

Doppel-MS – eine Surround- Aufnahmetechnik unter der Lupe

Double M/S – a Surround recording technique put to test

Helmut Wittek*,
Christopher Haut **, Daniel Keinath ***

* Schoeps Mikrofone GmbH, wittek@schoeps.de

** IHA, Fachhochschule Oldenburg, christopher.haut@fh-oldenburg.de

*** ETI, Musikhochschule Detmold, d_keinath@onlinehome.de

Abstract

Doppel-MS ist eine Aufnahmetechnik für Zweikanal- als auch Mehrkanalstereofonie, die sich nicht zuletzt aufgrund ihrer rein praktischen Vorteile großer Beliebtheit erfreut.

Aber was steckt tatsächlich hinter Doppel-MS?

Nähere Betrachtungen in theoretischer und praktischer Richtung zeigen: Doppel-MS kann gute Ergebnisse in verschiedenen Anwendungen liefern, wenn gewisse Parameter bei der Dematrixierung/Dekodierung beachtet werden. Vergleichsaufnahmen mit den Referenzsystemen IRT-Kreuz und OCT Surround belegen, dass in verschiedenen Aufnahmesituationen eine optimale Dekodierung der Doppel-MS Signale ähnlich bzw. genauso gute Ergebnisse liefert. Zu den entscheidenden Parametern zählen unter anderem Crosstalkpegel und Lokalisation sowie die Korrelation der Lautsprecher signale. Die Dekodierung lässt sich nur dann optimieren, wenn beide Kanal-Paare L/R und LS/RS aus allen drei Mikrofon signalen gewonnen werden - dies entspricht quasi einem Mitten-Signal mit variabler Richtcharakteristik. Diese Möglichkeit macht die Doppel-MS Aufnahmetechnik auch für Zweikanal-Aufnahme zu einer interessanten Option.

In Hörversuchen wurde der maximale Crosstalkpegel ermittelt, ab dem negative Konsequenzen für Klang und Abbildung entstehen.

Das Doppel-MS System stellt sich als System heraus, das jenseits seiner traditionellen Anwendung für Dokumentation und Hörspiel seine Berechtigung und seine Anhänger hat.

Abstract

Double-M/S is a recording technique for two-channel as well as multichannel stereo. It has considerable practical advantages which justify its growing popularity.

What is really behind Double M/S?

Closer examination shows that it can provide high-quality pickup not only of ambient sound but also of music, provided that certain criteria are considered in the dematrixing/decoding of the three microphone signals. Recordings made simultaneously with the "IRT Cross" and "OCT Surround" reference systems reveal that Double M/S can achieve comparable overall quality, given optimal decoding.

The decisive parameters for the quality of decoding are crosstalk level, localization, and the degree of correlation among the loudspeaker signals. These characteristics can be optimized only if the L/R and LS/RS channel pairs are both obtained from all three of the microphone signals. In other words, the optimal decoding needs a variable M pattern generated through a mix of front- and rear-facing cardioids. This capability also makes Double M/S an interesting option for two-channel recording, similar to Ambisonics.

Experiments have revealed the maximum allowable crosstalk level above which the sound color and sound image begin to be adversely affected.

The Double M/S system turns out to be a system that is suitable for more than just the usual applications in documentary sound and radio drama.

An English version of a paper on this topic will be available on www.schoeps.de.

Gliederung

Gliederung.....	2
Einleitung.....	2
1. Das MS-Prinzip.....	3
2. Beschreibung der Doppel-MS Anordnung	4
2.1. Mikrofonanordnung.....	4
2.2. Generierung von 2/0-Stereo und 3/2-Stereo-Signalen	5
2.3. Verwandtschaft des Doppel-MS Systems mit der Ambisonics-Aufnahmetechnik	6
3. Parameter zur theoretischen Analyse der Doppel-MS Technik	8
3.1. Richtungsabbildung (Lokalisation)	9
3.2. Kohärenz / Korrelation	13
3.3. Crosstalk.....	17
4. Praktische Untersuchung anhand von verschiedenen Testaufnahmen	20
4.1. Verschiedene Varianten der Dekodierung von Doppel-MS	21
4.2. Subjektive Erfahrungen mit Doppel-MS Aufnahmen	24
5. Werkzeuge zur Dekodierung von Doppel-MS	25
6. Danksagung.....	27
7. Literatur	27
Anhang	29

Einleitung

Die Doppel-MS Aufnahmetechnik erfreut sich einiger Beliebtheit und zählt zu den etablierten Aufnahmetechniken für bestimmte Anwendungen. Trotzdem ist über ihre Eigenschaften relativ wenig bekannt. Es besteht offensichtlich Nachholbedarf sowohl an einer Beschreibung der objektiven Eigenschaften als auch an Erfahrungsberichten mit dieser Mikrofonanordnung. Auch die verschiedenen Arten der Dekodierung (De-Matrizierung) der MS-Signale sind nicht allgemein bekannt, und erst seit kurzem existieren neue Werkzeuge zur optimalen Dekodierung.

Der Autor dieses Aufsatzes ist in den vergangenen Jahren selbst manchmal dem Vorurteil zum Opfer gefallen, dass MS-Technik per se eine den anderen Aufnahmetechniken unterlegene Aufnahmetechnik sei. Dies ist nicht der Fall. Eine MS-Aufnahme ist mitnichten ein Kompromiss (vgl. [9]) und auch die Verwendung der MS-Technik für Mehrkanalaufnahme ist kein Problem, wenn man die Eigenschaften der MS-Technik kennt. Es gilt auch allgemein: Nur wer die Eigenschaften einer Anordnung kennt, kann mit deren Vorzügen und Schwächen gut umgehen, keine Aufnahmetechnik bietet von vorn herein gute Ergebnisse. Zum Glück gibt es kein „Idioten-sicheres“ Aufnahmeverfahren. Wenn Werbung solches suggeriert, ist Vorsicht geboten, denn oft kann man mit dem beworbenen Produkt nicht nur nichts falsch, sondern auch nichts richtig machen.

So kann auch die Doppel-MS Anordnung, genau wie die MS-Technik überhaupt, durch schlechte Dekodierung unbrauchbare oder zumindest schlechtere Ergebnisse liefern. Genau wie bei der behutsamen Einstellung des Öffnungswinkels zweier Mikrofone im XY-Verfahren verlangt auch die MS-Technik eine vom Tonmeister bewusst kontrollierte Dekodierung. Diese scheinbar selbstverständliche Forderung wird anhand der Ergebnisse der Untersuchungen in diesem Aufsatz demonstriert.

Die Untersuchungen dieses Aufsatzes teilen sich in theoretische und praktische Analyse. Die theoretische Analyse wird anhand wichtiger Parameter der Mikrofonanordnung wie Korrelation der Kanäle im Diffusfeld, Abbildungseigenschaften und Crosstalkverhalten durchgeführt. Durch eine Analyse anhand dieser Parameter können wesentliche Eigenschaften der Anordnung vorausgesagt werden. Diese objektive Analyse vereinfacht die Bewertung einer bestehenden Anordnung wesentlich.

Leider ist die Aussagekraft der objektiven Analyse in den meisten Fällen trotzdem beschränkt. Nicht alles lässt sich erklären und voraussagen. Zum einen liegt das daran, dass die stereofone Wahrnehmung und die Wahrnehmung räumlicher Attribute komplizierten Prozessen unterliegen, die auch heute noch nicht ausreichend geklärt sind. Zum anderen beeinflussen auch ganz individuelle Parameter die Qualität einer Aufnahme wesentlich, sodass keine definitive allgemeine Aussage getroffen werden kann. Diese individuellen Parameter umfassen z.B. die Art des Aufnahmeraums, Art der Schallquelle(n), musikalischen Inhalt und nicht zuletzt den persönlichen Geschmack des Tonmeisters/Hörers.

Deshalb ist es gefährlich, rein theoretische Analysen durchzuführen und unabhängig von praktischer Überprüfung und Erfahrung Regeln für optimale Aufnahmen zu erstellen. Dies gilt insbesondere für die koinzidente Aufnahmetechnik, zu der (auch in Zusammenhang mit dem Ambisonics-Wiedergabeverfahren) reichlich theoretische Abhandlungen existieren, die deren Überlegenheit postulieren (z.B. [7], [15]). In solchen Untersuchungen wird allerdings teilweise von Annahmen ausgegangen, die einer kritischen Überprüfung bedürfen. Folglich sind auch die Schlussfolgerungen meist nur vom Thema Richtungsdarstellung bestimmt und vernachlässigen andere, wichtige Parameter. In diesem Aufsatz wird versucht, die koinzidente Aufnahmetechnik speziell für die Anwendung Surround Sound einer kritischen und wertfreien objektiven sowie subjektiven Analyse zuzuführen.

1. Das MS-Prinzip

Die Signale einer koinzidenten Mikrofonanordnung lassen sich nach dem MS-Prinzip (**M**id/**S**ide) matrizen, in dem deren Summe und Differenz gebildet werden. Das Mittel-Signal entspricht der Mono-Summe, das Seiten-Signal der Differenz.

Die De-Matrizierung (oder Dekodierung) wird erreicht, in dem aus Summe und Differenz die ursprünglichen Signale wieder gewonnen werden.

$$M = L + R$$

$$S = L - R$$

$$L = \frac{1}{2} * (M + S)$$

$$R = \frac{1}{2} * (M - S)$$

Ein Vorteil der MS-Kodierung liegt darin, dass Aufnahmewinkel bzw. Stereobreite direkt durch einfache Pegelung der MS-Signale zu variieren sind.

*Abbildung 1:
Beispiel für eine typische Anwendung der MS-Stereofonie:
Rohrrichtmikrofon mit aufgesetztem
Acht-Mikrofon für Montage im Windkorb.
aus [20]*



Bei der MS-Aufnahme wird direkt in matrizzierter Form aufgenommen, das heißt es wird ein Mikrofon für den M-Kanal und ein Mikrofon für den S-Kanal verwendet. Das M-Mikrofon ist nach vorne ausgerichtet, das S-Mikrofon ist eine quer liegende Acht.

Jede MS-Anordnung lässt sich in eine äquivalente XY-Anordnung umrechnen. Abbildung 3 zeigt verschiedene Kombinationen von MS-Signalen (M=Niere) in der oberen Reihe, mit der dazu gehörenden XY-Anordnung in der unteren Reihe.

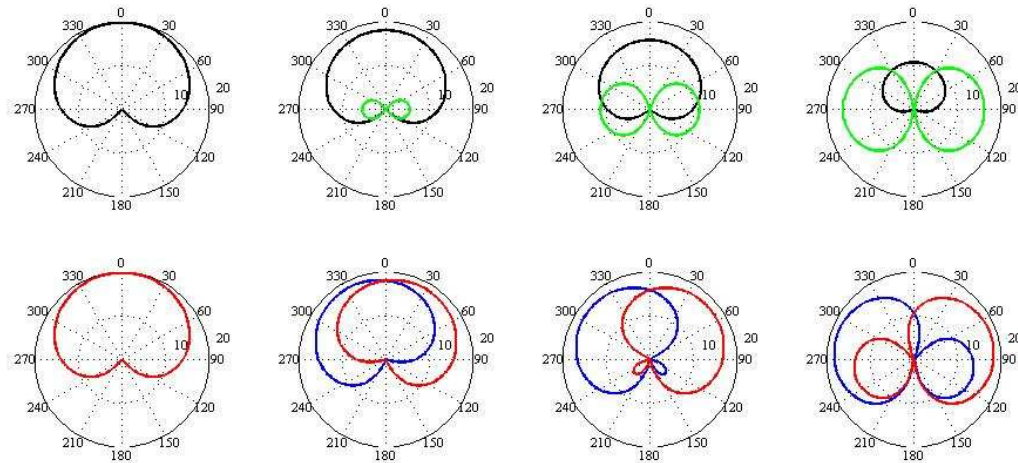


Abbildung 2: Illustration des MS-Prinzips:
 Reihe oben: M-Signal (Schwarz) und S-Signal (grün)
 Reihe unten: sich daraus ergebende Signale L (blau) und R (rot)

2. Beschreibung der Doppel-MS Anordnung

2.1. Mikrofonanordnung

Das Doppel-MS Verfahren ist eine Aufnahmetechnik für Zwei- und Mehrkanalstereofonie, die ausschließlich auf Pegeldifferenzen beruht.

Die Idee hinter der Doppel-MS Anordnung ist die Bündelung zweier MS-Mikrofonpaare. Dadurch kann ein Mikrofon eingespart werden, da die Acht für beide Paare verwendet wird. Das Doppel-MS System besteht somit aus drei Mikrofonen.

Abbildung 3 illustriert dieses Prinzip:

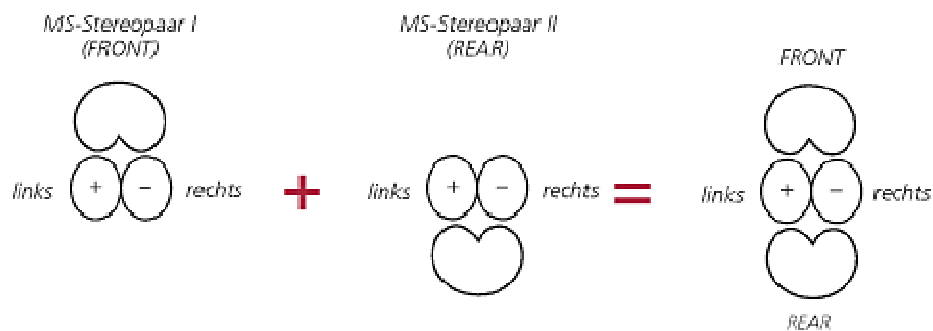


Abbildung 3: Prinzip der Doppel-MS Anordnung: Bündelung zweier MS-Paare, aus [20]

Die drei Mikrofone/Kanäle werden folgendermaßen benannt:

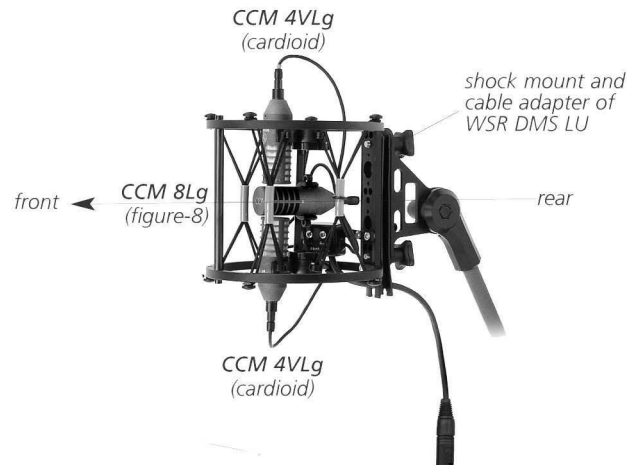
- M_{front}
- S
- M_{rear}

Durch die Verwendung von nur drei (Kleinmembran-)Mikrofonen kann erreicht werden, die Mikrofonmembrane quasi koinzident, also an einem Punkt anzubringen.

Abbildung 4:

Doppel-MS Anordnung

Die Verwendung von Kompaktmikrofonen (SCHOEPS CCM4V und CCM8) bewirkt einen optimal kleinen Abstand der drei Membranen. aus [20]



2.2. Generierung von 2/0-Stereo und 3/2-Stereo-Signalen

Aus diesen drei Mikrofonsignalen können nun die Signale sowohl für Zweikanal- als auch Mehrkanal-Stereofonie generiert werden.

Dies kann prinzipiell durch eine konventionelle MS-Matrix geschehen, wobei eine Matrix zur Generierung der Signale L/R aus M_{front}/S verwendet wird, und eine weitere Matrix zur Generierung der Signale LS/RS aus M_{rear}/S . Außerdem wird der Center optional mit dem Signal des M_{front} -Mikrofons gespeist (vgl. Abschnitt 5).

Die Doppel-MS Anordnung macht jedoch eine weitaus bessere Dekodierung möglich: Der Nachteil einer MS-Anordnung liegt schließlich darin, dass die Richtcharakteristik der dematrizierten virtuellen Mikrofone vom Mischungsverhältnis zwischen M und S abhängt. Dies ist in Abbildung 5 (oben) veranschaulicht. Mit zunehmendem Pegel des S-Signals verschiebt sich die resultierende Richtcharakteristik von der Niere zur Acht. Dadurch, dass sich in einer Doppel-MS Anordnung aus einer Mischung der nach vorne und der nach hinten gerichteten Nieren jede beliebige Richtcharakteristik 1. Ordnung erstellen lässt, kann die Richtcharakteristik des dekodierten virtuellen Mikrofons nun frei bestimmt werden (Abbildung 5 unten).

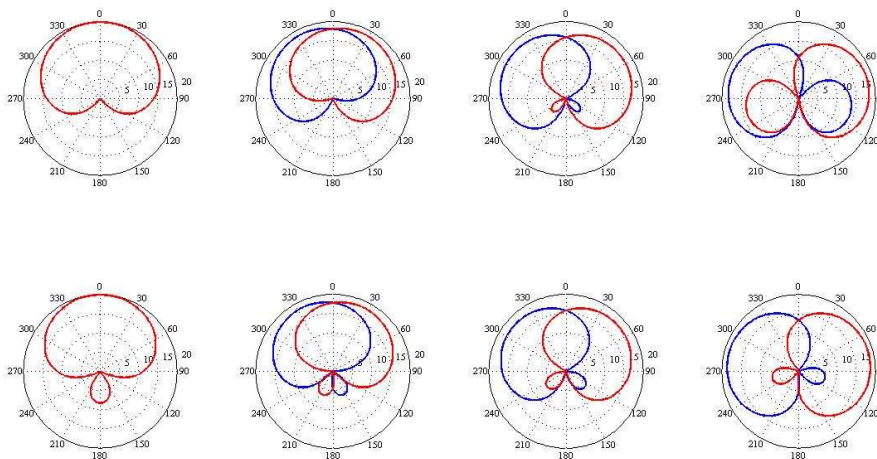


Abbildung 5:

Veranschaulichung des Unterschieds zwischen einer Dekodierung mit/ohne variable Richtcharakteristik des M_{front} -Signals. Abgebildet sind die resultierenden Polardiagramme

der dekodierten Kanäle L,R
 Reihe oben: mit fester Richtcharakteristik $M_{front}=Niere$
 Reihe unten: mit variabler Richtcharakteristik M_{front} , Vorgabe: L,R=Superniere.

Dieser Vorteil ist von wesentlicher Bedeutung für eine optimale koinzidente Aufnahme. Nur so ist eine Steuerung des Aufnahmewinkels möglich, ohne gleichzeitig Richtwirkung und Korrelation des resultierenden virtuellen Mikrofonpaars zu beeinflussen. Außerdem kann nur auf diese Weise die Unabhängigkeit (Dekorrelation) der Signale maximal werden – eine Forderung, die beim Doppel-MS System wichtig ist. Die Bedeutung dieser Parameter wird im Folgenden noch diskutiert. Praktische Werkzeuge zur Dekodierung von Doppel-MS werden in Abschnitt 5 beschrieben.

2.3. Verwandtschaft des Doppel-MS Systems mit der Ambisonics-Aufnahmetechnik

Ambisonics ist eine Aufnahme- und Wiedergabetechnik, die auf einer Erfindung von Michael Gerzon [7] beruht. Sie geht von einer koinzidenten Aufnahme aus. Grundlage der Ambisonics-Theorie ist die Zerlegung des Schallfelds in so genannte „spherical harmonics“ (Kugelflächenfunktionen), das sind Funktionen, die die Richtung des eintreffenden Schalls beschreiben. Die Kugelflächenfunktionen beschreiben das Schallfeld umso genauer, je höher ihre Ordnung wird.

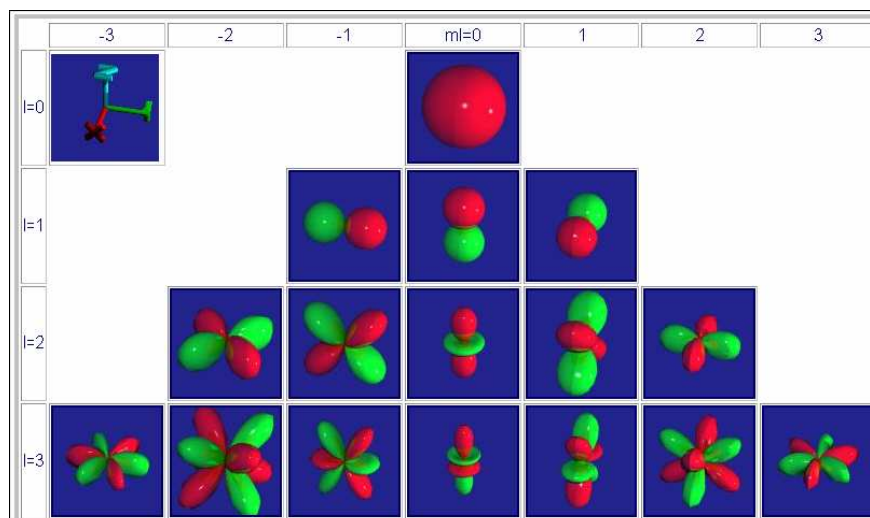


Abbildung 6: Visualisierung der Kugelflächenfunktionen bis zur dritten Ordnung ; l entspricht der Ordnung, ml der Dimension. aus [35]

In Abbildung 6 sind die Kugelflächenfunktionen bis zur dritten Ordnung abgebildet. Bisher ist mit konventionellen Mitteln nur die Mikrofonaufnahme der Kugelflächenfunktionen erster Ordnung möglich. Diese Aufnahme nach „Ambisonics erster Ordnung“ liefert folgende Signale, die auch „B-Format erster Ordnung“ genannt werden:

B-Format erster Ordnung:

0-te Ordnung:

$$W = 1;$$

1-te Ordnung:

$$X = \cos(\Theta) * \cos(\phi);$$

$$Y = \sin(\Theta) * \cos(\phi);$$

$$Z = \sin(\phi);$$

Θ beschreibt den Winkel in der x/y-Ebene (Horizontalebene $z=0$), ϕ beschreibt den Winkel in der z-Ebene.

Diese 4 Signale können auf unterschiedliche Weise durch eine Mikrofonaufnahme gewonnen werden:

Nach Gerzon lassen sich mit 4 Mikrofonen, die in einem gleichseitigen Tetraeder angebracht sind, die 4 Signale des B-Formats ermitteln. Solche Mikrofone existieren auch auf dem Markt, z.B. das „Soundfield“-Mikrofon der englischen Firma Soundfield (siehe [24]). Vorteil dieser Methode ist die gleichmäßige Anordnung im Tetraeder und die gute Koinzidenz der Anordnung in allen Raumrichtungen. Ein Nachteil ist die Notwendigkeit, zunächst die Tetraedersignale („A-Format“) ins B-Format umwandeln zu müssen.

Eine andere Methode ist eine Mikrofonaufnahme, die direkt das B-Format liefert. Dazu werden eine Kugel (W) und drei Acht-Mikrofone (X, Y, Z) verwendet. Dies wird oft mit „native B-format Recording“ bezeichnet (eine solche Anordnung ist z.B. in [3] abgebildet). Lässt man nun das dritte Acht-Mikrofon weg, weil man nicht an einer dreidimensionalen Wiedergabe interessiert ist, ergibt sich eine Anordnung, die geometrisch relativ einfach zu realisieren ist. Benjamin [1] bezeichnet das sich ergebende Format als „Horizontal B-Format“. Mit hinreichend kleinen Mikrofonen lässt sich eine gute Koinzidenz, in Bezug auf die Horizontalebene perfekte Koinzidenz erreichen.

Die Doppel-MS Anordnung liefert nun prinzipiell dieselben Signale, da sich die Mikrofon-signale ins „Horizontal B-Format“ durch einfache Summierung und Differenzbildung überführen lassen:

$$W = M_{\text{front}} + M_{\text{rear}};$$

$$X = M_{\text{front}} - M_{\text{rear}};$$

$$Y = S;$$

Es zeigt sich also, dass die Doppel-MS Signale mit einer Aufnahme nach Ambisonics erster Ordnung bis auf die fehlende Höheninformation (Z) identisch sind. Dieser Punkt spielt dann keine Rolle, wenn eine Wiedergabe auf einem konventionellen Lautsprecher-system ohne Höheninformation (2.0, 5.1, etc.) beabsichtigt ist.

Bei Benjamin [1] wird ein Vergleich zwischen den beiden Aufnahme-Methoden ange-stellt. Der Vergleich der sich ergebenden Polardiagramme des B-Formats ergibt, dass ein natives Array aus einzelnen Kapseln (in der Untersuchung Benjamins waren es Schoeps MK 2 und MK 8) die besseren Polardiagramme im B-Format bietet, wohl aber für Schalleinfall aus einer anderen Richtung als der Horizontalebene nicht die optimalen Polardiagramme behalten kann. Das Tetraeder-Mikrofonensystem zeigt gute, konsistente Polardiagramme, die unabhängig von der Schalleinfallsebene sind, jedoch unregelmäßi-gen Abweichungen ab ca. 6 kHz haben. Details liefert z.B. Flock [6].

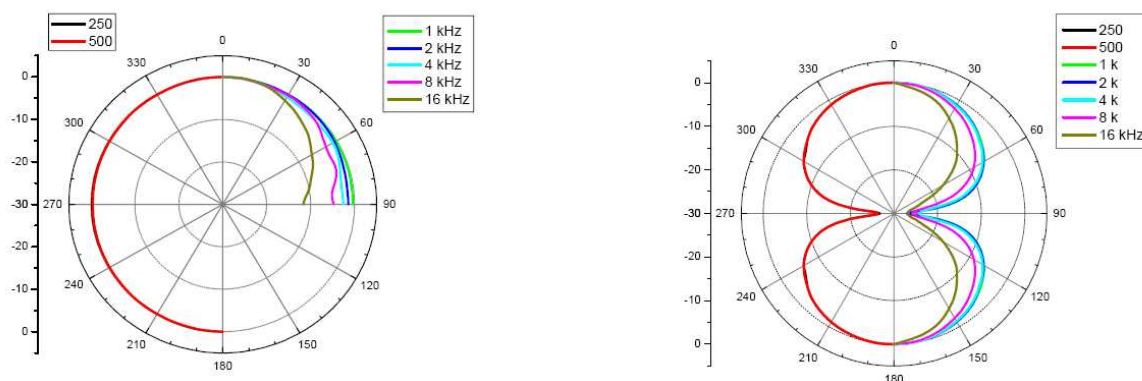


Abbildung 7: Polardiagramme aus einem "Horizontal B-Format"-Array aus Benjamin [1]: links: SCHOEPS MK 2 (=W-Signal, bei einer 90°-Stellung der Kugel wäre das Polardiagramm in der horizontalen Ebene sogar perfekt frequenzunabhängig) rechts: SCHOEPS MK 8 (=X-Signal)

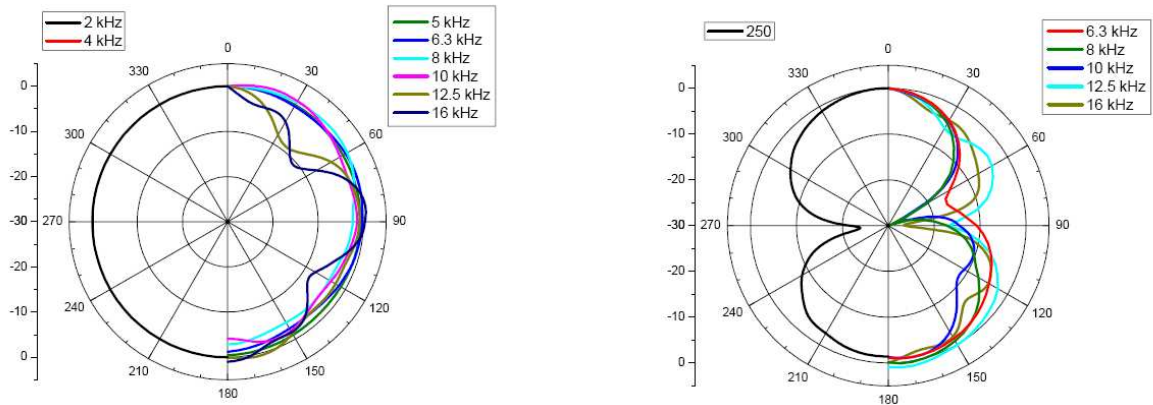


Abbildung 8: Polardiagramme aus einem Tetraeder-Array (Soundfield-Mikrofon MkV Mikrofonsystem) aus Benjamin [1]:
links: W-Signal
rechts: X-Signal
Achtung: Es sind andere Frequenzen als in Abbildung 7 verwendet worden, die Graphen sind nicht direkt vergleichbar.

Wie oben erwähnt, ist die Grundlage der Ambisonics-Theorie die Zerlegung des Schallfelds in mehrere Richtungskomponenten. Bei der Wiedergabe wird das Schallfeld also durch Überlagerung aller Lautsprecher signale wieder zusammengesetzt. Dabei können durchaus mehr als zwei Lautsprecher korrelierte Signale haben. Die Zerlegung des Schallfelds folgt nicht dem Gesetz, dass eine Phantomschallquelle aus Pegel- und Zeitdifferenzen zwischen zwei benachbarten Lautsprechern gebildet wird, sondern visiert eine physikalische Zusammensetzung des Signals im Sweet Spot an. Dadurch ergeben sich auch andere Gesetzmäßigkeiten bezüglich der in diesem Aufsatz besprochenen Parameter. Da der Ansatz der Ambisonics-Theorie zur Wiedergabe eines Surround-Signals so verschieden von der Phantomschallquellentheorie ist, kann er auch nicht direkt theoretisch verglichen werden. Würde man mit den in diesem Aufsatz genannten Parametern ein Ambisonics-Signal erster Ordnung beurteilen, käme man oft auf ein negatives Resultat. Tatsächlich gibt es viele Tonmeister, die auf diese Weise entstandene Mischungen ablehnen. Es sei jedoch ausdrücklich darauf hingewiesen, dass dies dann nicht an der Koinzidenz der Aufnahme liegen muss (wie es oft behauptet wird). Es gibt sehr verschiedene koinzidente Aufnahmetechniken und es gibt auch verschiedene Arten, die Qualität zu beurteilen und entsprechend zu optimieren.

3. Parameter zur theoretischen Analyse der Doppel-MS Technik

In diesem Kapitel werden verschiedene Parameter beschrieben, die bei der objektiven Bewertung der Doppel-MS Technik, sei es für Zweikanal-Stereo oder Surround, von Bedeutung sind. Sie sind speziell für die Diskussion des Doppel-MS Setups zusammengestellt, haben aber auch zur Diskussion anderer Anordnungen Geltung.

Die Parameter sind:

- Pegel- und Laufzeitdifferenzen zur Richtungsabbildung
- Korrelation
- Crosstalk (Übersprechverhalten)

Diese Parameter haben Einfluss auf verschiedene Attribute der Wahrnehmung wie Lokalisation, Klangfarbe und räumliche Wahrnehmung.

Die Relevanz der verschiedenen Parameter hängt von der jeweiligen Anwendung ab. Es sei darauf hingewiesen, dass nie ein Parameter allein beachtet werden sollte, auch wenn dieser noch so wichtig ist.

3.1. Richtungsabbildung (Lokalisation)

3.1.1. Allgemeine Beschreibung

Dieser Parameter beschreibt die Fähigkeit einer Mikrofonanordnung, das aufgenommene Schallfeld im Sinne der Absicht des Tonmeisters zwischen den Lautsprechern zu verteilen. Oft ist es gewünscht, dass das klangliche Geschehen im Aufnahme-raum proportional im Wiedergaberaum abgebildet wird. In diesem Fall ist der so genannte Aufnahme- winkel von Bedeutung. Der Aufnahmewinkel beschreibt denjenigen Winkel im Aufnahme- raum, der zwischen den vorderen Lautsprechern (L/C/R) abgebildet wird. Für eine detaillierte Beschreibung von Richtungsabbildung und Aufnahmewinkel sei auf [30], [31] verwiesen, im Folgenden wird diese Problematik nur kurz erläutert.

Die Auslenkung der Phantomschallquellen wird dadurch erreicht, dass zwischen den Mikrofonensignalen Pegel- und/oder Laufzeitdifferenzen bestehen, die die Phantomschall- quelle aus der Mitte zwischen zwei Lautsprechern auslenken. Theile erklärt, wie sich diese beiden Arten von Signaldifferenzen addieren. Die Gesamtauslenkung der Phan- tomschallquelle ergibt sich aus der Summe der jeweils durch die Pegel- oder Laufzeitdif- ferenz allein sich ergebenden Auslenkung.

$$\Phi_{\text{gesamt}} = \Phi_L + \Phi_t$$

siehe Theile: [25], [26]

Diese lineare Addition gilt allerdings nur in einem Bereich der Phantomschallquellenaus- lenkung bis ca. 50% der Maximalauslenkung. Danach ergibt sich eine allmähliche „Sät- tigung“ bis hin zur Abbildung in nur einem Lautsprecher. Der Autor dieses Aufsatzes beschreibt diesen Verlauf als mathematische Funktion ([30], [31], [32], [33], [34]). Diese Näherungsfunktion ist in Abbildung 9 gezeichnet. Zu beachten ist, dass die Aus- lenkung der Phantomschallquelle proportional zur Lautsprecher-Basisbreite ist. Die Auslenkung wird deshalb in % angegeben, um die Allgemeingültigkeit zu behalten. Im normalen Stereodreieck entspricht eine Auslenkung von +/- 100% der Auslenkung von +/-30°. Eine Auslenkung von 100% entspricht einer Abbildung in nur einem Lautspe- cher.

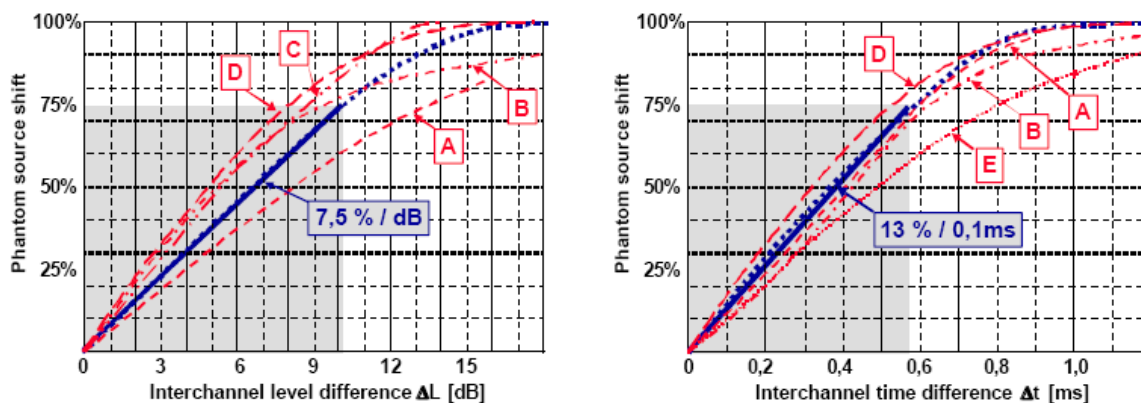


Abbildung 9: Zusammenhang zwischen Pegeldifferenz (links) bzw. Laufzeitdifferenz (rechts) und Auslenkung der Phantomschallquelle nach:

fettgedruckte Kurve: Wittek [34], 7,5% / dB bzw. 13% / 0,1 ms

A: Leakey [16]

B: Mertens [18]

C: Brittain and Leakey [2]

D: Simonson [23], Basis für die „Williams-Kurven“ [28]

E: Sengpiel [22]

Mithilfe dieser Näherungsfunktion lässt sich die resultierende Abbildung jeder stereofo- nen Anordnung aus zwei Mikrofonen berechnen. Dies ist im *Image Assistant* realisiert,

einem als JAVA-Applet ausgeführten Simulationsprogramm zu Berechnung von Lokalisationskurven. Dieses Applet kann online benützt werden (www.hauptmikrofon.de, siehe Abbildung 10, siehe Wittek [33]).

Eine Lokalisationskurve beschreibt die Auslenkung der Phantomschallquellen in Abhängigkeit vom Winkel der Schallquelle im Aufnahmebereich. Die Hauptgrafik im *Image Assistant* zeigt den Schallquellenwinkel in Grad auf der x-Achse und die Auslenkung zwischen den vorderen Lautsprechern in % auf der y-Achse an. Der Aufnahmewinkel lässt sich ablesen, indem, laut Definition, der Winkelbereich gesucht wird, in dem die Schallquellen zwischen den Lautsprechern abgebildet werden. Dieser Bereich ist hellblau eingefärbt, der Aufnahmewinkel (Recording Angle/ Recording Angle 75%) wird außerdem im Fenster in der Hauptgrafik links oben eingeblendet.

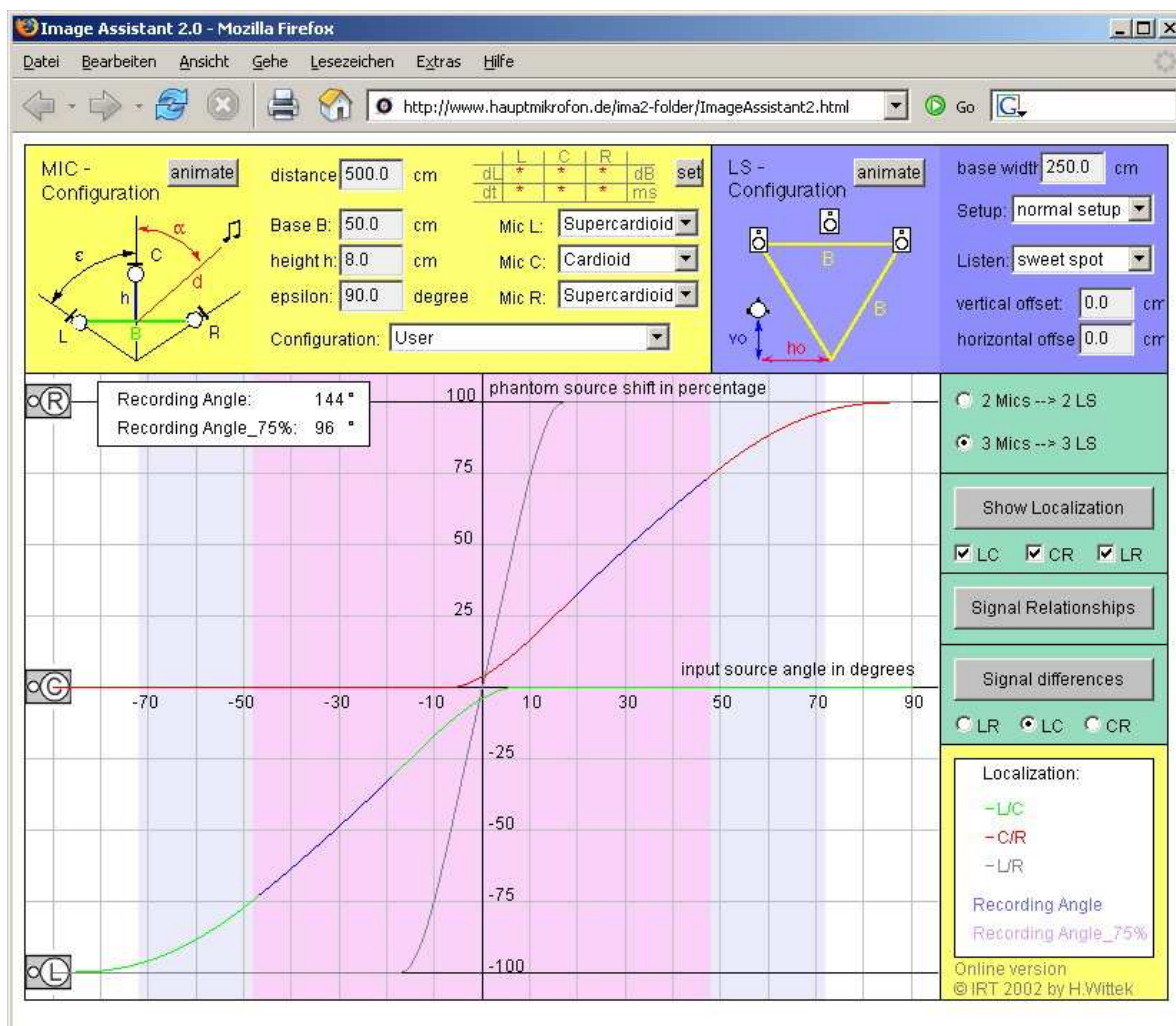


Abbildung 10: Simulation mit dem "Image Assistant" [33]: Lokalisationskurve der OCT-Anordnung.

3.1.2. Vergleich der Richtungsabbildungseigenschaften von stereofonen Anordnungen

Viele Autoren haben das Design einer Mikrofonanordnung wesentlich am Parameter Richtungsabbildung ausgerichtet. Das liegt nicht nur daran, dass dieser Parameter als wesentlich eingestuft wird. Vielmehr kann ein daraufhin optimiertes Design auch im Hinblick auf andere Parameter gut geeignet sein. Sicherlich ist es ratsam, bei der Suche nach geeigneten Kombinationen aber den Blick fürs Ganze nicht zu vergessen.

Die Vorschläge von Williams (MMA, siehe [27] und [29]) sind absichtlich in der Hauptsache auf eine 360°-Rundum-Abbildung des Schallfelds ausgerichtet. Nach Williams werden auch andere wesentliche Parameter wie die räumliche Abbildung wesentlich von der Fähigkeit zur Richtungsabbildung beeinflusst.

Auch Theile argumentiert ähnlich, allerdings empfiehlt er ausdrücklich bestimmte Kombinationen, um dafür Sorge zu tragen, dass neben der Richtungsabbildung andere Parameter optimal gestaltet werden. Ein Beispiel dafür ist die OCT-Anordnung (siehe Abbildung 11, [25], [20]), die mehrere Vorteile in sich vereint.

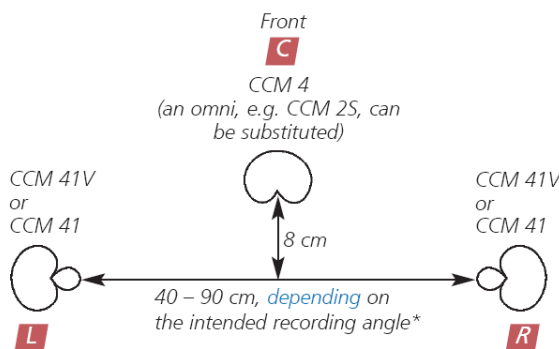


Abbildung 11: OCT-Anordnung nach Theile [25], aus [20]

Zum einen verläuft die Lokalisationskurve, die in Abbildung 10 gezeigt ist, sehr gerade und bietet somit eine sehr naturgetreue Richtungsabbildung ohne geometrische Verzerrungen. Zum anderen ist der Crosstalk zwischen den Abbildungsbereichen bei einer die OCT-Anordnung sehr gering, was Vorteile für die räumliche Abbildung, die Robustheit der Abbildung sowie die Klangfarbe mit sich bringt. Dieser Aspekt wird in Abschnitt 3.3 detailliert beschrieben.

Die Abbildungseigenschaften der Doppel-MS Anordnung sind abhängig von der Dekodierung der Signale. Dies ist einleuchtend, wenn man bedenkt, dass bei der Dekodierung sämtliche Varianten einer koinzidenten Anordnung aus bis zu 5 oder sogar mehr Mikrofonen möglich sind. 5 Mikrofone erster Ordnung können gar nicht hinreichend unterschiedliche Signale liefern, um sowohl klanglich als auch in Bezug auf die Richtungsabbildung optimale Ergebnisse zu liefern. Deshalb ist Vorsicht geboten, und oftmals ist die 4-kanalige Dekodierung einer 5-kanaligen klanglich und in Bezug auf eine 360°-Rundumabbildung überlegen.

Abbildung 12 zeigt die resultierende Richtungsabbildung einer 4-Kanal-Variante. Eine für MS/XY-Aufnahme typische, sehr lineare, das heißt proportionale Abbildung zwischen den Lautsprechern wird erreicht.

Die Richtungsabbildung einer 5-Kanal-Variante wird in Abbildung 13 gezeigt. In der Simulation ist deutlich zu erkennen, dass die theoretisch gewünschte, gleichmäßige Abbildung zwischen jeweils benachbarten Lautsprechern nicht mehr funktioniert. Der Grund dafür ist die schlechtere Kanaltrennung (durch Übersprechen, Crosstalk, siehe Abschnitt 3.3). Die Simulation des *Image Assistant* zeigt, dass im mittleren Bereich eine Abbildung zwischen drei Lautsprecherpaaren existiert. Diese theoretische Mehrfachabbildung äußert sich in reduzierter Abbildungsschärfe und schlechterer Lokalisiertheit der Phantomschallquelle. Dies muss nicht unbedingt schlechte Eigenschaften der Aufnahme zur Folge haben, denn Crosstalk existiert auch bei anderen Aufnahmetechniken. Jedoch zeigen die anderen Abschnitte dieses Aufsatzes weitere Nachteile von Crosstalk auf.

Hier existiert ein tatsächlicher Widerspruch zwischen zwei verschiedenen Theorien der Lokalisation: Die Ambisonics-Theorie erlaubt eine beliebige Anzahl Lautsprecher und betrachtet nur die überlagerten Signale im „Sweet Spot“. Auf der anderen Seite ist bei einer Phantomschallquellenbildung (siehe Theile [25]) ein korreliertes Signal auf mehr

als zwei Lautsprechern negativ für Klangfarbe und Lokalisation. Würde man Lokalisation nach der Ambisonics-Theorie berechnen wollen, müsste man die Signale aller Lautsprecher gleichzeitig auswerten und nicht nur, wie im *Image Assistant*, die bestimmter Lautsprecherpaare. Crosstalk ist kein negativer Parameter im Sinne der Ambisonics-Theorie.

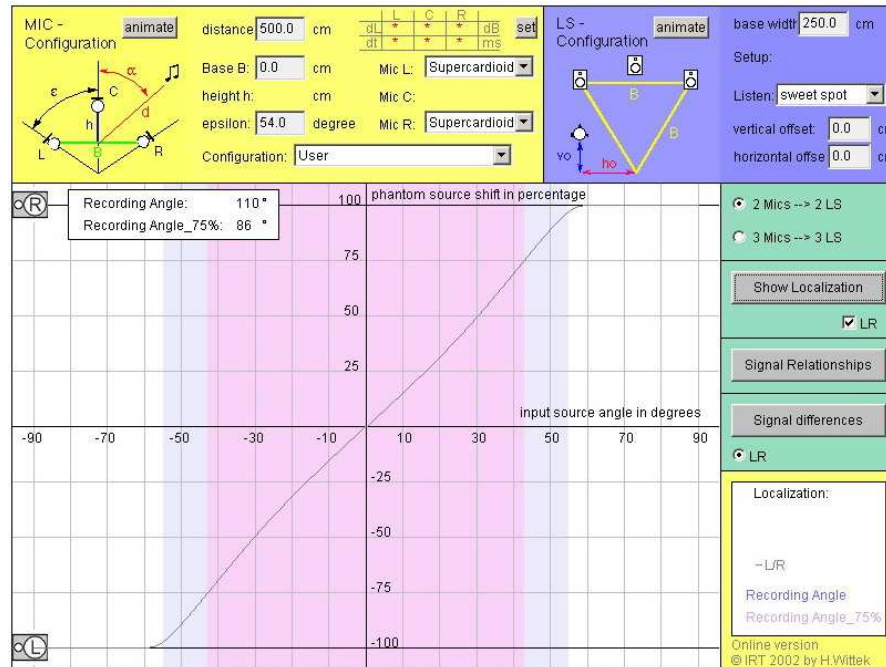


Abbildung 12: Abbildung zwischen den Lautsprechern L/C/R bei einer 4-Kanal-DMS-Dekodierung (=4ch Setting der Schoeps Hardware Matrix, siehe Abschnitt 5), (Simulation mit dem "Image Assistant" [33])

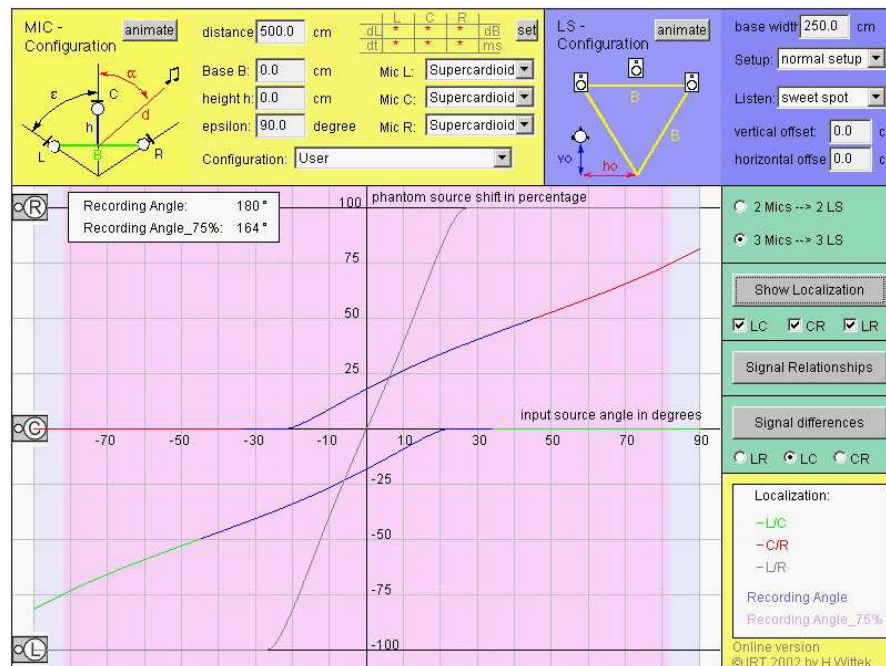


Abbildung 13: Abbildung zwischen den Lautsprechern L/C/R bei einer 5-Kanal-DMS-Dekodierung, (Simulation mit dem "Image Assistant" [33])

3.2. Kohärenz / Korrelation

3.2.1. Die Korrelation und ihre Aussagekraft bei MS-Stereofonie

Die Kohärenz bzw. die Korrelation¹ im Diffusfeld (=im Nachhall) zwischen den Kanälen einer stereofonen Aufnahme wird oft als Parameter angesehen, der auf Attribute wie räumliche Wahrnehmung und Klangfarbe einen wichtigen Einfluss hat (siehe [5], [8], [17], [19], [25]). Die Diffusfeld-Korrelation wird als ein entscheidender Parameter für die Unterschiede in der räumlichen Wahrnehmung zwischen verschiedenen stereofonen Anordnungen angenommen (vgl. Abbildungen 30 bis 33). So sind Anordnungen, die Abstände zwischen den Mikrofonen enthalten, eher für eine gute räumliche Abbildung bekannt (siehe z.B. [36]).

Gerade bei der Thematik MS-Aufnahme sollte der Parameter „Korrelation zwischen den Kanälen“ (interchannel correlation) im Diffusfeld erhöhte Aufmerksamkeit bekommen, da er neben der Richtungsabbildung entscheidend für klangliche Unterschiede zwischen den Dekodierungen ist. Diese beiden Parameter sind allerdings bei einer koinzidenten Aufnahme nicht unabhängig, sondern eng miteinander verbunden. Vergrößert man den Aufnahmewinkel, steigt auch die Korrelation. In der Auswahl der Richtcharakteristik der Mikrofone beeinflusst man die Korrelation und das Abbildungsverhalten. In Abbildung 14 wird der Korrelationsgrad einer koinzidenten Mikrofonanordnung im Diffusfeld in Abhängigkeit vom Öffnungswinkel² angegeben. Der Korrelationsgrad ist dabei abhängig von der Richtcharakteristik der beiden Mikrofone. Abzulesen ist unter anderem, dass der Diffusfeld-Korrelationsgrad zweier koinzidenter Nieren niemals unter 0.5 sinken kann. Ist ein Diffusfeld-Korrelationsgrad von 0 gewünscht, kann dies nur mit Richtcharakteristika erreicht werden, die (in etwa) maximal den Kugelanteil einer Superniere haben (vgl. auch Griesinger [11]).

Abbildung 14 zeigt, dass mit Vergrößern des Öffnungswinkels die Korrelation im Diffusfeld sinkt. Es muss allerdings beachtet werden, dass gleichzeitig der Aufnahmewinkel kleiner wird. Die folgende Tabelle zeigt die Korrelation und den Aufnahmewinkel dreier Anordnungen mit gleichem Öffnungswinkel, aber verschiedener Richtcharakteristika.

Öffnungswinkel=90°	Nieren	Supernieren	Achten
Korrelationsgrad	0.75	0.49	0.00
Aufnahmewinkel, (Aufnahmewinkel 75% [31])	180° (142°)	130° (104°)	72° (58°)

¹ Zur Klärung der Begriffe und Definition von Korrelation und Kohärenz ist im Anhang eine ausführliche Diskussion enthalten. Die Kohärenz zweier Kanäle ist ein Maß für die Ähnlichkeit der Signale im Frequenzbereich, ohne Rücksicht auf die Phasenlage. Die Korrelation zwischen zwei Kanälen ist ein Maß für die Ähnlichkeit der Signale im Zeitbereich.

² Der Begriff Öffnungswinkel bezeichnet den Winkel, den zwei Mikrofone einschließen. Es existieren auch andere Begriffe für diesen Winkel, u.a. Hauptachsenwinkel, Versatzwinkel, etc.

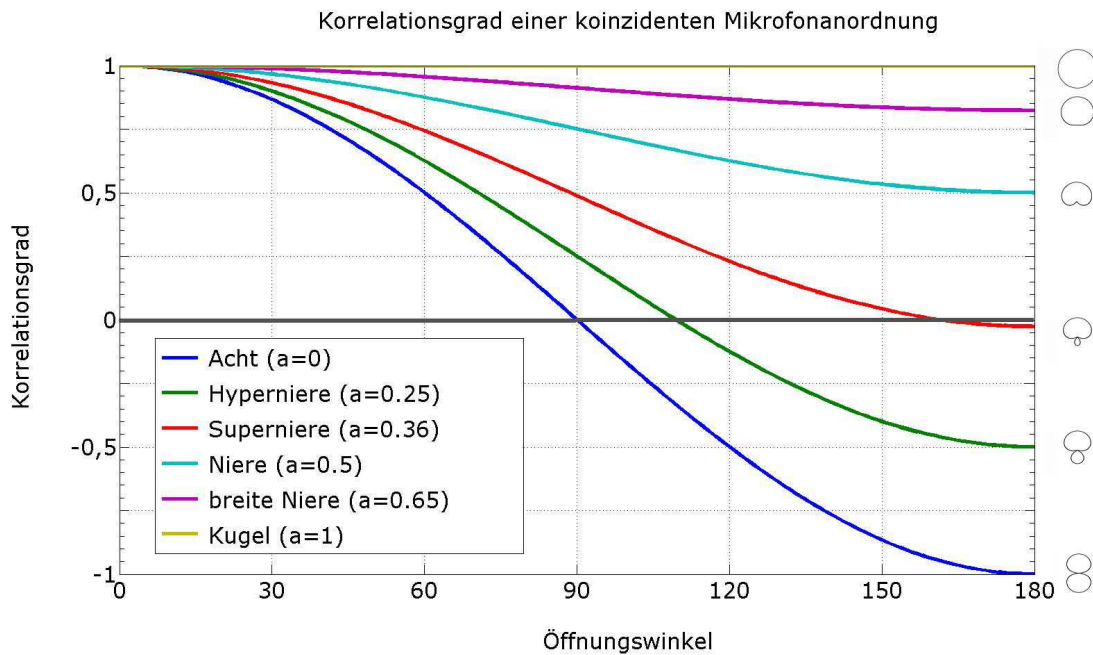


Abbildung 14: Korrelationsgrad einer koinzidenten Mikrofonanordnung im Diffusfeld in Abhängigkeit vom Öffnungswinkel der Mikrofonanordnung für verschiedene Richtcharakteristika. Angegeben ist auch der Kugelanteil des Mikrofons nach der Mikrofonformel: $sens = a + (1-a) * \cos(\phi)$

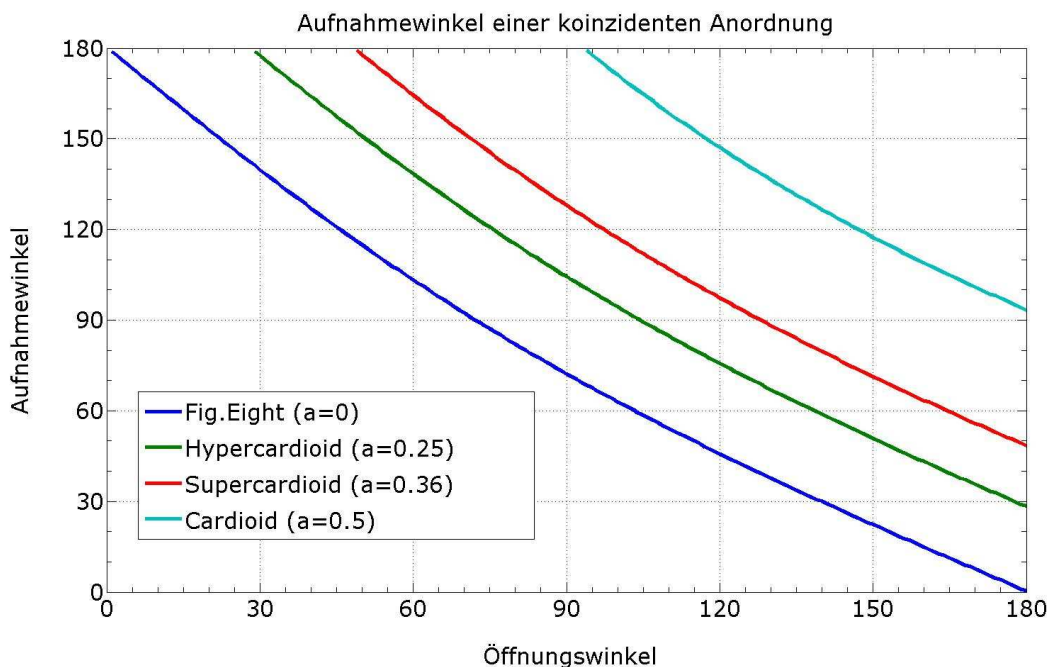


Abbildung 15: Aufnahmewinkel einer koinzidenten Mikrofonanordnung in Abhängigkeit vom Öffnungswinkel zwischen den beiden Mikrofonen für verschiedene Richtcharakteristika. Es existiert kein Aufnahmewinkel für eine koinzidente Anordnung aus breiten Nieren oder Kugeln.

Interessant wäre es, die Korrelation der Anordnung in Abhängigkeit vom resultierenden Aufnahmewinkel anzugeben. Dann kann beobachtet werden, welche Mikrofonanordnung bei gegebenem Aufnahmewinkel die optimale Dekorrelation im Diffusfeld erreicht. Dies ist möglich, wenn zunächst der Aufnahmewinkel in Abhängigkeit von Öffnungswinkel und Richtcharakteristik berechnet wird. Das Ergebnis dieser Berechnung zeigt Abbildung 15.

Die Berechnung wurde folgendermaßen durchgeführt: Der Aufnahmewinkel wird als der doppelte kleinste Versatzwinkel definiert, bei dem ein Pegelunterschied von mind. 16dB zwischen den beiden Mikrofonsignalen auftritt.

Mithilfe dieser Werte kann nun der Korrelationsgrad einer koinzidenten Anordnung aus zwei Mikrofonen beliebiger Richtcharakteristik in Abhängigkeit vom resultierenden Aufnahmewinkel angegeben werden:

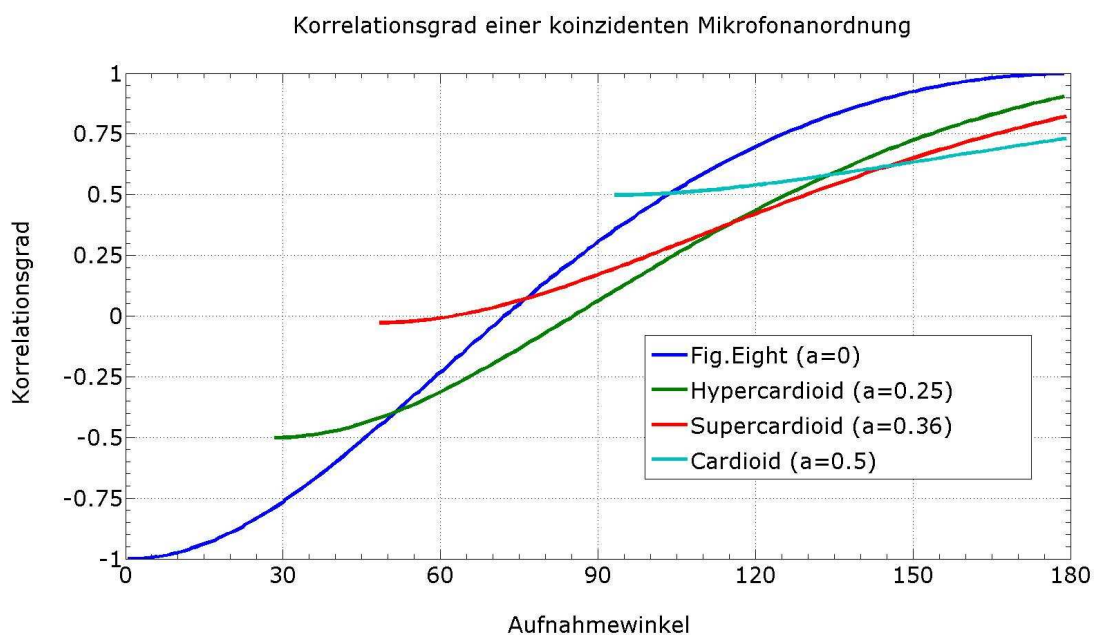


Abbildung 16: Korrelationsgrad einer koinzidenten Mikrofonanordnung im Diffusfeld in Abhängigkeit vom Aufnahmewinkel der Mikrofonanordnung für verschiedene Richtcharakteristika.

Es existiert kein Aufnahmewinkel für eine koinzidente Anordnung aus breiten Nieren oder Kugeln.

Abbildung 16 zeigt, dass der Diffusfeld-Korrelationsgrad einer koinzidenten Mikrofonanordnung mit festem Aufnahmewinkel durchaus von der Richtcharakteristik der Mikrofone abhängt. Allerdings ist man beschränkt bei der Auswahl der Richtcharakteristik, da durchaus nicht alle in Abbildung 16 angegebenen Werte eine sinnvolle Anordnung ergeben und schließlich die Unterschiede in der Korrelation oft klein sind. Ob die bestehenden Unterschiede einen hörbaren Effekt haben, kann hier nicht gesagt werden. Zu groß sind auch die gleichzeitig auftretenden Unterschiede in der Abbildung des Direktschalls.

Zusammenfassend kann gesagt werden, dass die Korrelation im Diffusfeld zwischen den Kanälen einer stereofonen Aufnahme ein wichtiger Parameter für die räumliche Wahrnehmung ist, der sich allerdings kaum unabhängig vom Aufnahmewinkel steuern lässt. Damit ist die Korrelation bei koinzidenten Aufnahmen hauptsächlich vom gewünschten Aufnahmewinkel abhängig.

3.2.2. Optimierung des Doppel-MS Systems hinsichtlich der Diffusfeld-Korrelation

Das Doppel-MS System entspricht nach der Dekodierung einer koinzidenten Anordnung aus 4 oder 5 Mikrofonen erster Ordnung. Größtmögliche Signaltrennung und Gleichmäßigkeit der Richtungsabbildung sowie kleinstmögliche Diffusfeld-Korrelation zwischen den Kanälen wird erreicht, wenn folgende Forderungen erfüllt werden:

- der Winkel zwischen den virtuellen Mikrofonen ist maximal
- die Richtwirkung der virtuellen Mikrofone ist maximal

Forderung a) lässt sich leicht erfüllen, wenn die Ergebnisse der Dekodierung kontrolliert werden. Forderung b) führt zur Wahl einer Richtcharakteristik, die den besten Kompromiss zwischen großer Richtwirkung und störender Rückkeule darstellt - der Superniere. Eine diesbezüglich optimale Anordnung ist demnach eine Anordnung aus 4 oder 5 Supernieren, die im Winkel maximal gegeneinander versetzt sind.

Diese Forderungen führen zum Beispiel zu folgenden 4-Kanal und 5-Kanal Dekodierungen:

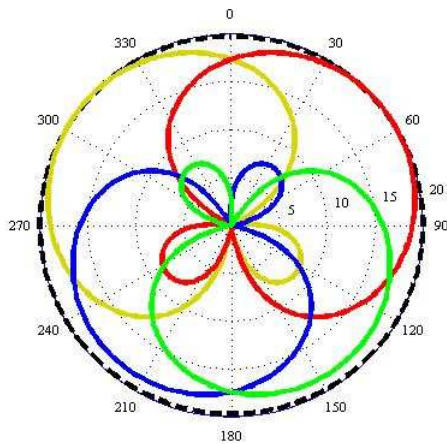


Abbildung 17a: eine optimierte 4-Kanal Dekodierung des Doppel-MS Systems

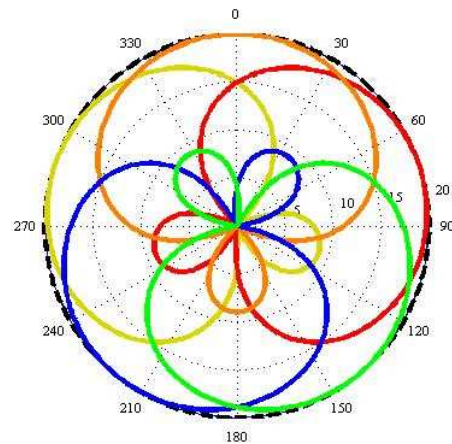


Abbildung 17b: eine optimierte 5-Kanal Dekodierung des Doppel-MS Systems

Die sich ergebenden Korrelationsgrade dieser optimierten Doppel-MS-Dekodierung aus 4 bzw. 5 Supernieren sind:

5-ch Anordnung:

	L-C	L-R	L-LS	LS-RS	LS-C	L-RS
Versatzwinkel	72°	144°	72°	72°	144°	144°
Korrelationsgrad	0.66	0.11	0.66	0.66	0.11	0.11

4-ch Anordnung:

	L-C	L-R	L-LS	LS-RS	LS-C	L-RS
Versatzwinkel	-	108°	72°	72°	-	198°
Korrelationsgrad	-	0.36	0.51	0.66	-	0.04

In obiger Tabelle ist abzulesen, dass eine möglichst geringe Diffusfeld-Korrelation besonders für die 4-Kanal-Variante erreicht wird. Die Bereiche höherer Korrelation sind nur im hinteren Bereich. Die 5-Kanal-Variante zeigt eine höhere Korrelation durch die kleineren Öffnungswinkel der virtuellen Mikrofone.

Es kann vermutet werden, dass eine Doppel-MS Anordnung mit hoher Korrelation eher für Aufnahmeszenarien eignen, die gute Richtungsabbildung benötigen. Eine gute Abbildung von Räumlichkeit und Umhüllung („spaciousness“ und „envelopment“) erfordert laut Griesinger ([8], [10], [11]) eine hohe Dekorrelation der Kanäle.

Bei einer zweikanaligen MS-Aufnahme bzw. XY-Aufnahme ist eine Diffusfeld-Korrelation von 0 möglich. Diese Eigenschaft kann allerdings mit einer XY-Anordnung aus zwei Nieren nicht ausgenutzt werden, da die Korrelation hier nie unter 0.5 sinken kann. Wohl auch deshalb genießt die MS- bzw. XY-Aufnahme generell einen schlechteren Ruf als sie verdient. XY-Aufnahmen werden in der Fachliteratur oft als mittig, eng und schlecht für die Abbildung des Raums beschrieben. Diese Eigenschaften sind allerdings die einer XY-Aufnahme mit hoher Korrelation – treffen also auf eine optimale XY-Anordnung nicht oder wesentlich weniger zu.

In Abbildung 14 kann abgelesen werden, welche Öffnungswinkel eine in diesem Sinne ideale XY-Anordnung ergeben, in dem man den jeweiligen Wert auf der x-Achse für eine Korrelation von 0 abliest. Folgende Werte lassen sich ermitteln:

- Achten ($a=0$): 90° (Blumlein-Technik, Aufnahmewinkel 72°)
- Hypernieren ($a=0.25$): 110° (Aufnahmewinkel 85°)
- Supernieren ($a=0.36$): 160° (Aufnahmewinkel 64°)

Dies ist auch in [11] beschrieben. Mithilfe der Doppel-MS Anordnung können diese und alle anderen koinzidenten Anordnungen erzeugt und somit auf den Parameter Korrelation optimiert werden.

Bei mehrkanaliger koinzidenter Aufnahmetechnik ist eine Diffusfeld-Korrelation von 0 nicht für alle Mikrofonpaare möglich. Je mehr Kanäle erzeugt werden sollen, desto höher ist die Gefahr, stark korrelierte Mikrofonpaare zu erzeugen, wie aus obiger Diskussion hervorgeht.

Die praktischen Konsequenzen für die Aufnahme sind in Abschnitt 4 beschrieben. Es zeigt sich: Die Gewinnung von 4 Kanälen ist durch geeignete Dekodierung gut möglich. Eine Gewinnung von 5 Kanälen aus einem koinzidenten System erster Ordnung ist schwierig und führt nur dann zu guten Ergebnissen, wenn der Tonmeister die Dekodierung kritisch kontrolliert und gegebenenfalls geeignete Maßnahmen wie Delay, Pegel-senkung etc. vornimmt (vgl. Abschnitt 4). Eine Gewinnung von mehr als 5 Kanälen ohne eine starke Ähnlichkeit ist nicht möglich – deshalb muss davor gewarnt werden, allzu vollmundige Versprechungen für neue Surroundformate wie 7.1 oder 10.2 abzugeben. Allein die Möglichkeit, beliebig viele Kanäle zu gewinnen sagt noch lange nichts über deren Qualität aus. Leider existieren sich auch bestimmte Mikrofonssysteme, die nur eine ungenügende Unabhängigkeit der Kanäle bieten. Dieser Aufsatz versucht, ein einfaches Mikrofonssystem zu beschreiben, das trotz der Einfachheit hohe Qualität liefert.

3.3. Crosstalk

3.3.1. Theoretische Analyse

Eine Phantomschallquelle entsteht aus der Wiedergabe eines kohärenten Signals auf zwei Lautsprechern. Gibt man ein kohärentes Signal zusätzlich auf einem dritten Lautsprecher wieder, so entstehen verstärkt Kammfiltereffekte, die störend hörbar sein können.

Dieses dritte Signal wird Crosstalk-Signal genannt. Besteht der Crosstalk aus einem gegenphasigen Signal, ist dies weniger störend als bei einem gleichphasigen.

In der Literatur wird dieses Thema z.B. von Theile [25] behandelt, der Mehrfach-Abbildungen durch Crosstalk durch geeignete Mikrofonierung vermeiden will. Lee und Rumsey [14] untersuchten verschiedene Mehrkanal-Mikrofonanordnungen und stellten

negative Effekte von Crosstalk vor allem auf Ausdehnung und Lokalisiertheit der Phantomschallquellen fest.

Crosstalk bewirkt auch eine Verkleinerung der Hörzone, da durch den Präzedenzeffekt schon eine geringe Annäherung an diesen Lautsprecher eine Fehllokalisierung zur Folge haben kann. Der Effekt auf die Lokalisation abseits des „Sweet Spot“ kann z.B. mit dem *Image Assistant* [33] simuliert werden.

Beim Design einer Mikrofonanordnung ist es wichtig, den Crosstalkpegel so gering wie möglich zu halten. Zur Optimierung der Doppel-MS Anordnung im Hinblick auf minimalen Crosstalk zwischen den Kanälen können dieselben Forderungen (a und b) des vorigen Abschnitts 3.2.2 aufgestellt werden. Auch hier ist eine Anordnung aus virtuellen Supernieren mit größtmöglichem Winkelversatz optimal.

Die in Abbildung 17 gezeigten optimierten Dekodierungen ergeben folgendes Crosstalkverhalten, eine Illustrierung zeigt Abbildung 18:

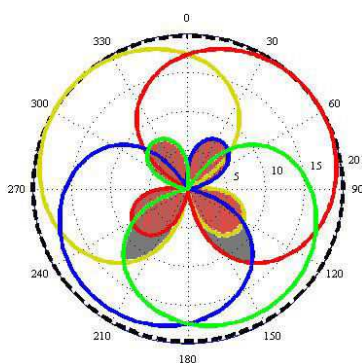


Abbildung 18a:
Crosstalkpegel der optimierten 4-Kanal Dekodierung des Doppel-MS Systems (schwarz eingefärbte Bereiche: gleichphasig, rot eingefärbt: gegenphasig)

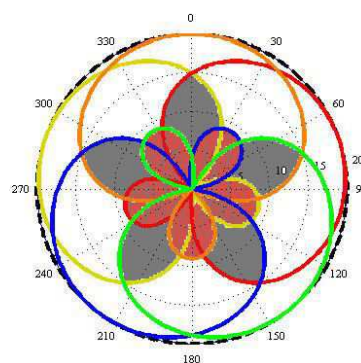


Abbildung 18b:
Crosstalkpegel der optimierten 5-Kanal Dekodierung des Doppel-MS Systems (schwarz eingefärbte Bereiche: gleichphasig, rot eingefärbt: gegenphasig)

Die 4-Kanal Dekodierung aus Abbildung 17a hat einen maximalen Crosstalkpegel von ca. -7.5 dB. Dieser Pegel wird an zwei Stellen erreicht, jeweils bei gleichzeitiger Wiedergabe durch die Lautsprecher L, LS und RS bzw. R, RS und LS. Der Crosstalk im vorderen Bereich besteht aus einem gegenphasigen Signal und hat nur einen geringen Pegel. Die 5-Kanal Dekodierung aus Abbildung 17b hat einen maximalen Crosstalkpegel von ca. -5 dB. Dieser Pegel wird an mehreren Stellen erreicht.

Der Vergleich zwischen den beiden Varianten ergibt eine bessere theoretische Qualität der 4-kanaligen Dekodierung gegenüber der 5-kanaligen. In wieweit ein solches schlechteres theoretisches Verhalten auch tatsächlich schlechtere praktische Eigenschaften nach sich zieht, soll die Untersuchung der nächsten Kapitel zeigen.

Zu einer optimalen Dekodierung des Doppel-MS Systems gehört nicht zuletzt auch ein gleichmäßiger Pegel der Phantomschallquellen. Die in Abbildung 17 und Abbildung 18 gestrichelt gezeichneten Linien zeigen die gleichmäßige Verteilung des Summenleistungspegels in allen Richtungen. Der angegebene Crosstalkpegel ist auf diesen Summenleistungspegel bezogen.

3.3.2. Praktische Analyse

Zur Ermittlung der Wirkungen des Crosstalks auf verschiedene Attribute der Wahrnehmung wurde ein Hörversuch durchgeführt (zu den Details siehe [11]). Ziel der Untersuchung war die Ermittlung der Wahrnehmbarkeitsschwelle eines Crosstalksignals in Bezug auf folgende Attribute der Phantomschallquelle:

- Ausdehnung
- Richtung
- Lokalisiertheit („Ist die Quelle leicht zu lokalisieren?“)
- Klangfarbe

Die Phantomschallquelle wurde zwischen den Lautsprechern Center und Rechts gebildet, die Pegeldifferenz betrug $\Delta L (C/R) = 3.7 \text{ dB}$.

Das Crosstalksignal wurde im linken Lautsprecher erzeugt, und zwar mit verschiedenen Pegeln von -20 bis -5 dB.

Die Hörversuchsteilnehmer, die im „Sweet Spot“ saßen, hörten jeweils eine Testreihe aus 2 Vierergruppen der Form A-B-A-B und A-C-A-C. A war das Referenzsignal ohne Crosstalk, genauso wie eines der Signale B oder C. Das andere Signal war der jeweilige Stimulus mit Crosstalk. Die Zuordnung zu B oder C war randomisiert, wechselte also ständig und war den Versuchsteilnehmern nicht bekannt. Mit diesem Verfahren sollte sichergestellt werden, dass keine negative Bewertung aufgrund eines Vorurteils des Versuchsteilnehmers zustande kommen konnte.

Die Teilnehmer sollten auf folgender Skala die Bewertung vornehmen. Wurde der Stimulus in der falsche Vierergruppe vermutet, wurde die Bewertung negativ gezählt.

Ausdehnung

Ist die Phantomschallquelle bei Beispiel 2 ausgedehnter als bei Beispiel 1?



Es nahmen 15 Personen am Versuch teil.

Der Test wurde mit verschiedenen Testsignalen durchgeführt, darunter trockene Sprache, Sprachaufnahme im Raum, trockene Kastagnetten und eine Kastagnettenaufnahme im Raum.

Die Ergebnisse dieses Versuchs sind in Abbildung 19 bis Abbildung 23 dargestellt. Die Abbildungen 18 bis 21 enthalten beispielhaft die Ergebnisse für das Testsignal Sprachaufnahme im Raum. Die Diagramme zeigen die Hörbarkeit einer Änderung des jeweiligen Attributs. Dabei wurde die oben gezeigte Skala um eins reduziert, um den Wert 0 als „keine Änderung“ zu definieren. Die wahrgenommene Änderung wird in Abhängigkeit vom Crosstalkpegel im dritten Lautsprecher angezeigt.

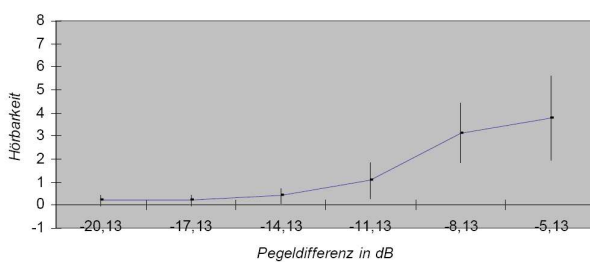


Abbildung 19: aus [11]: Hörbarkeit der Ausdehnungsänderung. Signal: Sprache, trocken. Arithmetischer Mittelwert inkl. 95%-Konfidenzintervalle.

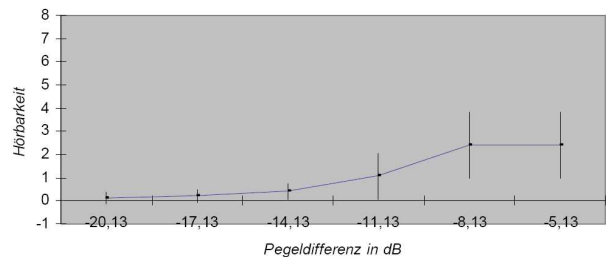


Abbildung 20: aus [11]: Hörbarkeit der Änderung der Lokalisiertheit. Signal: Sprache, trocken. Arithmetischer Mittelwert inkl. 95%-Konfidenzintervalle.

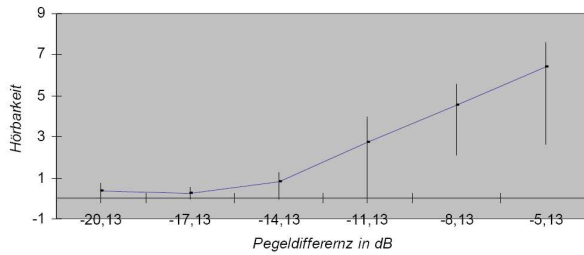


Abbildung 21: aus [11]:
Hörbarkeit der Richtungsänderung.
Signal: Sprache, trocken.
Arithmetischer Mittelwert inkl. 95%-
Konfidenzintervalle.

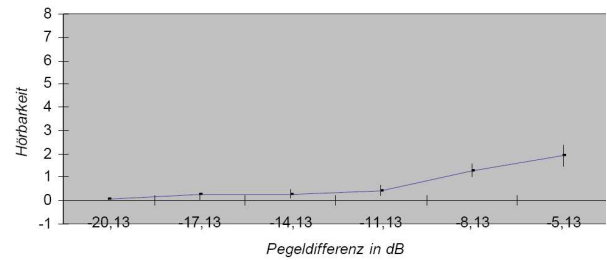


Abbildung 22: aus [11]:
Hörbarkeit der Klangfarbenänderung.
Signal: Sprache, trocken.
Arithmetischer Mittelwert inkl. 95%-
Konfidenzintervalle.

Es zeigt sich, dass Richtungs- und Ausdehnungsänderungen eher wahrgenommen werden als Änderungen der Lokalisiertheit und der Klangfarbe. Die Wahrnehmbarkeitsschwelle (relativ zum Summenpegel der Phantomschallquelle C/R) liegt für Richtungsänderungen bei ca. -12 dB, für Ausdehnungsänderungen bei ca. -9 dB. Die Lokalisiertheit nimmt ab ca. -9 dB ab und die Klangfarbe der Phantomschallquelle ändert sich ab ca. -6 dB.

Abbildung 23 zeigt die Ergebnisse aller Testsignale für das Attribut Ausdehnungsänderung.

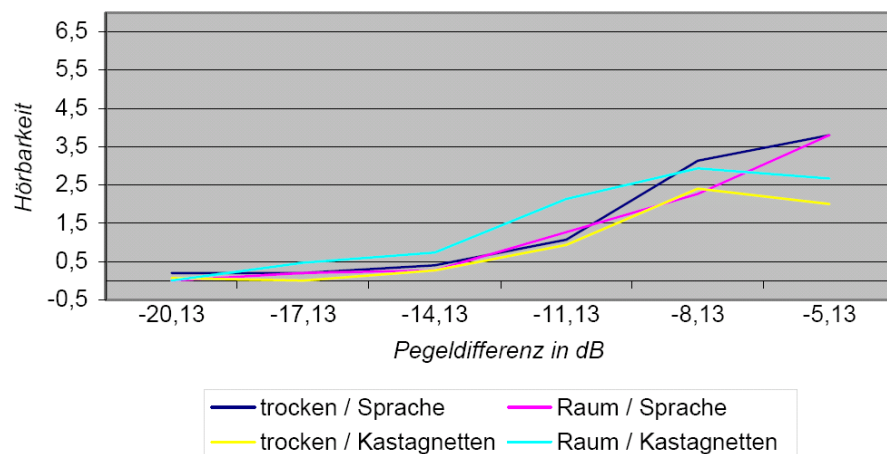


Abbildung 23: aus [11]:
Hörbarkeit der Ausdehnungsänderung, Ergebnisse für alle 4 Testsignale.
Arithmetischer Mittelwerte.

Diese Ergebnisse machen deutlich, dass Crosstalk soweit wie möglich vermieden werden sollte, um negative Effekte zu vermeiden. Dies ist möglich, wenn das Doppel-MS Signal optimal dekodiert wird. Der Crosstalk der optimierten Dekodierungen ist, abhängig von der Variante, allerdings immer noch knapp über der Wahrnehmbarkeitsschwelle für einzelne Attribute.

4. Praktische Untersuchung anhand von verschiedenen Testaufnahmen

Es gilt zwar: „Nichts ist praktischer als eine gute Theorie“ (Gerhard Steinke), aber dennoch hat die Theorie ihre Grenzen und kann noch lange nicht alles erklären, was unsere Wahrnehmung bestimmt. Deshalb ist für eine Diskussion der Doppel-MS Anordnung auch eine praktische Untersuchung notwendig.

Die Zielsetzungen der praktischen Untersuchung sind:

- Test verschiedener Dekodierungsvarianten
- Überprüfung der in den vorangegangenen Kapiteln besprochenen Vor- und Nachteile der Doppel-MS Aufnahmetechnik
- Test der Anwendbarkeit der Doppel-MS Aufnahmetechnik in verschiedenen Aufnahmeszenarien, darunter Musik, Atmo, Theater-Aufzeichnung, Hörspiel, Dokumentation/Film, Fernsehstudio
- Vergleich des Doppel-MS Systems mit anderen Referenzanordnungen

Die Qualität einer Aufnahme ist mit wissenschaftlichen Mitteln nicht leicht zu bewerten. Zu sehr schwanken die unterschiedlichen Erwartungshaltungen und Prioritäten der verschiedenen Hörer. Deshalb sollen hier auch keine allgemeinen Ergebnisse postuliert, sondern nur unsere Erfahrungen und die gesammelten Kommentare anderer beschrieben werden.

4.1. Verschiedene Varianten der Dekodierung von Doppel-MS

Die folgenden, mit den Buchstaben A-F bezeichneten Varianten der Dekodierung der Doppel-MS Signale sind in der praktischen Untersuchung betrachtet worden (siehe [13]). Neben den Polardiagrammen und der Beschreibung der Öffnungswinkel der resultierenden Ausgangskanäle ist jeweils die genaue Pegelmatrix abgebildet, mit der die Dekodierung erreicht wurde.

Außerdem wird auch das jeweils verwendete Delay der hinteren Kanäle angegeben.

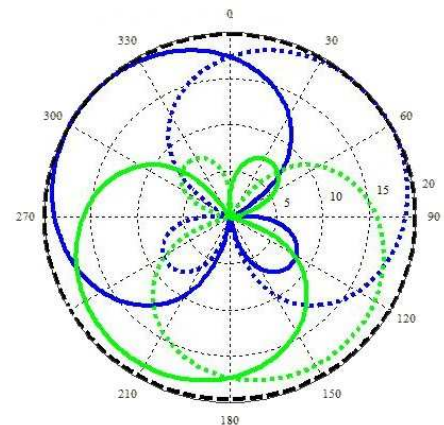
A) 4 Supernieren → Dekodierung wie MDMS U (4ch), siehe Abschnitt 5 b)

Die Richtcharakteristik Superniere hat sich als guter Kompromiss herausgestellt. Aufgrund der starken Richtwirkung der Superniere bei einer maximalen Schallunterdrückung im hinteren Halbraum ist diese Einstellung nur von sehr geringem Crosstalk behaftet. Vorteilhaft ist der Verzicht auf den Center-Kanal, da so größere Öffnungswinkel möglich werden.

Die Richtungsabbildung ist im gesamten Bereich ausgewogen und präzise. Der Aufnahmewinkel für die Basis L/R im vorderen Bereich beträgt 110°.

Die Unterdrückung von Direktschall in den hinteren Kanälen funktioniert gut.

Diese Eigenschaften prädestinieren diese Dekodierung für Musikaufnahme und Hörspiel, ob allerdings die Varianten mit Center bevorzugt werden, hängt nach der Erfahrung des Autors stark von der Intention und dem Geschmack des einzelnen ab.



Loudspeaker Feeds	L	R	C	LS	RS
Front: Cardioid	-0.0	-0.0	-inf	-14.9	-14.9
Fig.8: <input checked="" type="checkbox"/> MK 8	-0.9	-0.9	-inf	-5.1	-5.1
Rear: Cardioid	-inf	-inf	-inf	0.0	0.0

normalize dB yellow Color marks inverted phase

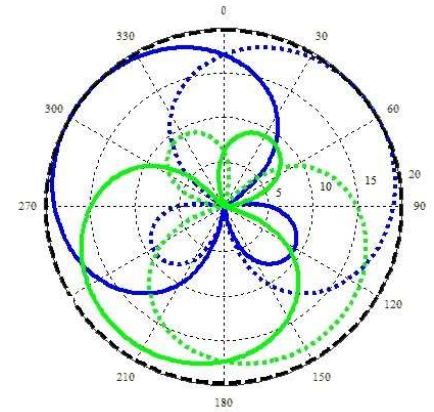
Virtual Microphones:	L,R	LS,RS
Polar Pattern	0.369	0.36
Angle (+/-)	54.26	144
Level	1.5	0

0=Fig. of 8
1=Omini

B) 4 Supernieren, etwas breitere Abbildung

Um den Aufnahmewinkel einer Stereoabbildung zu verkleinern, muss entweder der Versatzwinkel zwischen den Mikrofonen vergrößert oder deren Richtwirkung erhöht werden. Dieses Beispiel unterscheidet sich nur sehr geringfügig vom vorangegangenen, denn jeder Versuch, den L/R-Aufnahmewinkel wesentlich zu verkleinern und gleichzeitig eine stereofone Abbildung in LS/RS zu behalten, würde zu einer sehr ungleichmäßigen Energieverteilung und starkem Übersprechen der Kanäle führen. Bei dieser Einstellung beträgt der L/R-Aufnahmewinkel ca. 90° (im Vergleich zu ca. 110° im Variante A).

Loudspeaker Feeds					
	L	R	C	LS	RS
Front: Cardioid	-0.3	-0.3	-inf	-11.2	-11.2
Fig.8 <input checked="" type="checkbox"/> MK 8	-0.5	-0.5	-inf	-5.7	-5.7
Rear: Cardioid	-inf	-inf	-inf	0.0	0.0
<input checked="" type="checkbox"/> normalize	<input checked="" type="checkbox"/> dB	yellow Color marks inverted phase			

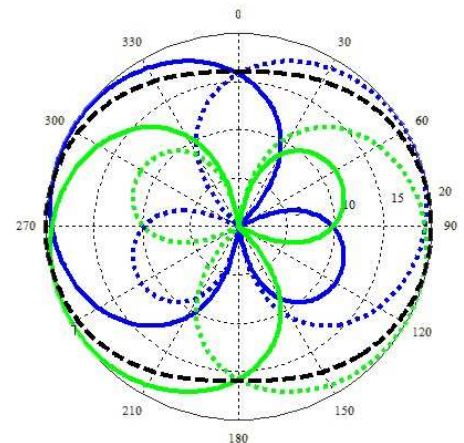


Virtual Microphones:			
	L,R	LS,RS	
Polar Pattern	0.355	0.325	0=Fig. of 8 1=Omni
Angle (+/-)	56.55	148	°
Level	0.472	-1.24	dB

C) „konventionelle“ MS-Dekodierung 4ch

Diese Dekodierung wurde als Negativbeispiel mit in die Vergleiche aufgenommen. Die starke Rückwärtsempfindlichkeit der Mikrone bewirkt ein deutliches Übersprechen und damit starke klangliche Unregelmäßigkeiten für Hörer außerhalb des „Sweet Spots“. Die Energieverteilung ist nicht optimal. Die Hörerergebnisse zeigen, dass diese Variante klanglich und räumlich schlecht beurteilt wird.

Loudspeaker Feeds					
	L	R	C	LS	RS
Front: Cardioid	-2.3	-2.3	-inf	-inf	-inf
Fig.8 <input checked="" type="checkbox"/> MK 8	0.0	0.0	-inf	0.0	0.0
Rear: Cardioid	-inf	-inf	-inf	-2.3	-2.3
<input checked="" type="checkbox"/> normalize	<input checked="" type="checkbox"/> dB	yellow Color marks inverted phase			

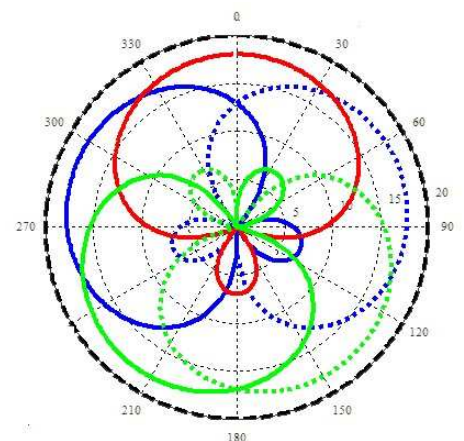


Virtual Microphones:			
	L,R	LS,RS	
Polar Pattern	0.309	0.309	0=Fig. of 8 1=Omni
Angle (+/-)	63.43	116.6	°
Level	-1.84	-1.84	dB

D) 5 Supernieren → Dekodierung wie MDMS U (5ch), siehe Abschnitt 5 b)

Um eine gleichmäßige Energieverteilung und ausgewogene Lokalisation zu erhalten, sind die Supernieren für L/R gegenüber der 4-kanaligen Ausführung weiter nach außen gedreht. Die Abbildung ist ausgewogen, die Stabilität durch den Center-Kanal gegeben. Ob diese oder Variante A für Aufnahmen mit ausgewogener Abbildung Verwendung findet ist auch Geschmackssache.

Der Center-Pegel kann reduziert werden, um Crosstalk vorne zu vermeiden und die Richtungsabbildung zu verbessern. Klanglich ist die 5-Kanal-Variante ohne



weitere Maßnahmen der 4-Kanal-Variante unterlegen, allerdings kann der Center-Kanal durchaus eine wichtige Rolle bei Programm-wichtigem Inhalt übernehmen, z.B. für Soloinstrumente oder im Hörspiel. Außerdem kann er Stabilität bieten.

Loudspeaker Feeds					
	L	R	C	LS	RS
Front: Cardioid	-5.1	-5.1	0.0	-16.0	-16.0
Fig.8 <input checked="" type="checkbox"/> MK 8	-2.0	-2.0	-inf	-6.2	-6.2
Rear: Cardioid	-15.8	-15.8	-11.1	-1.1	-1.1
<input checked="" type="checkbox"/> normalize <input checked="" type="checkbox"/> dB	yellow Color marks inverted phase				

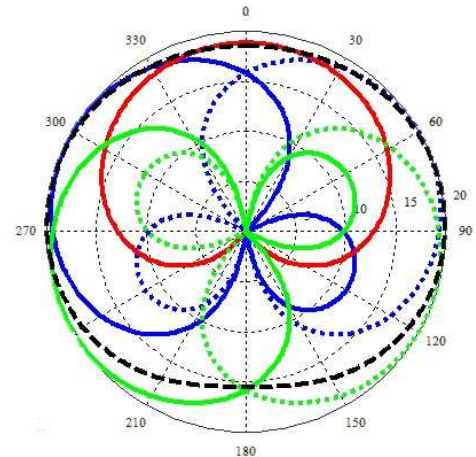
Virtual Microphones:			
	L,R	C	LS,RS
Polar Pattern	0.36	0.36	0.36
Angle (+/-)	71.99	0	144
Level	-1.97	-1.97	-1.97

0=Fig.of 8
1=Omni

E) „konventionelle“ MS-Dematrixierung 5ch

Diese Dekodierung wurde ebenfalls als Negativbeispiel mit in die Vergleiche aufgenommen. In dieser Einstellung überlappen sich die Kanäle besonders stark.

Diese Einstellung zeigt auch im praktischen Versuch, dass eine schlechte Dekodierung ohne Kontrolle schlechte Ergebnisse liefert. Unangenehme Klangfarbe und mangelnde räumliche Transparenz sind die Folge. Die Klangfarbe ändert sich schon mit kleinen Kopfbewegungen.



Loudspeaker Feeds					
	L	R	C	LS	RS
Front: Cardioid	-3.0	-3.0	0.0	-inf	-inf
Fig.8 <input checked="" type="checkbox"/> MK 8	-0.7	-0.7	-inf	-0.7	-0.7
Rear: Cardioid	-inf	-inf	-inf	-3.0	-3.0
<input checked="" type="checkbox"/> normalize <input checked="" type="checkbox"/> dB	yellow Color marks inverted phase				

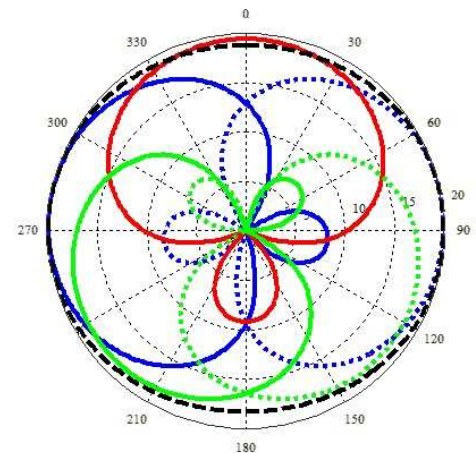
Virtual Microphones:			
	L,R	C	LS,RS
Polar Pattern	0.309	0.5	0.309
Angle (+/-)	63.43	0	116.6
Level	-1.84	-3	-1.84

0=Fig.of 8
1=Omni

F) Verzögerte Surroundkanäle

Diese Einstellung stellt eine Verbesserung der 5-Kanal-Variante D dar, da die Unabhängigkeit der Kanäle vergrößert wird: sowohl vorne durch die Vergrößerung des Öffnungswinkels als auch zwischen vorne und hinten durch das Delay ($\Delta t = 10\text{ms}$).

Diese Variante kann für viele Aufnahmeszenarien die optimale Lösung sein, wenn der Center-Kanal gewünscht wird. Allerdings muss natürlich ein vorne-bezogenes Klangbild vorliegen, denn das Delay verhindert die stabile Hinten-Ortung.



Loudspeaker Feeds					
	L	R	C	LS	RS
Front: Cardioid	-6.5	-6.5	0.0	-27.1	-27.1
Fig.8 <input checked="" type="checkbox"/> MK 8	-1.2	-1.2	-inf	-5.0	-5.0
Rear: Cardioid	-10.1	-10.1	-10.2	-3.0	-3.0
<input checked="" type="checkbox"/> normalize <input checked="" type="checkbox"/> dB	yellow Color marks inverted phase				

Virtual Microphones:			
	L,R	C	LS,RS
Polar Pattern	0.369	0.346	0.368
Angle (+/-)	83.16	0	131.3
Level	-0.593	-1.15	-2.01

0=Fig.of 8
1=Omni

4.2. Subjektive Erfahrungen mit Doppel-MS Aufnahmen

Die Erfahrungen mit verschiedenen Aufnahmeszenarien zeigten in vielen Fällen, dass die Doppel-MS Aufnahmetechnik durchaus gute bis sehr gute Ergebnisse lieferte. Hier sind unsere subjektiven Erfahrungen wiedergegeben, die auch mit einigen anderen diskutiert werden konnten und sich oft mit deren Erfahrungen deckten.

So war z.B. eine Atmo mit vielen Direktschallanteilen in allen Richtungen (Marktplatz) gut darstellbar und zeigte in der 4-Kanal-Dekodierung ähnliche Qualität wie das simultan aufgezeichnete IRT-Kreuz. Die Flexibilität der Richtungsabbildung nach der Doppel-MS Aufnahme stellte sich für die schräg vorbeifahrende Straßenbahn als echter Vorteil heraus.

Eine Atmo, die mehr auf eine effektvolle Räumlichkeit zielte (Feuerwerk am Volksfest), konnte mit dem IRT-Kreuz besser und effektvoller wiedergegeben werden. Das IRT-Kreuz kann gleichzeitig eine gute 360°-Abbildung liefern und einen guten Eindruck von Umhüllung und Räumlichkeit liefern.

Einige Musikaufnahmen (Klavierkonzert, Kammermusik, großes Orchester) gelangen überraschend gut mit dem Doppel-MS System, das parallel aufgezeichnete OCT Surround Setup lieferte jedoch sehr gute, effektvollere und raumbetonte Ergebnisse. Eine Entscheidung zwischen diesen Verfahren bleibt dem Geschmack des Tonmeisters überlassen. In unserer Erfahrung kann Doppel-MS eher kleine Klangkörper darstellen, bei größeren Klangkörpern in großen Räumen bieten sich eher OCT Surround oder andere Verfahren an. Bei Musikaufnahmen ist die Verwendung einer zusätzlichen Tiefpaß-gefilterten Kugel zur Aufnahme tiefster Frequenzen empfehlenswert. Die Mischung der Doppel-MS Signale mit einer Groß-AB Kugel-Aufstellung im Raum brachte außerdem den gewünschten weiträumigen Klang des Raums. Letztere Lösung liefert außerdem dekorrelierte Tieffrequenzsignale.

Ein A-Cappella Ensemble, das eine schöne Abbildung zwischen den Lautsprechern erforderte, wurde mit Doppel-MS sehr gut wiedergegeben, beide Dekodierungsvarianten mit und ohne Center waren möglich.

Ein Jazz-Ensemble in einem Jazzclub mit Publikum wurde mit Doppel-MS und einzelnen Mikrofonen für die Instrumente aufgezeichnet. Die Atmosphäre und der Raumklang in den Doppel-MS Signalen ergänzte gut die Mischung der Instrumentenmikrofone. Ein parallel aufgezeichnetes Doppel-ORTF-Mikrofon lieferte eine ähnlich gute, allerdings weniger flexible Raumabbildung. Die Flexibilität war wichtig, da in der Mischung der Wunsch nach einer unterschiedlichen Einstellung der Stereobreite bei Musikpassagen und Applauspassagen entstand.



Abbildung 24: Testaufnahmen mit dem Doppel-MS System und Referenzanordnungen:
links: Doppel-MS und OCT Surround System bei einer Theater-Live-Aufzeichnung
rechts: Doppel-MS und IRT-Kreuz bei einer Atmoaufnahme am Durlacher Marktplatz.

Die Theater-Live-Aufnahme gelang mit dem Doppel-MS besser als mit OCT. Grund dafür war wohl die schlechte Räumlichkeit, da die Mikrofone schon im Bühnenhaus hängen

mussten und dessen Akustik einfügen. Das Doppel-MS System stellte den Raum weniger in den Vordergrund.

Eine Live-Aufzeichnung einer TV-Sendung mit Publikum gelang sowohl mit IRT-Kreuz als auch Doppel-MS gut.

Die Verwendung des Doppel-MS Setups bei Hörspielaufnahmen bietet sich schon deshalb an, da hier koinzidente Techniken gerne verwendet werden und die Doppel-MS Technik mehr Flexibilität und natürlich die Mehrkanalfähigkeit bietet. Außerdem ist gerade bei Hörspielen Abwärts- bzw. sogar Monokompatibilität gefordert, die mit der Doppel-MS Technik ideal erreicht wird.

Die Anwendung der Doppel-MS Technik im Filmtone ist seit einiger Zeit bewährte Praxis. Eine Anwendung des Doppel-MS Systems im Dokumentarfilm brachte neue Ideen bei der Auswahl der Mikrofone, siehe auch [37]: Es ist manchmal sinnvoll, die nach vorne gerichtete Niere durch eine stärker gerichtete Kapsel zu ersetzen, da dieses Mikrofon diskret für das Dialogsignal im Center verwendet werden soll. Deshalb kann entweder eine Superniere verwendet werden, oder aber es wird ein Rohrichtmikrofon eingesetzt. Letzteres erfordert allerdings eine spezielle Anordnung, die weiterhin eine optimale Koinzidenz ermöglicht, siehe Abbildung 25. Die Doppel-MS Technik bewährte sich gut in dieser Anwendung, die Vorteile des Surround-Tons an der Angel zeigten sich vor allem dadurch, dass bei der Dokumentarfilmproduktion die Authentizität im Surround-Ton eine wichtige Rolle spielt und deshalb oft die subjektive Tonperspektive verwendet werden kann.

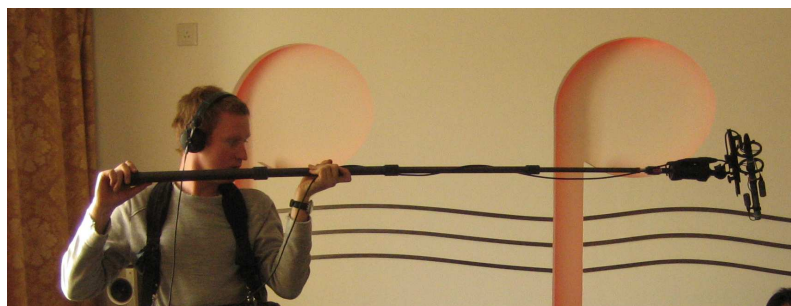


Abbildung 25: Die Doppel-MS Anordnung mit Rohrichtmikrofon
links: Realisierung mit dem Schoeps Double-M/S Set „CMIT“
(verwendete Mikrofone: CMIT 5U, CCM 4 und CCM 8)
rechts: bei der Anwendung, [37], Foto: André Zacher

5. Werkzeuge zur Dekodierung von Doppel-MS

Da die Anwendungsbereiche der Doppel-MS Anordnung sehr verschieden sind, existieren mehrere Möglichkeiten, die Dekodierung vorzunehmen. Dabei ist es egal, ob während der Aufnahme oder in der Nachbearbeitung (Postproduktion) dekodiert wird.

Folgende drei grundsätzliche Möglichkeiten zur Dekodierung werden erläutert:

- a) zwei MS-Matrizen in Mischpult oder Schnittsoftware
 - b) Hardware Doppel-MS Matrix MDMS U
 - c) Software VST Plugin „Double M/S Tool“
- zu a) Doppel-MS Signale können grundsätzlich genauso wie MS-Aufnahmen dekodiert werden, dazu werden dann zwei statt nur einer MS-Matrix verwendet. Außerdem kann die vordere Niere auch als Center-Signal verwendet werden. Diese Methode wird in vielen Fällen gute Ergebnisse liefern, das Ergebnis sollte aber vom Tonmeister kontrolliert werden, da auch ungünstige Dekodierungen entstehen könnten (vgl. Abschnitt 4). Falls die Doppel-MS Technik mit einem Rohrichtmikrofon

als M-Signal verwendet wird, wird diese Methode empfohlen, da eine Mischung zwischen hinterer Niere und Rohrriechmikrofon keinen Sinn macht.

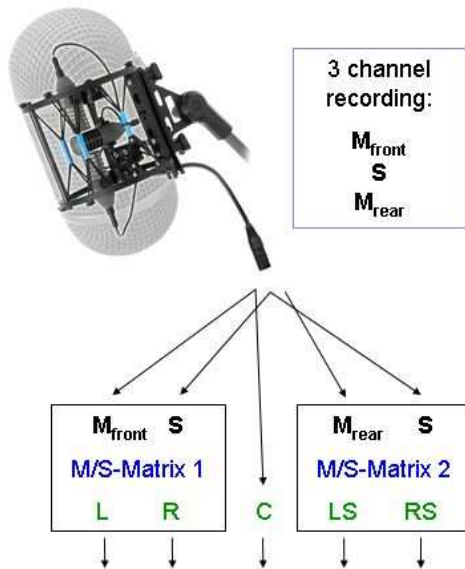


Abbildung 26:

konventionelle Doppel-MS
Dekodierung nach Variante a)

Wie in den vorigen Abschnitten besprochen, lassen sich die Eigenschaften der Doppel-MS Anordnung mithilfe einer geeigneten Dekodierung verbessern. Diese Dekodierung erfordert die Einbeziehung aller drei Kapseln zur Synthese der Signale L/R/LS/RS und der zwei Nieren für den Center-Kanal. Diese Dekodierung ist zwar mithilfe von Pegel-Matrizen oder im Mischpult möglich, aber generell umständlich. Deshalb ist die Verwendung spezieller Werkzeuge hilfreich, die diese optimierte Dekodierung ermöglichen:

zu b) Hardware-Matrix Schoeps MDMS U

Diese analoge, passive Matrix liefert die 4 bzw. 5 dekodierten Kanäle L/R/C/LS/RS aus den drei Doppel-MS Signalen. Sie kann direkt bei der Aufnahme im Mikrofonkabel verwendet werden oder aber in der Nachbearbeitung. Sie bietet die beiden Dekodierungsvarianten A und D aus Abschnitt 4. In der Matrix wird die Empfindlichkeit der Schoeps-Kapseln angeglichen.

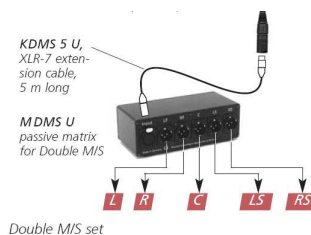


Abbildung 27:

Hardware Matrix
Schoeps MDMS U

zu c) Software VST PlugIn „Double M/S Tool“

Dieses PlugIn wird in einem Sequencerprogramm mit Multichannel VST-Schnittstelle verwendet. Es eignet sich hervorragend zur flexiblen und intuitiven Dekodierung der Doppel-MS Signale. Die Bedienung ist selbsterklärend, da alle Änderungen sofort in den Polardiagrammen visualisiert werden. Auch die Audio-Ausgangssignale werden in Echtzeit mit den aktuellen Parametern geändert. Ebenso wie die Hardware hat das PlugIn drei Eingangskanäle (die Signale der Doppel-MS Anordnung) und 5 Ausgangskanäle (L/C/R/LS/RS). Im PlugIn wird automatisch die unterschiedliche Empfindlichkeit der Schoeps-Kapseln angepasst und eine Entzerrung der Kapsel CCM/MK 8 vorgenommen. Das PlugIn ist kostenlos zum Download verfügbar [21]. Außerdem werden kurze Audio-Samples von Doppel-MS verfügbar sein, damit das PlugIn ausprobiert wer-

den kann.
Es ist bisher nur für Windows erhältlich.

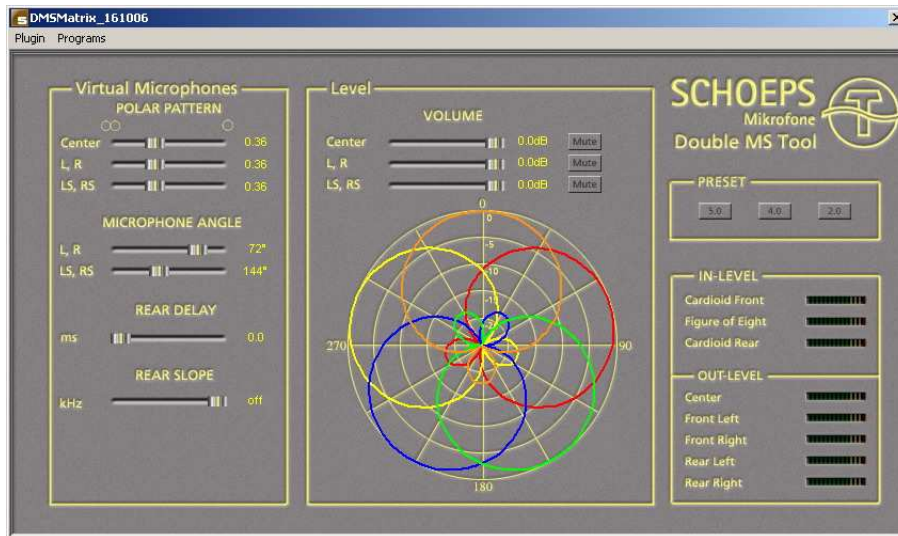


Abbildung 28:

VST PlugIn
„Double M/S Tool“

6. Danksagung

Ich will mich für die hervorragende Arbeit bedanken, die Christopher Haut innerhalb seiner Diplomarbeit und Daniel Keinath innerhalb seines Praktikums in ihrer Zeit bei der Firma Schoeps an diesem Thema geleistet haben.

Herzlicher Dank geht auch an Prof.Dr.Jörg Bitzer vom IHA der Fachhochschule Oldenburg für seine Unterstützung.

7. Literatur

- [1] Benjamin E., Chen T.: "The native B-format Microphone, Part I", 119th AES Convention, New York, 2005, Preprint No.6621
- [2] Brittain, F. H., Leakey, D. M.: "Two-channel stereophonic sound systems. Wireless World 206-210, 1956
- [3] Camerer, F.: „Practical Surround-Sound-Production - Part 2: TV-Documentaries“, AES 19th International Conference, Ellmau, May 2001
- [4] Cremer: „Zur Verwendung der Worte Korrelationsgrad und Kohärenzgrad“, ACUSTICA, Vol.35, No.3, p.215-218, Juni 1976
- [5] Damaske, P.: „Subjektive Untersuchung von Schallfeldern“, ACUSTICA, Vol.19, No.4, p.199-213, 1967/1968
- [6] Flock, S.: "Theoretische sowie objektive und subjektive Messungen am Soundfield-Mikrofon", Diplomarbeit im Studiengang Ton und Bild, FH Düsseldorf
- [7] Gerzon, M.: "Periphony: With-Height Sound Reproduction", Journal of the Audio Engineering Society, 21(1):2-10, 1973.
- [8] Griesinger, D.: „Spaciousness and Localization in Listening rooms – how to make coincident recordings sound as spacious as spaced microphone arrays“, 79th AES Convention, New York, 1985, Preprint No.2294
- [9] Griesinger, D.: „New Perspectives on Coincident and Semi Coincident Microphone Arrays“, 82nd AES Convention, London, 1987, Preprint No.2464
- [10] Griesinger, D.: „Physik, Psychoakustik und Surround-Technik“, Production Partner Spezial 21.Tonmeistertagung, Musik-Media-Verlag Ulm, 2000, 2.

- [11] Griesinger, D.: „Räumliches Hören in Theorie und Praxis: Wie ergänzt man Tiefe und Halligkeit mit künstlichem Nachhall ohne Beeinträchtigung der Deutlichkeit - The Theory and Practice of Perceptual Modeling - How to use Electronic Reverberation to Add Depth and Envelopment Without Reducing Clarity“, www.world.std.com/~griesngr, Stand Oktober 2006
- [12] Haut, C.: „Programmierung und Evaluierung einer parametrisierbaren Doppel-MS Dekodierung für eine 5.1 Surround-Abmischung“, Diplomarbeit des Studienganges Hörtechnik & Audiologie an der FH Oldenburg, 2006.
- [13] Keinath, D.: „Testaufnahmen mit dem Doppel-MS Verfahren“, interner Praktikumsbericht der Firma Schoeps, Karlsruhe, 2005.
- [14] Lee, H. K. und Rumsey F.: „Investigation into the effect of interchannel Crosstalk in Multichannel Microphone Technique“, 119th AES Convention, New York 2005, Preprint No.6374.
- [15] Lipshitz, S.P.: "Stereo Microphone Techniques: Are the Purists Wrong?", AES Convention, Anaheim, May 1985, Preprint No.2261
- [16] Leakey, D. M.: "Further thoughts on stereophonic sound systems". Wireless World, 154-160, 1960
- [17] Martin, G.: „The Significance of Interchannel Correlation, Phase and Amplitude Differences on Multichannel Microphone Techniques“, 113th AES Convention, 2002, Los Angeles, Preprint No.5671
- [18] Mertens, H.: "Directional hearing in stereophony theory and experimental verification", Europ. Broadcasting Union Rev. Part A, 92, 1-14, 1965
- [19] Pfanzagl, E.: „Über die Wichtigkeit ausreichender Dekorrelation bei 5.1 Surround-Mikrofonsignalen zur Erzielung besserer Räumlichkeit“, Bericht zur Tonmeistertagung 2002, Hannover
- [20] Schoeps-website: www.schoeps.de, Stand 2006
- [21] Schoeps-website: Download des VST-PlugIn „Double M/S Tool“ unter www.schoeps.de/dmsplugin.htm, Stand 2006
- [22] Sengpiel, E., Vorlesungsunterlagen: „Theoriegrundlagen:„Intensitäts“-Stereofonie“, „Theoriegrundlagen: Laufzeitstereofonie“, <http://www.sengpielaudio.com/TheorieGrundlaIntensitaet.pdf>, <http://www.sengpielaudio.com/TheorieGrundlaLaufzeit.pdf>, Stand 12. Mai 2006
- [23] Simonson, G.: Master´s Thesis. Lyngby, Denmark, 1984
- [24] Soundfield-website: www.soundfield.com, Stand 2006
- [25] Theile, G.: „Multichannel natural recording based on psychoacoustics principles“, 108th AES convention, 2000, Preprint No. 5156.
- [26] Theile, G.: "On the performance of two-channel and multi-channel stereophony", 88th AES Convention, 1990, Preprint No.2932
- [27] Williams, M.: „Multichannel sound recording practice using microphone arrays“, 24th AES International Conference.
- [28] Williams, M.: „The Stereophonic Zoom: A Practical Approach to Determining the Characteristics of a Spaced Pair of Directional Microphones“, 75th AES Convention, 1984, Preprint No.2072
- [29] Williams, M.: „Microphone Arrays for Stereo and Multichannel Sound recording“, Il Rostro, Milan, 2000
- [30] Wittek, H., Theile, G.: Investigations on directional imaging using L-C-R stereo microphones. (German), 21. Tonmeistertagung 2000, Proceedings pp. 432-455
- [31] Wittek, H., Theile, G.: „The recording angle – based on localization curves“, 112th AES Convention, Munich 2002, Preprint No.5568
- [32] Wittek, H.: „Untersuchungen zur Richtungsabbildung mit L-C-R *Hauptmikrofonen*“, Diplomarbeit des Studienganges Ton- und Bildtechnik an der FH Düsseldorf, 2000.

- [33] Wittek, H., „Image Assistant“, Java Applet auf www.hauptmikrofon.de, Stand 2006
- [34] Wittek, H., Neumann, O., Schaeffler, M, Millet, C.: „Studies on Main and Room Microphone Optimization“, AES 19th Int.Conference, Ellmau, 2001
- [35] Wolfram-Homepage, Lösungen zu Kugelflächenfunktionen, <http://mathworld.wolfram.com/SphericalHarmonic.html>, Stand: Juni 2006.
- [36] Wuttke, J.: Principles of microphones and stereo recordings, in: Mikrofonaufsätze, auf www.schoeps.de, Stand 2006
- [37] Zacher, A.: „Dragonsongs – Lang Lang in China, Dokumentarischer Dreh“, Bericht auf <http://www.proaudio.de/reports/praxis/langlang/index.html>, Stand: Oktober 2006

Anhang

Zur Begrifflichkeit der Korrelation und der Kohärenz im Zusammenhang mit der Analyse der Aufnahmetechnik:

Cremer beschreibt in [4] die Definitionen der Begrifflichkeiten recht anschaulich. Auf diese Quelle bezieht sich der Autor dieses Aufsatzes. Die Korrelation ist ein Maß für die Ähnlichkeit zweier Signale im Zeitbereich. Es existieren zwei wichtige Begriffe: Korrelationsgrad und Korrelationsfunktion.

Der *Korrelationsgrad* gibt die Korrelation bei der Laufzeit $\tau = 0$ an:

$$\chi = \frac{\overline{p_1(t)p_2(t)}}{\tilde{p}_1\tilde{p}_2} \quad \text{mit} \quad \overline{p_1(t)p_2(t)} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{+T/2} p_1(t)p_2(t)dt$$

Erläuterung: Der Korrelationsgrad χ ist der Quotient aus dem Produkt der Signale p_1 und p_2 , gemittelt über der Zeit T , und dem Produkt der Effektivwerte (=Autokorrelationsfunktionen zum Zeitpunkt 0). Der Nenner sorgt dafür, dass der Korrelationsgrad „normiert“ ist, das heißt, er nimmt nur Werte zwischen -1 und 1 an. Das heißt auch: Der Korrelationsgrad ist nicht abhängig vom Pegelverhältnis der beiden Signale. Der Korrelationsgrad ist kein guter Wert, um Stereophonie mit Laufzeiten zu überprüfen, da er nur den Zeitpunkt $\tau = 0$ analysiert. Trotzdem wird der Korrelationsgrad meistens größer 0 sein, wenn AB-Stereophonie verwendet wird, da die tiefen Frequenzen bei AB-Stereophonie, abhängig vom Abstand, stets korreliert sind.

Die *Korrelationsfunktion* analysiert nun die Korrelation auch bei allen anderen Werten für τ :

$$\chi(\tau) = \frac{\overline{p_1(t)p_2(t+\tau)}}{\tilde{p}_1\tilde{p}_2}$$

Erläuterung: Es können damit auch solche Signale als „korreliert“ genannt werden, die über einen Zeitversatz τ verfügen. Allerdings muss dieser Zeitversatz für alle Frequenzen gleich sein, da ja im Zeitbereich analysiert wird und damit sonst die Korrelationsfunktion einen geringeren Wert annimmt. Wenn die Korrelationsfunktion abhängig von der Frequenz betrachtet wird, das heißt

$$\chi(\tau, \omega) = \frac{\overline{p_1(t, \omega)p_2(t+\tau, \omega)}}{\tilde{p}_1(\omega)\tilde{p}_2(\omega)}$$

ergibt sich eine Funktion, die für jede Frequenz separat die Korrelation bestimmen kann, und dann können auch unterschiedliche τ vorliegen. Damit könnte jetzt auch ab-Stereofonie analysiert werden.

Wenn die Analyse der Ähnlichkeit nicht im Zeitbereich, sondern im Frequenzbereich passiert, spricht man von der

Kohärenzfunktion:

$$\gamma(\omega) = \frac{|\Phi_{12}(\omega)|}{[\Phi_{11}(\omega)\Phi_{22}(\omega)]^{1/2}}$$

Die Kohärenzfunktion ist der Quotient aus der Fourier-Transformierten der Kreuzkorrelationsfunktion und dem Produkt der Fourier-Transformierten der Autokorrelationsfunktionen. Achtung: Meistens wird die Kohärenz(-funktion) im Quadrat angegeben ($\gamma^2(\omega)$). Die Verwandtschaft zur frequenzabhängigen Korrelationsfunktion wird ersichtlich, wenn bei letzterer das Maximum über der Variable τ für jede einzelne Frequenz gesucht wird. Dann gilt:

$$|\chi'(\omega)|_{\max} = \gamma(\omega)$$

Das „Strichlein“ des χ' bedeutet, dass diese Beziehung nur gilt, wenn die Frequenzen einzeln beobachtet werden. Der Betrag des Maximums der Korrelationsfunktion kann also als Maß für die Kohärenz angesehen werden.

Grundsätzlich kann man trennen zwischen der Korrelation, die eine Ähnlichkeit der Signale im Zeitbereich betrachtet und die Gruppenlaufzeiten τ liefert, und der Kohärenz, die die Ähnlichkeit zweier Signale im Frequenzbereich untersucht und die Phasenbeziehung dabei außer Acht lässt.

Welches Maß ist nun geeignet, eine Aussage über eine stereofone Anordnung zu liefern? Zunächst ist klar: Wird auch Laufzeit-Stereofonie untersucht, muss eine frequenzabhängige Größe gesucht werden, somit kommt der Korrelationsgrad nicht in Frage. Wenn von der „Korrelation“ oder „korreliert“ die Rede ist, sollte darauf geachtet werden, was damit gemeint wird. Meist ist die Korrelationsfunktion gemeint, und für die gilt zum Beispiel, dass zwei kohärente Signale mit unterschiedlichen Gruppenlaufzeiten eine Korrelation kleiner 1 ergeben. Bei ab-Stereofonie ist die Kohärenzfunktion im Diffusfeld frequenzabhängig, wie Abbildung 33 zeigt. In diesem Fall zeigt der Korrelationsgradmesser ungefähr den Wert 0.2. Dies geht aus den Untersuchungen Damaskes [5] hervor. Der Korrelationsgradmesser liefert also in etwa eine gewichtete Mittelung aus den Einzelkorrelationen, wobei die tiefen Frequenzen eine wesentliche Rolle spielen. Abbildung 29 zeigt den Korrelationsgrad einer ab-Anordnung im diffusen Schallfeld in Abhängigkeit vom Mikrofonabstand.

Es ist ein Erfahrungswert der Tonmeister, welche angezeigte Korrelation am Korrelationsgradmesser bei AB-Stereofonie welchen Klang ergibt. Die Aussagekraft der Korrelations-/Kohärenzfunktion für die räumliche Wahrnehmung ist in höchstem Maße interessant. Wie nimmt das Gehör die Kohärenz bei unterschiedlichen Frequenzen wahr? Offensichtlich ist das Gehör in der Lage, eine Kohärenz zu erkennen, also die Ähnlichkeit der Signale trotz unterschiedlicher Gruppenlaufzeiten zu analysieren. Wie wertet das Gehör die Kohärenz in unterschiedlichen Frequenzbereichen aus?

Zur Erläuterung der Aussagekraft der Kohärenz sind nun drei Beispiele einer Kohärenzfunktion für verschiedene stereofone Anordnungen abgebildet. Es ist eine spannende Aufgabe, diese Daten psychoakustisch oder empirisch (in der Tonmeisterpraxis) auszuwerten.

Die Graphen zeigen die quadrierte Kohärenzfunktion in Abhängigkeit von der Frequenz. Diese ist nur bis 1200 Hz gezeichnet. Die Zahl im Graphen gibt die Kohärenz bei $f=0$ an.

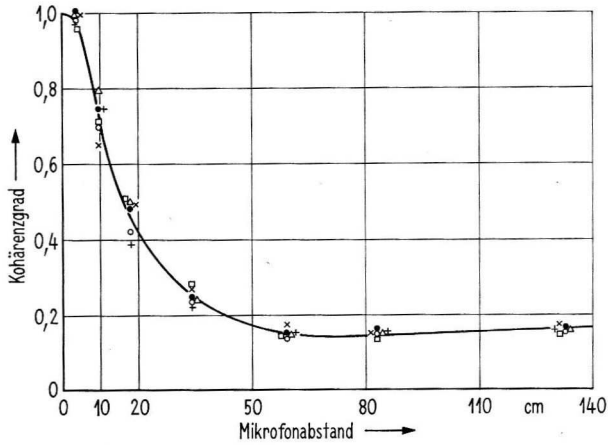


Fig. 3. Mittlerer Kohärenzgrad \bar{k} der Signale von vier Mikrofonen im Hallraum; gemessen mit gefilterten δ -Impulsen (0,25 kHz bis 2 kHz). Die sechs verschiedenen Symbole entsprechen den Messungen für die sechs verschiedenen Kohärenzgrade der vier Signale.

Abbildung 29: aus Damaske [5]

Abgebildet ist tatsächlich der Korrelationsgrad, der aus dem Maximum der Korrelationsfunktion ermittelt wurde. (Definitionen siehe Anhang) Da anzunehmen ist, dass in allen Fällen gilt: $\tau=0$, entspricht dies dem Wert, den ein Korrelationsgradmesser liefert.

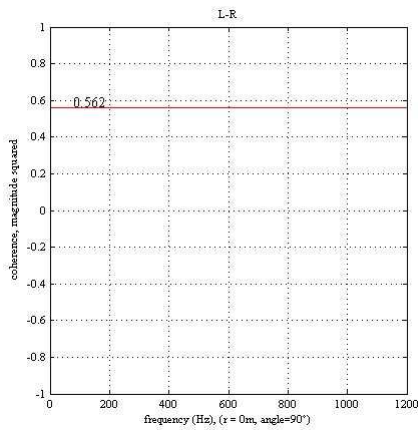


Abbildung 30: Kohärenzfunktion γ^2 : XY, 2 Nieren 90°, $b=0\text{cm}$

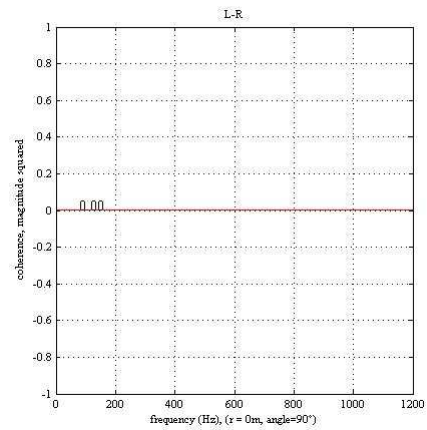


Abbildung 31: Kohärenzfunktion γ^2 : Blumlein, 2 Achten 90°, $b=0\text{cm}$

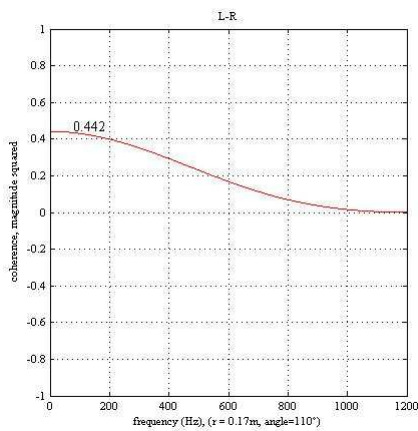


Abbildung 32: Kohärenzfunktion γ^2 : ORTF, 2 Nieren 110°, $b=17\text{cm}$

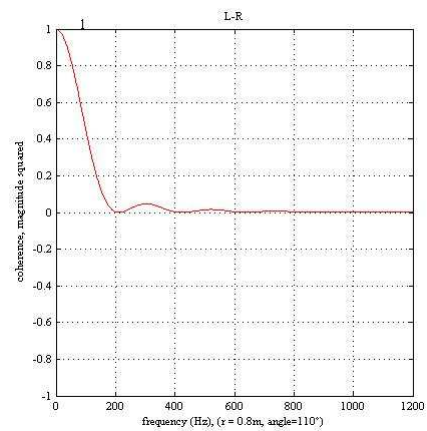


Abbildung 33: Kohärenzfunktion γ^2 : ab, 2 Kugeln, $b=80\text{cm}$

Ein paar Beispiele zum Unterschied korreliert – kohärent (korreliert heißt nun, dass die Korrelationsfunktion bei einem beliebigen τ ein Maximum von 1 hat):

	kohärent	Korreliert	Korrelationsgrad
Signal A: $\sin(\phi)$ Signal B: $\sin(\phi)$	ja	Ja	1
Signal A: $\sin(\phi)$ Signal B: $-\sin(\phi)$	ja	Ja	-1
Signal A: $\sin(\phi)$ Signal B: $\cos(\phi)$	ja	Ja	0
Signal A: s Signal B: s, um 10 ms verzögert	Ja	Ja	0
Signal A: ein Ton (mehrere Sinus-Schwingungen verschiedener Frequenzen) Signal B: ein anderer Ton (Sinus-Schwingungen der gleichen Frequenzen, aber unter- schiedlicher Phasenlagen)	Ja	Nein	0
Signal A/Signal B: zwei Mikrofone im diffusen Schallfeld mit Abstand r (ab-Stereofonie)	frequenzab- hängige Kohärenz	Nein	<1, siehe Abbildung 29

Da der Autor selbst Lernender auf diesem Gebiet ist, freut er sich über Verbesserungen und Korrekturen dieses Anhangs.