und diese besitzen keine Anlage.

### 8.4 Fazit

Im Großen und Ganzen wurden die in den Selbstversuchen gewonnenen Erkenntnisse durch die Hörversuche bestätigt. Auch hat sich wieder einmal gezeigt, dass es im künstlerisch-gestalterischen Bereich oftmals kein eindeutiges Besser oder Schlechter gibt. Gerade bei der Frage der Mittenabbildung vorne fällt auf, dass die ästhetischen Vorlieben der Probanden sich stark unterscheiden. Während der eine die Phantommitte als „sehr unkonkret“ oder „zu räumlich“ und das reine Centersignal als „präziser, besser“ bezeichnet, empfindet der andere genau umgekehrt den Center als „zu eng“ und die Phantommitte als „natürlicher, räumlicher“.

Zwar war bei den meisten Beispielen eine ganz leichte Tendenz zugunsten der L-C-R-Version zu erkennen, der Median lag aber fast immer in der Antwortklasse *gleich gut*. Vor allem bei den Schlagzeug-Beispielen äußerten viele der weniger hörerfahrenen Probanden, dass es ihnen sehr schwer falle, überhaupt Unterschiede wahrzunehmen oder eine Bewertung abzugeben. Bei den Versuchen mit der Stimme entschieden sich im Wiederholungsversuch ein Großteil der Versuchsteilnehmer anders als im ersten Durchgang.

Bei der seitlichen Abbildung liefern Stereosignale, bei denen der hintere Lautsprecher im Pegel reduziert ist, die überzeugendsten Resultate bezüglich Abbildung und Klangfarbe. Es ist daher bei Surround-Produktionen zu empfehlen, sich bereits vor den Aufnahmen Gedanken über die Platzierung der Instrumente in der Mischung zu machen und im Zweifelsfall immer stereofon aufzunehmen.

Eine räumliche und gleichzeitig sehr stabile Abbildung von Schallquellen im Surroundfeld gelingt vor allem mit einer Kombination aus Laufzeitdifferenzen und Höhenabsenkung auf den hinteren Lautsprechern. Auch in dieser Frage herrschte bei den Probanden große Einigkeit. Trotz Phasenauslöschungen können je nach Verzögerung auch klanglich sehr gute Ergebnisse erzielt werden. Man sollte allerdings bedenken, dass bei anderen Abständen zwischen den Lautsprechern andere Frequenzen angehoben und abgesenkt werden, so dass dieselbe Mischung auf einer anderen Anlage nachteilig klingen kann. Daher empfiehlt es sich zusätzlich auch immer mit Pegeldifferenzen zu arbeiten.

Auch bei den Umhüllungsversuchen sind relativ klare Favoriten zu erkennen. Beste Ergebnisse erzielt man mit fünfmal eingespielten Spuren. Die gepitchten Versionen erzeugen auch ein sehr homogenes Umhüllungsgefühl, sind jedoch klanglich weniger überzeugend. Besonders bei den Gitarrenbeispielen wurde oftmals der Klang kritisiert, und als „phasig“ oder „künstlich“ bezeichnet.

Das Ergebnis des Vergleichs der zwei verschiedenen Mischungen zeigt einmal mehr, dass es in der Musikproduktion verschiedene Möglichkeiten gibt, ein ansprechendes Resultat zu erzielen.

Sämtliche Beispiele des Hörversuchs sind auf der DTS-CD im Anhang zu finden.

### 8. Ausblick

Die Fünfkanal-Stereofonie ist in Bezug auf Stabilität und Raumeindruck herkömmlichen Zweikanal-Produktionen deutlich überlegen und bietet ein gänzlich neues Hörerlebnis. Eine gute Surround-Mischung ist jedoch mit erheblich größerem Aufwand verbunden. Instrumente müssen mit mehreren Mikrofonen auf getrennte Spuren aufgenommen und/oder mehrfach eingespielt werden.

Das auf Pegeldifferenzen basierende Surround-Panpot ist für eine stabile und gute räumliche Abbildung allein nicht ausreichend. Es müssen zusätzlich Laufzeitunterschiede erzeugt und Filterungen vorgenommen werden. Dazu sind mehr Kanäle, Busse, Delays, Equalizer und größere DSP-Ressourcen notwendig. Ein „richtiges“ Surround-Panning-Tool, bei welchem Pegel, Laufzeiten und Filter direkt für den jeweiligen Kanal eingestellt werden können, wäre dabei eine große Hilfe. Soweit mir bekannt ist, existiert ein solches Tool (VSP = Virtual Surround Panning) bisher nur in einigen großen Digitalpulten der Firma Studer (D 950 S). Mehr und vor allem preisgünstigeres DVD-A/SACD-Authoring-Equipment wären ein weiterer Schritt in die richtige Richtung.

Viele Künstler, selbst international erfolgreiche, haben noch kaum die Möglichkeiten, die ihnen die Mehrkanaltechnik bietet, voll erkannt, oder machen sich zumindest derzeit noch wenig Gedanken über eigene Surround-Produktionen. Einer der Hauptgründe dafür ist sicherlich, die bisher sehr geringe Verbreitung der DVD-Audio und SACD, wie auch der entsprechenden DVD-Player. Preisgünstige Multi-Player, die alle Formate von der CD über die DVD-Video bis hin zu den hochauflösenden Audioformaten abspielen können, könnten die Akzeptanz beim Konsumenten verbessern.

Letztendlich wird sich der Konsument jedoch nur dann für die hochauflösenden Surround-Formate entscheiden, wenn die Produkte überzeugende Inhalte liefern. Es bleibt also die Forderung nach musikalisch-künstlerisch ansprechenden und qualitativ hochwertigen Produktionen, die dem Konsumenten ein neues, eindrucksvolles Hörelebnis bieten.

### Quellenverzeichnis

#### ...und weiterführende Literatur:

- Blauert, Jens, Räumliches Hören, S. Hirzel Verlag, Stuttgart 1974
- Blauert, Jens, Räumliches Hören - Nachschrift, Neue Ergebnisse und Trends seit 1972, S. Hirzel Verlag, Stuttgart 1985
- Blauert, Jens, Räumliches Hören - 2. Nachschrift, Neue Ergebnisse und Trends seit 1982, S. Hirzel Verlag, Stuttgart 1997
- Burgtorff, W./Wagener, B., Subjektiv diffuse Schallfelder, in: Acustica Nr. 19, S. Hirzel Verlag, Stuttgart 1968
- Damaske, P., Subjektive Untersuchung von Schallfeldern, Acustica Vo. 19, 1967/1968
- Dickreiter, Michael, Handbuch der Tonstudioteknik, Band 1 und 2, K. G. Saur Verlag, München 1997
- Gernemann, A., Kohärenz und Korrelation in der Tonstudioteknik, <http://www.ksdigital.de/deutsch/index.html>, Seitenaufruf vom 21.03.2005
- Görne, Thomas, Mikrofone in Theorie und Praxis, Elektor-Verlag GmbH, Aachen 1996
- Görne, Thomas, Studioteknik, Elektor-Verlag GmbH, Aachen 1996
- Griesinger, David, Spatial impression and envelopment in small rooms, <http://world.std.com/~griesngr>, Seitenaufruf vom 26.03.2005
- Griesinger, David, Speaker placement, externalization, and envelopment in home listening rooms <http://world.std.com/~griesngr>, Seitenaufruf vom 26.03.2005
- Griesinger, David, The Science of Surround, <http://world.std.com/~griesngr>, Seitenaufruf vom 26.03.2005
- Herrmann, U./Henkels, V., Vergleich fünf verschiedener Surround-Hauptmikrofonverfahren, Diplomarbeit, FH Düsseldorf, März 1997
- ITU-R BS. 775-1, Multichannel stereophonic sound system with and without accompanying picture, Rec., International Telecommunications Union, Geneva, Switzerland (1992/1994)
- ITU-R BS. 1116-1, Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems, Rec., International Telecommunications Union, Geneva, Switzerland (1994-1997)
- Martin, G., Sound Source Lokalization in a Five-Channel Surround Sound Reproduction System, <http://www.music.mcgill.ca/mmt/publications/martin99b.html>, Seitenaufruf vom 26.03.2005
- Meyer, J., Akustik und musikalische Aufführungspraxis, Verlag Erwin Bochinski, Frankfurt a. M. 1999
- Nousaine, T., Multiple subwoofers for home theatre, J. Audio Eng. Soc. (Abstracts), Vol. 45, S. 1015 (Nov. 1997), Preprint 4558.
- Rogge, K.-E., Methodenatlas, Springer-Verlag, Berlin/Heidelberg 1995

## Anhang

---

Sengpiel, Eberhard, Der Ohrabstand-Welcher,

<http://www.sengpielaudio.com/DerOhrabstand-Welcher.pdf>, Seitenaufruf vom 10.03.2005

Sengpiel, Eberhard, Die Bedeutung der Blauertschen Bänder,

<http://www.sengpielaudio.com/DieBedeutungDerBlauertschenBaender.pdf>,

Seitenaufruf vom 10.03.2005

Sengpiel, Eberhard, Kurven gleicher Hörereignisrichtungen,

<http://www.sengpielaudio.com/KurvenGleiHoerereignis.pdf>, Seitenaufruf vom 10.03.2005

Theile, Günther, Multichannel Natural Music Recording Based on Psychoacoustic Principles,

[http://www.irt.de/wittek/hauptmikrofon/theile/Multich\\_Recording.pdf](http://www.irt.de/wittek/hauptmikrofon/theile/Multich_Recording.pdf), Seitenaufruf vom 21.03.2005

Theile, Günther, Möglichkeiten und Grenzen der 3/2-Stereo-Aufnahme,

Tagungsbericht der 19. Tonmeistertagung 1996

Theile, Günther, Über die Lokalisation im überlagerten Schallfeld (Diss. 1980),

[http://www.irt.de/wittek/hauptmikrofon/theile/UEBER\\_DIE\\_LOKALISATION\\_deutsch.pdf](http://www.irt.de/wittek/hauptmikrofon/theile/UEBER_DIE_LOKALISATION_deutsch.pdf),

Seitenaufruf vom 21.03.2005

Wittek, Helmut, Untersuchungen zur Richtungsabbildung mit L-C-R Hauptmikrofonen, (Dipl. 2000),

[http://www.irt.de/wittek/hauptmikrofon/DA\\_Helmut\\_Wittek.pdf](http://www.irt.de/wittek/hauptmikrofon/DA_Helmut_Wittek.pdf), Seitenaufruf vom 21.3.2005

Zieglmeier, W./Theile, G., Darstellung seitlicher Schallquellen bei Anwendung des 3/2-Stereo Formates,

[http://www.irt.de/wittek/hauptmikrofon/theile/Seitl.Lokalisation\\_TMT1996.pdf](http://www.irt.de/wittek/hauptmikrofon/theile/Seitl.Lokalisation_TMT1996.pdf),

Seitenaufruf vom 21.03.2005

Zöfel, P., Statistik in der Praxis, Fischer Verlag, Stuttgart, 1985

Zacharov, N./Bech, S./Meares, D., The use of subwoofers in the context of surround sound program reproduction, J. Audio Eng. Soc. (Abstracts), Vol. 46, S.276–287 (Apr. 1998).