



UG

Bakkalaureatsarbeit

Institut für Elektronische Musik und Akustik

Universität für Musik und darstellende Kunst Graz

Andrés Ureta Staackmann

Studienrichtung: Elektrotechnik-Toningenieur V 033 213

Trinnov Surroundmikrofonanordnung

Lehrveranstaltung: Aufnahmetechnik 1 Seminar

Beurteilung: Dr. Alois Sontacchi

Juli 2010

Inhaltsverzeichnis

1. Einleitung	3
1.1 Einführung in Ambisonics	4
1.2 Das Prinzip	7
1.3 Dreidimensionale Beschreibung	7
1.4 Räumliche Abtastung	10
1.5 Räumliche Kodierung	11
2. Mikrofonanordnungen	14
2.1 Zeitliche Güte	14
2.2 Räumliche Güte	15
2.3 Ergebnisse	17
2.4 Zusammenfassung	20
3. Räumliche Auflösung bei gängigen merhkanaligen Aufnahmeverfahren	23
3.1 Qualität	25
3.2 Herangehensweise zu einer höheren räumlichen Auflösung	26
3.3 Eigenschaften	29
3.4 Implementierung	31
3.5 Optimierung	32
4. Zusammenfassung	38
Quellenverzeichnis	40
Anhang	42

1. Einleitung

Surroundmikrofonanordnungen wurden seit mehreren Jahren mit dem Ziel eingesetzt, Schallfelder mit einem mehrkanaligen System wiedergeben zu können. Der Surround Sound Ansatz ist aber im Gegenteil zu Ambisonics¹ durch das Wiedergabesystem auf zwei Dimensionen begrenzt.

Durch Festlegen einer Wiedergabeordnung soll eine optimale Mikrofonanordnung gefunden werden und somit die gewünschte räumliche Auflösung² erlangen. Um möglichst viele Informationen aus dem Schallfeld zu gewinnen und auf diese Weise eine hohe räumliche Auflösung bei der Wiedergabe zu ermöglichen, werden diejenige Mikrofonanordnungen eingesetzt, die Informationen aus vielen Raumrichtungen des Schallfeldes entnehmen. Wie es sich herausgestellt hat, lässt sich diese Aufgabe schwer lösen. Erstens sind viele Mikrofone nötig und zweitens entspricht die diskrete Ausführung von herkömmlichen Mikrofonen nicht dem Ideal. Kommerzielle Mikrofone können Richtcharakteristiken³ 0. und 1.Ordnung⁴ aufweisen (siehe [1]).

In dieser Arbeit wird eine Herangehensweise präsentiert, die auf eine mathematische Beschreibung des Schallfeldes basiert und eine dreidimensionale Darstellung dieses durch die Mikrofone ermöglicht. Mit dieser Technik ist es nun möglich, Mikrofone 0. und 1. Ordnung einzusetzen, diese frei im Raum zu positionieren und trotzdem eine qualitative räumliche Abtastung des Schallfeldes zu erreichen. Sie kann als eine Erweiterung des Ambisonics Prinzips gesehen werden. Der Unterschied liegt bei der Abtastung des Schallfeldes, welche nicht nur aus einem Punkt im Raum erfolgt (siehe [2] und [3]). Im Gegenteil zu anderen Verfahren, wo die Koinzidenz der Mikrofonen (vgl. Kap.3.1) neben dem Einsetzen von Kapseln höherer Ordnung vorausgesetzt wird, können mit dieser Herangehensweise Mikrofone verschiedener Richtcharakteristiken verwendet werden (vgl. Kap.3 bzw. Abb.3.1) und trotzdem eine qualitative räumliche

¹ Verfahren zur Aufnahme und Wiedergabe eines Schallfeldes, siehe Kap. 1.1

² Gibt an wie genau das Schallfeld im Raum abgetastet bzw. wiedergegeben werden kann, siehe Kap. 2.2

³ Die Richtcharakteristik im Polardiagramm beschreibt die Empfindlichkeit eines Mikrofons in Abhängigkeit vom Schalleinfallswinkel

⁴ *Mikrofone*, die nur auf den (ungerichteten) Schalldruck reagieren, werden als Druckempfänger *nullter Ordnung* bezeichnet, auch ungerichtete Mikrofone genannt

Auflösung des Schallfeldes erzielt werden. Die Mikrofone können also willkürlich im Raum positioniert werden⁵.

Vorge stellt wurde dieses Konzept von *Trinnov Audio*⁶ im Rahmen des „*Immersive Sound Project*“ bei der 114th AES Convention im Jahre 2003 in Amsterdam [4] sowie bei der 116th AES Convention im Jahre 2004 in Berlin [5].

Ziel der Arbeit ist nun dieses Verfahrensprinzip zu beschreiben und herzuleiten sowie die Systemeigenschaften und deren Anwendung für den Aufnahmebereich zu untersuchen.

1.1 Einführung in Ambisonics

Ambisonics ist ein mehrkanaliges Aufnahme- und Wiedergabesystem, das in den 70er Jahren von Michael Gerzon [6] entwickelt wurde. Es wird versucht eine dreidimensionale Reproduktion des Schallfeldes zu schaffen. Ein großer Vorteil dieses Verfahrens besteht darin, dass sowohl für die Wiedergabe wie für die Aufnahme keine feste Anzahl von Lautsprechern bzw. Mikrofonen vorgegeben ist.

Bei Ambisonics erster Ordnung wird die Information in vier Kanälen (W, X, Y, Z) kodiert, das sogenannte Ambisonics *B-Format*. Das dreidimensionale Schallfeld wird also in einem Punkt im Raum durch mehrere Mikrofone aufgenommen, üblich ist der Einsatz von Soundfield Mikrofonen (*SFM*)⁷. Der W Kanal entspricht das Ausgangssignal eines ungerichteten Druckempfängers, während die X, Y und Z Kanäle die gerichtete Komponente des Signals darstellen. Diese drei Signale entsprechen den Ausgangssignalen drei orthogonal geortete Mikrofone mit einer Achtercharakteristik.

⁵ Diese ist der Ausgangspunkt um eine gute Räumlichkeit zu schaffen. Der Frequenzbereich, in dem Schallquellen noch akkurat abgetastet werden können, wird durch die Geometrie der Mikrofonanordnung bestimmt (siehe Kap.1.5)

⁶ Arnaud Laborie, Rémy Bruno und Sébastien Montoya (Paris, Frankreich)

⁷ Die Anzahl der Mikrofonen ist frei wählbar und hängt von der Zielsetzung ab, bei SFM werden vier gerichtete Mikrofone in Form eines Tetraeders angeordnet

Dabei werden folgende Voraussetzungen an das Schallfeld gestellt

- das aufgenommene Schallfeld ist eine ebene Welle
- die Lautsprecher strahlen bei der Wiedergabe ebenfalls eine ebene Welle aus

Diese Bedingungen sind dann erfüllt, wenn sich die Schallquelle weit genug von den Mikrofonen befindet bzw. wenn der Zuhörer sich auch weit genug von den Lautsprechern befindet. Die Mikrofone liefern die Signale im *A-Format* und ein Kodierer transformiert diese ins *B-Format* (siehe [7]). Die Signale im *B-Format* basieren auf einer Zerlegung des Schallfeldes in vier Komponenten, *W* die ungerichtete Komponente und *X*, *Y* und *Z* dementsprechend die gerichtete Komponenten.

Für eine genaue Beschreibung zur Herleitung des *B-Formats* (siehe [2] und [6]). Ein anderer Vorteil dieses Verfahrens ist die Unabhängigkeit der Signalen vom Aufnahme- und Wiedergabesystem. Die Lautsprechersignale entstehen durch einen Decoder im Rekonstruktionsprozess. Dieser wird hingegen durch die Lautsprecherkonfiguration bestimmt.

Ausgehend von der Gleichung für die ebene Welle im Aufpunkt (Referenzwelle)

$$S(r, \phi) = P_\psi e^{ikr \cos(\phi - \psi)} \quad (\text{Gl.1})$$

erhält man nach einer Reihenentwicklung Gleichung 2. Mit dieser Gleichung kann eine ebene Welle aus einer Summation von (*N*) Lautsprechersignalen beschrieben werden. Wobei (*k*) für die Wellenzahl ($k = 2\pi / \lambda$) und (*r*) für die Entfernung des Aufpunktes vom Koordinatenursprung stehen. (*l*) stellt die Ordnung des Systems dar, bis welcher die Reihenentwicklung angenähert wird (siehe [3]).

$$S = \sum_{n=1}^N P_n j_0(kr) + \sum_{l=1}^{\infty} 2 i^l j_l(kr) \left(\sum_{n=1}^N P_n [\cos(l\phi_n) \cos(l\phi) + \sin(l\phi_n) \sin(l\phi)] \right) \quad (\text{Gl.2})$$

Dieses Verfahren wurde mehrmals erweitert (siehe [3], [8], [9] und [10]), trotzdem traten verschiedene Probleme auf⁸, die das Verfahren begrenzt haben. Man musste insbesondere die Koinzidenz der Mikrofonen berücksichtigen, Gegebenheit die bei diesem Prinzip vorausgesetzt wird (siehe Kap.3.1) aber nicht erfüllt werden kann. Bei Ambisonics geht man von einer Approximation aus und dadurch entstehen also erste Ungenauigkeiten in den Signalen. In Folge dessen verschlechtert sich die Lokalisationsschärfe der Schallquellen und der Sweet-Spot verkleinert sich beträchtlich. Das Hauptproblem stellt aber die Ordnung (l) des Systems dar, bis welcher die Reihe approximiert wird. Je größer die Ordnung, desto besser kann die Referenzwelle reproduziert werden. Mit der Ordnung nimmt die Anzahl der Übertragungskanälen sowie die Anzahl der notwendigen Lautsprechern zu. Für $l=0$ ergibt sich eine kugelförmige Richtcharakteristik (siehe Abb.3.4). Bei höheren Ordnungen entstehen Richtcharakteristiken, die nicht als solche verfügbar sind. Sie können unter Umständen aus einer Superposition von niedrigeren Ordnungen zusammengesetzt werden.

In dieser Arbeit wird ein Aufnahmeverfahren präsentiert, das auf ein dreidimensionales Modell des Schallfeldes basiert und auf die akkurate Wiedergabe von Schallquellen zielt. Es können nicht alle Probleme vollkommen gelöst werden, es soll eher eine alternative Lösung zur Aufgabenstellung vorgestellt werden.

Die Zielsetzung ist demnach so viele Richtungsinformationen⁹ wie möglich aus dem Schallfeld zu erhalten. Das Ergebnis ist eine Darstellung vom Schallfeld in einer dreidimensionalen Form, die mit Ambisonics kompatibel ist.

Das in dieser Arbeit präsentierte Verfahren wird nicht aus Ambisonics abgeleitet, da Schallereignisse nicht nur aus einem beliebigen Punkt im Raum betrachtet werden. Es kann aber wie eine Erweiterung von Ambisonics aufgefasst werden. Mit diesem Verfahren ist es möglich unter dem Einsatz von Mikrofonen mit verschiedenen Richtcharakteristiken und die

⁸ Es ist physikalisch unmöglich zwei Objekte (Mikrofone) am gleichen Punkt im Raum zu positionieren

⁹ Informationen über den Schalldruck aus verschiedenen Richtungen und Punkten im Raum

freie Positionierung dieser im Raum, Signale höherer Ordnung¹⁰ zu erzeugen.

1.2 Das Prinzip

Es wird von einem dreidimensionalen Modell des Schallfeldes ausgegangen. Dies entspricht dem realen Fall und ermöglicht es durch die Näherung der Reihenentwicklung, jedes Schallereignis so getreu wie möglich mit einer endlichen Anzahl an Signalen (die Mikrofonensignale) zu beschreiben (vgl. Kap.1.3). Es soll allein mit der Kenntnis der aufgenommenen Signale sowie der Position der Mikrofonen im Raum und deren Richtcharakteristiken eine genaue Repräsentation des unbekannten Schallfeldes abgeleitet werden. Das Prinzip ist in Abbildung 1.1 genau beschrieben.

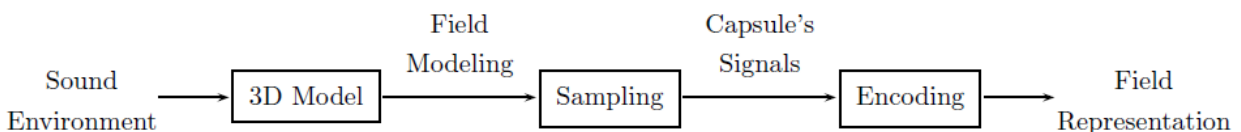


Abb.1.1:Das Prinzip

1.3 Dreidimensionale Beschreibung

In einem quellfreien Punkt¹¹ soll ein Schallfeld mit einer endlichen Anzahl von Signalen modelliert¹² werden. Für diese Aufgabe wird ein dreidimensionales Modell in Kugel-Koordinaten mit dem Namen „Fourier-Bessel-Zerlegung“ verwendet. Die verwendete Größe um das Schallfeld zu beschreiben ist der Schalldruck (skalare Größe). Aus der Funktion $P(r, \theta, \phi, t)$ erhält man nach der Fouriertransformation und weitere Annahmen (siehe Anhang) für den Schalldruck folgende Gleichung

¹⁰ Allgemein ein Set von orthogonalen Kugelflächenfunktionen

¹¹ Ein Volumen, wo sich die Mikrofonen befinden. Keine Schallquellen dürfen sich in diesem Volumen befinden

¹² Unter Modellierung ist eine Approximation des Schallfeldes gemeint

$$P(r, \theta, \phi, f) = 4\pi \sum_{l=0}^{\infty} i^l j_l(kr) \sum_{m=-l}^l p_{l,m}(f) Y_l^m(\theta, \phi) \quad (\text{Gl.3})$$

$Y_l^m(\theta, \phi)$Kugelflächenfunktionen

$j_l(kr)$sphärische Besselfunktionen

$p_{l,m}(f)$sphärisches Wellenspektrum

Die Fourier-Bessel Zerlegung¹³ liefert die richtungsunabhängigen Koeffizienten $p_{l,m}(t)$, wobei (l) und (m) ganze Zahlen sind und diese Bedingungen erfüllen, $l \geq 0$ und $-l \leq m \leq l$. Die Ordnung wird hier auch (l) genannt, (m) stellt den Grad dar. Abbildung 1.2 zeigt die Kugelflächenfunktionen für verschiedene Werte von (m) und (l) .

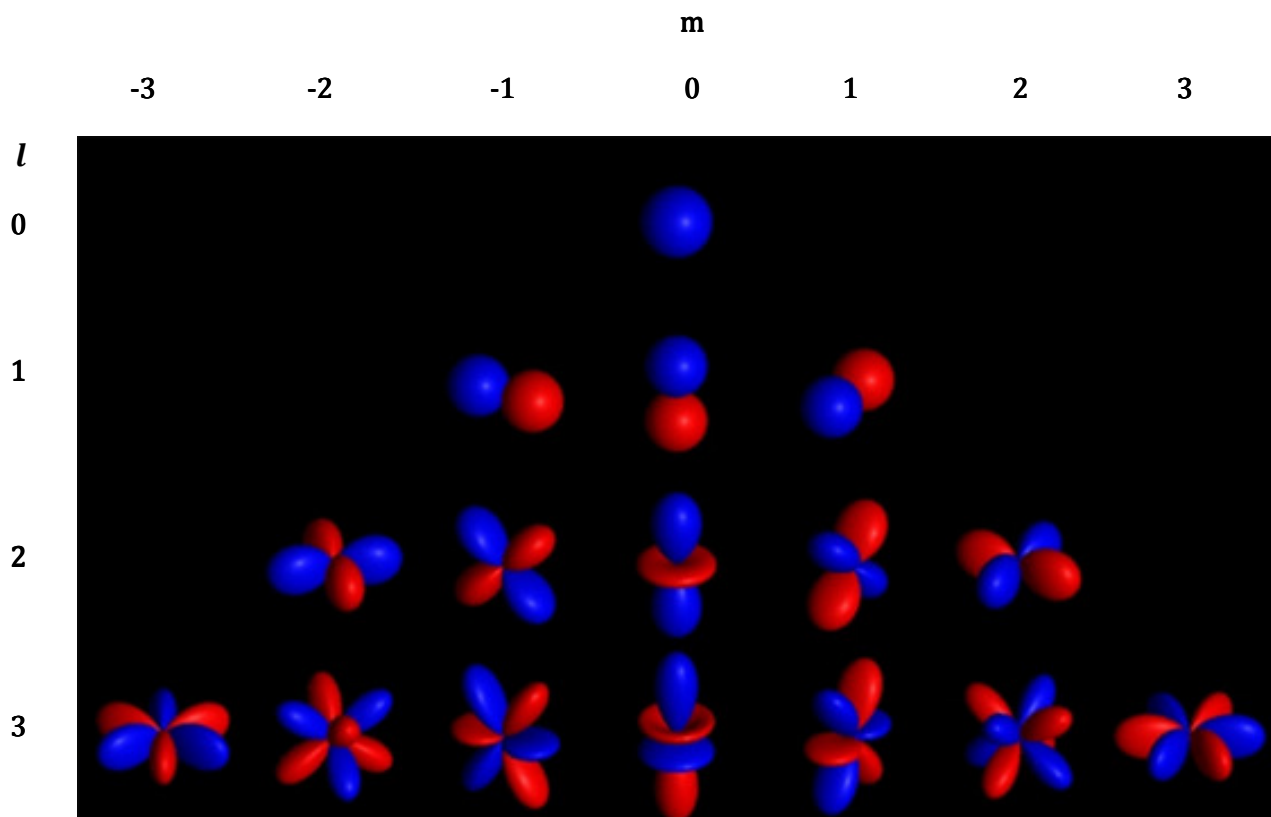


Abb.1.2: Dreidimensionale Ansicht der Kugelflächenfunktionen

¹³ Die Fourier-Bessel Zerlegung beschreibt ein akustisches Schallfeld aus einer Überlagerung von elementaren akustischen Wellenfronten mit verschiedenen Richtungsinformationen (spektrale Repräsentation eines akustischen Schallfeldes) siehe [15]

Mit den Kugelkoordinaten (siehe Abb.1.3) ist die Periodizität vom Schalldruck durch θ und ϕ ersichtlich. Eine Beschreibung des Schallfeldes mit einer endlichen Anzahl von Koeffizienten ist nun möglich, wenn die Reihe (siehe Gl.3) bis zu einer bestimmten Ordnung (l) approximiert wird. Der Unterschied zu Ambisonics liegt darin, dass beide Anteile berücksichtigt werden, sowohl der richtungsabhängige Teil $Y_l^m(\theta, \phi)$ als auch der radiale Teil $j_l(kr)$. In den sphärischen Besselfunktionen ist das Argument (kr) enthalten. Diese Funktionen sind vom Vielfachen der Wellenzahl abhängig und besitzen die Ordnung (l). Es werden die Wellenlänge (λ)¹⁴ und der Abstand (r) im Vergleich zu einander gesetzt.

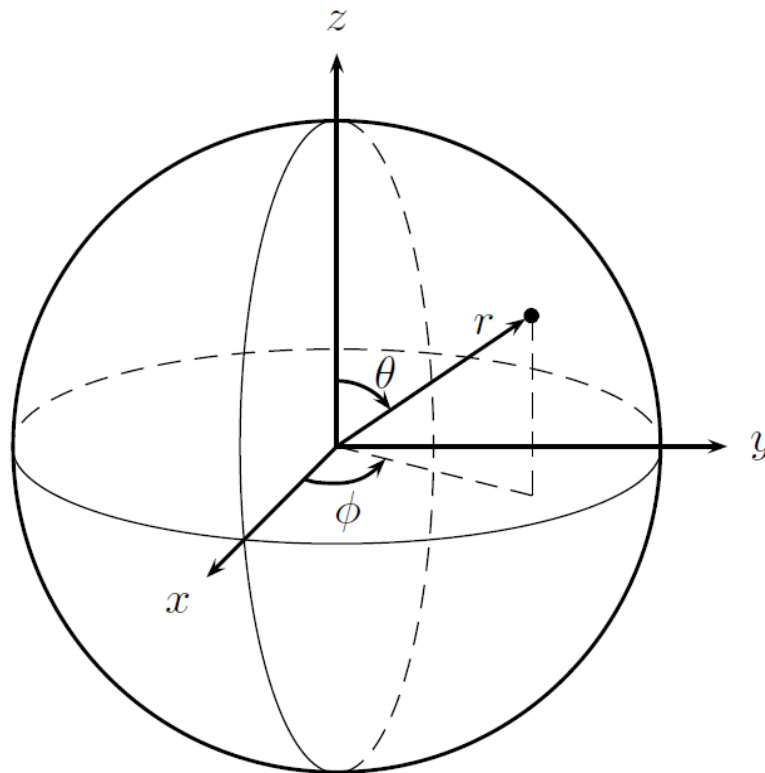


Abb.1.3: Kugel Koordinaten System

¹⁴ $\lambda = f^{-1} * c$, wobei c die Schallgeschwindigkeit (hängt vom Medium ab) und f die Frequenz ist

1.4 Räumliche Abtastung

Es ist bereits bekannt, dass um eine gute räumliche Abbildung von Schallquellen zu erreichen, mehrere Lautsprecher bei der Wiedergabe zum Einsatz kommen müssen. Dies ist nur dann realisierbar, wenn bei der Aufnahme auch mehrere Mikrofone verwendet wurden, die Richtungsinformationen aus dem Schallfeld liefern. Die Mikrofone können verschiedene Richtcharakteristiken aufweisen und können auch frei im Raum positioniert werden. Mit einer fixen Mikrofonanordnung kann nun ein Schallereignis aufgenommen werden. Das Schallfeld wird also an bestimmten Punkten abgetastet.

Das Eingangssignal der Mikrofone ist das Schallfeld und der Ausgang dieser sind also die Mikrofonsignale. Der Vorgang ist in Abbildung 1.4 dargestellt. Die Signale beschreiben das Schallfeld nur dort, wo die einzelne Mikrofone positioniert wurden. Eine räumliche Abtastung ist das räumliche Äquivalent von einer zeitlichen Abtastung.

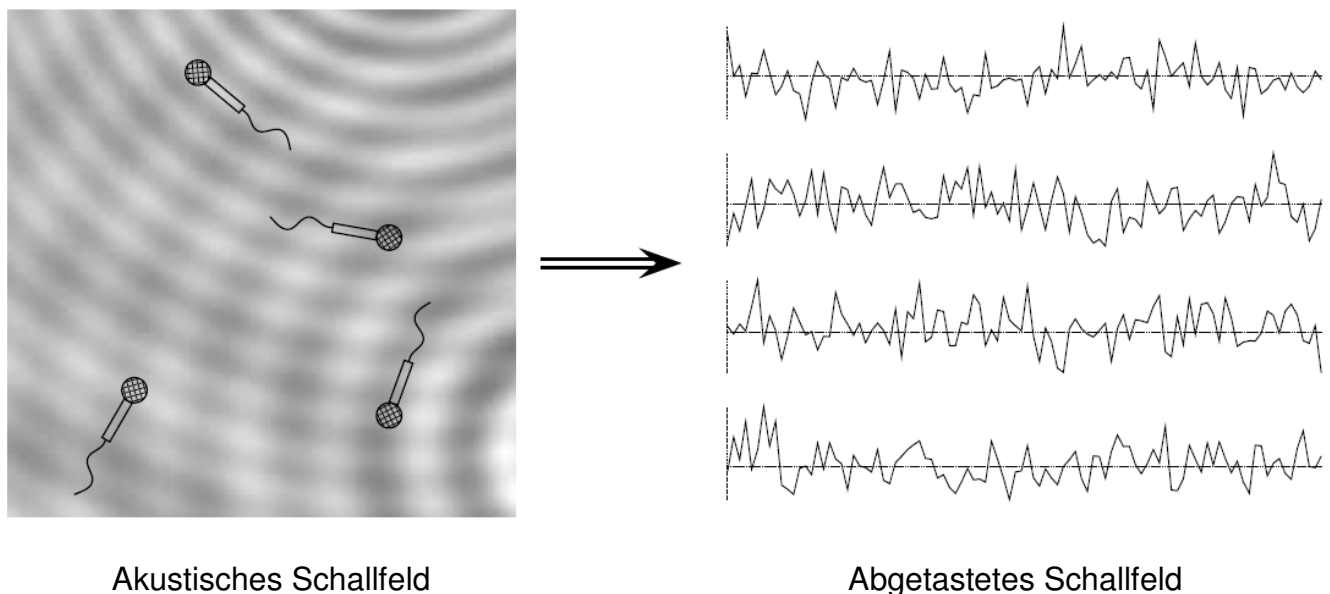


Abb.1.4: Prozess des Abtastens

Um diesen Begriff besser verstehen zu können, kann man sich ein zeitliches Sample wie eine Information aus einem zeitlichen Signal (z.B. ein audio Signal) vorstellen und gleichzeitig ist dieses zeitliche Signal eine Information des aufgenommenen Schallfeldes.

Das räumliche Äquivalent zum Shannon'schen Abtasttheorem soll zunächst eingeführt werden. Um spektrales Aliasing¹⁵ zu verhindern, muss im dreidimensionalen Fall folgende Beziehung eingehalten werden

$$(L + 1)^2 \leq N \quad (\text{Gl.4})$$

Dabei ist (N) die Anzahl der verwendeten Mikrofonen und (L) die Ordnung des Systems. Wenn (N) Mikrofone zur Verfügung stehen, dann wird die Ordnung (L) so gewählt, dass Gl.4 eingehalten wird. Mit der Abtastbeziehung (Gl.5), können die Signale berechnet werden, die die Mikrofone durch Abtastung des Schallfeldes erzeugen würden. Man erhält eine lineare Matrix Beziehung zwischen den Mikrofonensignalen (c) und das sphärische Wellenspektrum (p) im Frequenzbereich.

$$c = B * p \quad (\text{Gl.5})$$

(B) stellt die Sampling Matrix dar. Diese hängt nur von der Mikrofonanordnung ab, sie beschreibt also wie das Array Informationen aus dem Schallfeld abtastet und extrahiert.

1.5 Räumliche Kodierung

Ziel von diesem Prozess ist das dreidimensionale Schallfeld aus den Signalen der Mikrofonen zu berechnen. Das Array liefert die Mikrofonensignale (c) und mit der Sampling Matrix (B) kann im Frequenzbereich das sphärische Wellenspektrum durch Inversion obiger Beziehung berechnet werden.

$$p = B^{-1} * c \quad (\text{Gl.6})$$

¹⁵ Unter spektrales Aliasing versteht man den Ortungsfehler bei der Wiedergabe von Schallquellen

Dadurch entstehen stets mehrere Probleme

- die Sampling Matrix (B) ist nicht immer quadratisch¹⁶ und infolgedessen existiert nicht immer die Inverse (B^{-1})
- die reale Position der Mikrofonen weicht von der angenommenen Position ab
- die Kennlinie realer Mikrofonen weicht ebenso vom Ideal ab
- jedes Mikrofon besitzt ein Eigenrauschen

Durch allgemeine Inversionsmethoden kann das erste Problem gelöst werden (z.B. durch Bildung der Pseudoinverse bei der Lösung des Systems). Werden die andere Probleme nicht berücksichtigt, dann können keine brauchbare Ergebnisse erreicht werden (u.a. schlechte Abbildung von Schallquellen bzw. Verkleinerung des Sweet-Spots). Die gesuchte Richtungsinformationen ergeben sich aus den Unterschieden der Mikrofonensignalen (c). Im tieffrequenten Bereich, wo die Wellenlänge (λ) bereits mehrere Meter beträgt, stellt dies ein Problem dar. Die Signale werden mit einer kleiner werdenden Mikrofonanordnung sowie für tiefe Frequenzen ähnlicher zu einander (die Phaseninformationen der Mikrofonensignalen unterscheiden sich wenig von einander). Aus diesem Grund und mit dem Ziel eine akkurate Räumlichkeit zu schaffen, muss dieser Faktor beachtet werden.

Um einen Kompromiss zwischen räumlicher Auflösung und der Verstärkung von Rauschen¹⁷ zu finden, soll der Parameter (μ) eingeführt werden. Durch eine direkte Inversion der Sampling Matrix (B), also $\mu = 1$, wird der Einfluss von Rauschen in den berechneten Signalen größer. Je kleiner der Wert von (μ), desto kleiner wird auch der Einfluss von Rauschen, dafür verschlechtert sich dementsprechend die räumliche Auflösung. Dabei gilt

$$0 \leq \mu \leq 1 \quad (\text{Gl.7})$$

¹⁶ Eine Matrix ist quadratisch wenn sie die Dimensionen ($n \times n$) besitzt, dabei ist (n) eine ganze Zahl und diese Bedingung gilt $B * B^{-1} = I$. (I) ist die Einheitsmatrix

¹⁷ Ungenauigkeiten bei der Inversion von Matritzen (z.B. Division durch sehr kleine Werte) führen zu Ungenauigkeiten bei der Wiedergabe von Schallquellen (spektrales Aliasing)

Der Prozess der Kodierung erhält man so

$$\hat{p} = E * c \quad (\text{Gl.8})$$

wobei (E) die Kodier Matrix ist und (c) wiederum die Mikrofon-signale enthält. ($\hat{p}$) stellt hier das berechnete sphärische Wellenspektrum dar. Die Kodiermatrix lässt sich so berechnen

$$E = \mu B^T (\mu B B^T + (1 - \mu) I)^{-1} \quad (\text{Gl.9})$$

(I) ist die Identitätsmatrix oder auch Einheitsmatrix genannt. Jeder Koeffizient der Matrix (E) ist eine frequenzabhängige Funktion. Gleichung 9 kann in der Literatur mit einem Normalisierungsfaktor bei der Lösung der Inversion gefunden werden. Die Genauigkeit, mit der die Koeffizienten vom akustischen Feld abgebildet werden, könnte mit diesem Faktor festgestellt werden.

Der große Vorteil besteht darin, dass die räumliche Auflösung für jede Raumrichtung eingestellt werden kann und irreguläre Anordnungen optimiert werden können¹⁸. Für eine eine Ordnung (L) und eine Mikrofonanordnung bestehend aus (N)-Mikrofonen ergeben sich $N * (L + 1)^2$ Filter. Das Kodier Schema ist in Abbildung 1.4 dargestellt.

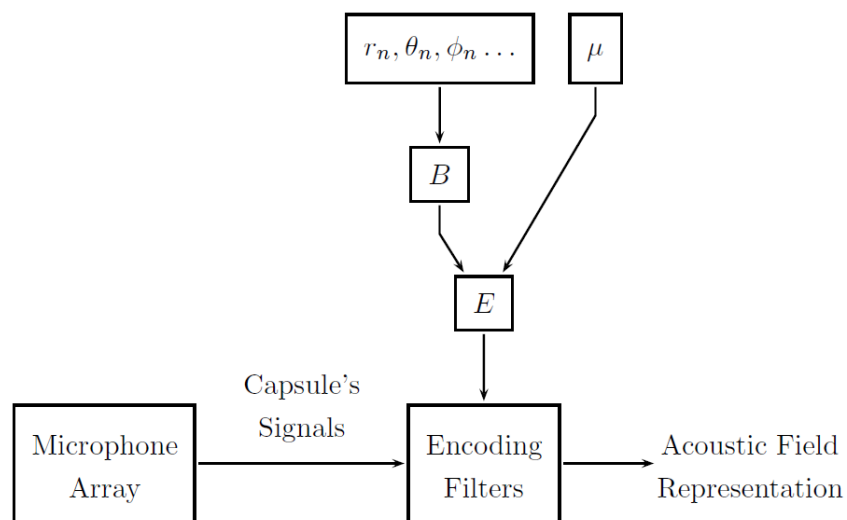


Abb.1.4: Kodier Schema

¹⁸ Die räumliche Auflösung kann anwendungsspezifisch eingestellt werden. Bei einer Kugelfläche kann z.B. 5.Ordnung für die vordere Fläche, zu den Seiten 3.Ordnung und im hinteren Bereich 1.Ordnung gewählt werden.

2. Mikrofonanordnungen

Als nächstes werden verschiedene Mikrofonanordnungen näher betrachtet. Das Ziel ist viele Richtungsinformationen zu gewinnen und diese auch in einer mit Ambisonics kompatiblen Form zu liefern. Dies ist möglich, weil Ambisonics auch auf der Theorie der Kugelflächenfunktionen basiert.

Die mit dieser Herangehensweise berechneten Signalen sind auf diese Weise auf ein Ambisonicssystem anwendbar. Bei diesen Untersuchungen wurden zwei Arten von Güten im Bezug auf die Mikrofonanordnungen berücksichtigt.

2.1 Zeitliche Güte

Die zeitliche Güte gibt die Qualität des berechneten Schallfeldes an. Es wird ein Vergleich zwischen dem Signal, das von einem idealen Mikrofon geliefert wird, und dem an der gleichen Stelle berechneten Signal. Parameter sind hier die Verstärkung des Eigenrauschens der Mikrofonen sowie die Verzerrung des Signals.

Die zeitliche Güte hängt eng mit den Eigenschaften der Mikrofonen zusammen. Der Wert von (μ) bestimmt nicht nur die Qualität der räumlichen Auflösung, dieser legt auch den Einfluss vom Eigenrauschen der Mikrofonen auf den Signalen fest. Dies kann durch das Kodierverfahren erklärt werden. Beim Kodierprozess handelt es sich von einer linearen Berechnung von Matrizen und dadurch entstehen keine weitere Verzerrungen. Andererseits kann unter Umständen das Eigenrauschen der Mikrofonen verstärkt werden.

Für die in dieser Arbeit präsentierten Versuchen wurde $\mu = 0,9$ gewählt. Der Grund ist durch das gesetzte Ziel klar; es soll eine hohe räumliche Auflösung erreicht werden.

2.2 Räumliche Güte

Die räumliche Auflösung beschreibt wie genau ein dreidimensionales Schallfeld dargestellt werden kann. Es werden das berechnete und das reale Schallfeld im Vergleich zu einander gesetzt. Argumente sind die räumliche Verzerrung und der Quellpositionierungsfehler.

Die räumliche Güte¹⁹ hängt hauptsächlich von der Sampling Matrix (B) ab. Tatsächlich bestimmt diese Matrix wie eine Mikrofonanordnung Informationen aus dem Schallfeld entnimmt. Sollte eine Mikrofonanordnung keine Informationen über einen bestimmten Koeffizient messen können, dann ist es auch unmöglich das Schallfeld mit diesem Koeffizient zu beschreiben. Folglich verschlechtert sich die räumliche Auflösung.

Um diesen Fall besser verstehen zu können wird in Abbildung 2.2 als Beispiel eine Mikrofonanordnung präsentiert. Sie besteht aus 24 Mikrofonen, die auf der *X-Achse* im Abstand von -10 cm und 10 cm positioniert sind. In Abbildung 2.3 wird die Sampling Matrix (B) 3.Ordnung für eine Frequenz von 1kHz dargestellt. Aus Abbildung 2.1 sind die Amplituden der Matrix Koeffizienten im Bezug zur höchsten vorkommenden Amplitude zu entnehmen.

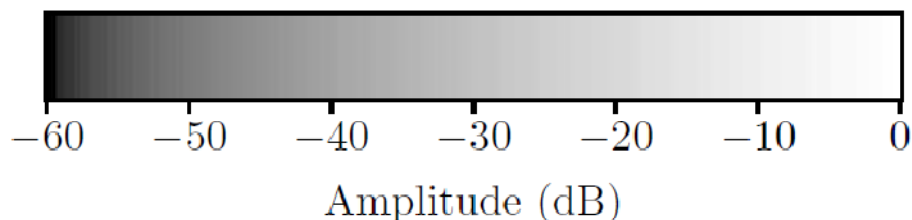


Abb.2.1: Kodierung der Amplituden der Sampling Matrix (B) in Abb. 2.3

¹⁹ Hier ist die Qualität der räumlichen Abbildung von Schallquellen bzw. die Qualität der Richtungsinformationen gemeint

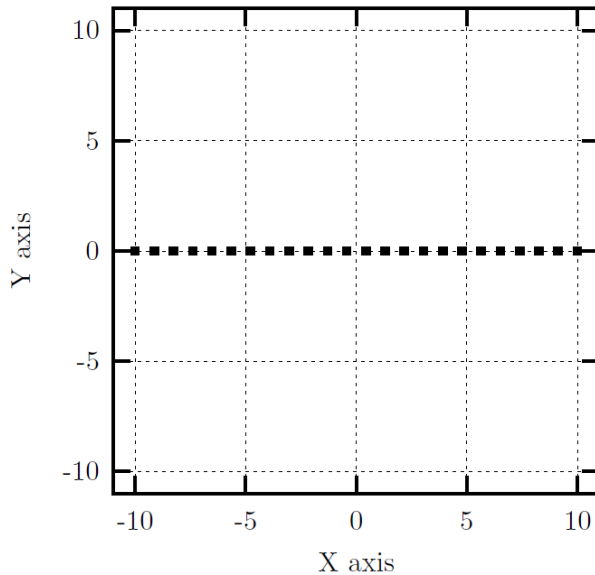


Abb.2.2: Lineare Mikrofonanordnung

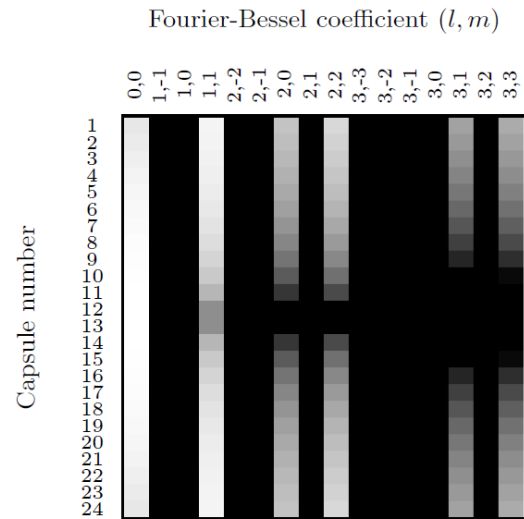


Abb.2.3: Sampling Matrix (B) der linearen Mikrofonanordnung in Abb. 2.2

Abbildung 2.3 zeigt für welche Koeffizienten die Mikrofonanordnung keine Informationen aus dem Schallfeld entnehmen kann (schwarze Spalten z.B. $(1,-1)$, $(1,0)$ oder $(3,-3)$). Die Mikrofonanordnung misst also keine Informationen aus der Y Komponente. Bei anderen Spalten wie $(3,1)$ ergibt sich ein schwarzes Feld für sämtliche Mikrofone. Das Schallfeld kann also nur durch bestimmte Koeffizienten beschrieben werden.

Um die räumliche Güte einer Mikrofonanordnung beschreiben zu können, wird ein neues Maß eingeführt, der räumliche Signalstörabstand $(p_s)^{20}$. Damit kann untersucht werden, wie viel Schall richtig im Raum abgebildet wird. Das „*Signal*“ ist hier als gut abgebildeter Schall und das „*Rauschen*“ als schlecht abgebildeter Schall im Raum zu verstehen.

$$p_s(dB) = -10 \log_{10} |diag[(EB) - I]^T * (EB - I)| \quad (Gl.10)$$

$diag(X)$ ist die Diagonale der Matrix X .

²⁰ Auch als „Spatial Signal to Noise Ratio“ in englischer Literatur zu finden

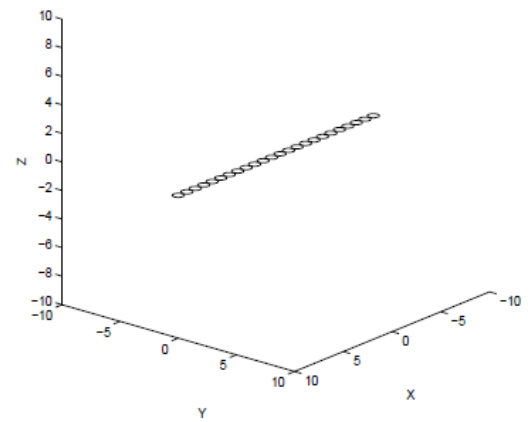
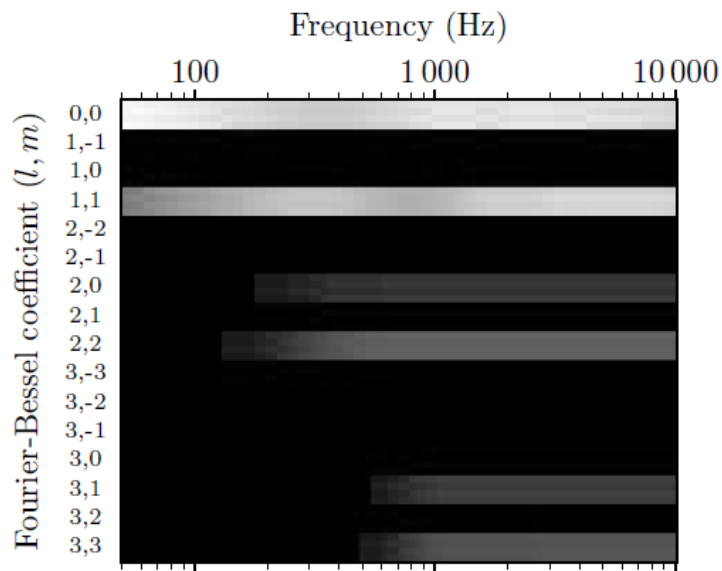
2.3 Ergebnisse

Es wurden lineare, zirkulare, kugelförmige und irreguläre Mikrofonanordnungen auf räumlichen Signalstörabstand untersucht. Die besten Ergebnisse erzielten allerdings letztere Anordnungen, die aus 24 in einem Sphärevolumen irregulär positionierten Mikrofonen mit einer kugelförmigen Richtcharakteristik bestehen.

Bei den anderen Anordnungen wird der *räumliche* Signalstörabstand für mehrere Frequenzen gleich Null. Dies kann durch die richtungsunabhängigen Funktionen $p_{l,m}(f)$ (siehe Gl.3) erklärt werden. Wenn diese Funktionen für eine gewisse Frequenz $f = f_0$ gleich Null werden, dann heben sich die Elemente $B_{n,l,m}$ der Sampling Matrix (B) für die Ordnung $l = l_0$ bei allen Mikrofonen auf. Es können also keine Informationen aus den richtungsunabhängigen Koeffizienten der Ordnung (l_0) bei der Frequenz (f_0) vom Schallfeld gewonnen werden. Wie es ersichtlich ist, nimmt der räumliche Signalstörabstand ab.

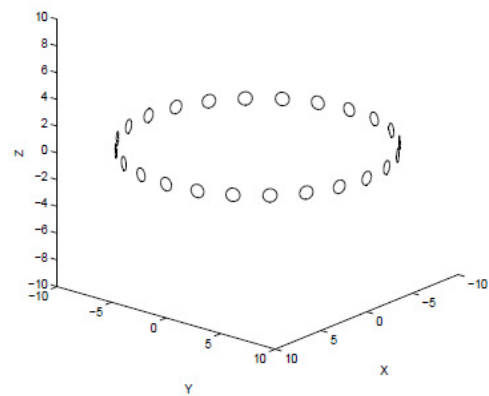
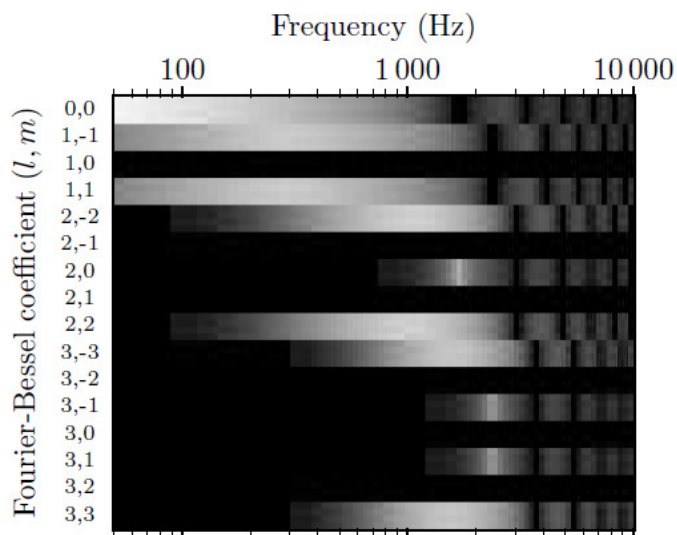
In Abbildungen 2.4 bis 2.7 ist der *räumliche* Signalstörabstand (siehe Gl.10) für verschiedene Mikrofonanordnungen dargestellt. Für diese Messungen wurden nur Mikrofone mit einer kugelförmigen Richtcharakteristik eingesetzt. Würde man hingegen Mikrofone mit einer *Nieren* Richtcharakteristik²¹ verwenden, dann ist nur eine kleine Verbesserung im tieffrequenten Bereich sowie für Koeffizienten 1.Ordnung feststellbar. Was hier als einen 'tieffrequenten Bereich' bezeichnet wird, wird nicht allein durch die Frequenz bestimmt, die Größe der Mikrofonanordnung spielt dabei eine wesentliche Rolle. Eine exakte Frequenz anzugeben ist aus diesem Grund und vor allem für diese Betrachtungen nicht von größer Relevanz.

²¹ Mikrofon dessen Polardiagramm die Form einer Niere besitzt (siehe [1])



$$\langle \rho_s \rangle = 0.97 \text{ dB}$$

Abb.2.4: Räumlicher Signalstörabstand für eine lineare Mikrofonanordnung



$$\langle \rho_s \rangle = 1.66 \text{ dB}$$

Abb.2.5: Räumlicher Signalstörabstand für eine zirkulare Mikrofonanordnung

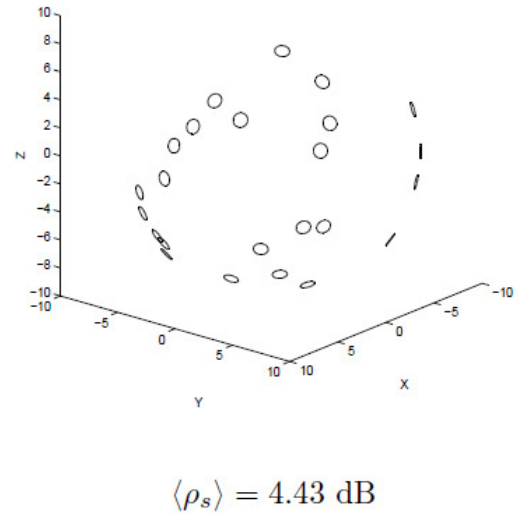
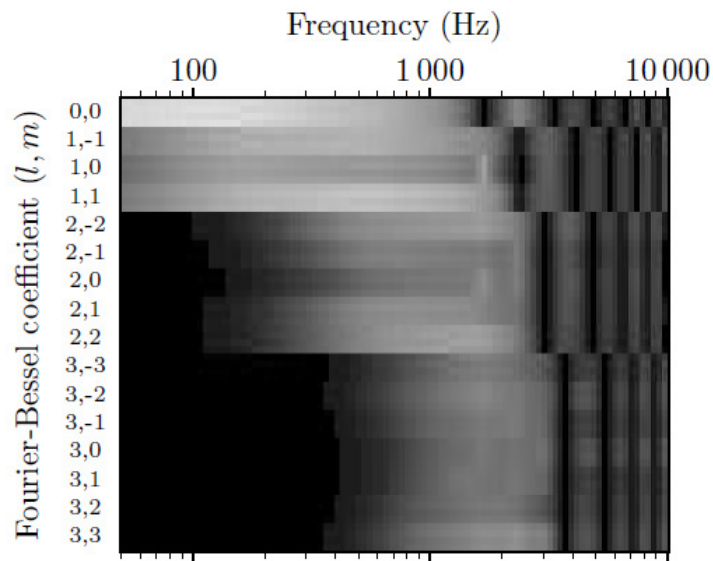


Abb.2.6: Räumlicher Signalstörabstand für eine kugelförmige Mikrofonanordnung

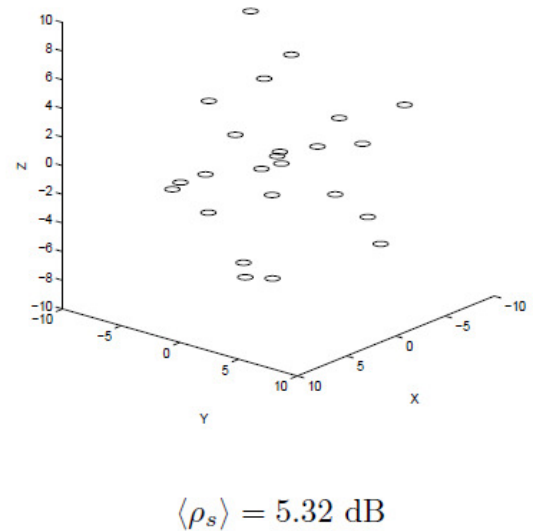
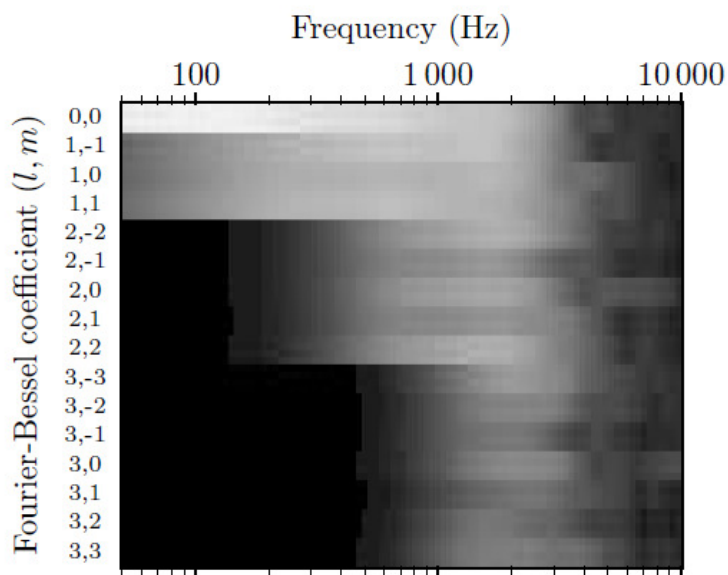


Abb.2.7: Räumlicher Signalstörabstand für eine dreidimensionale Mikrofonanordnung (irregulär)

2.4 Zusammenfassung

Für $\mu \rightarrow 1$ verbessert sich die räumliche Güte (die Inversion der Sampling Matrix (B) ist genauer, $E = B^{-1}$, vgl. Gl.9). Gleichzeitig ist der Einfluss vom Rauschen in den Signalen größer. Ganz ersichtlich ist die Tatsache, dass irreguläre Mikrofonanordnungen bessere Ergebnisse bzw. einen besseren *räumlichen* Signalstörabstand liefern wie reguläre Anordnungen. Dies resultiert aus der Geometrie der Mikrofonanordnung. Bei regulären Anordnungen sind die Mikrofone symmetrisch angeordnet, die Mikrofonanordnung entnimmt also Informationen aus dem Schallfeld an symmetrischen Punkten im Raum (siehe Mikrofonanordnung in Abb.2.4). Bei irregulären Anordnungen ergibt sich diese Symmetrie nicht mehr. Die Ergebnisse (siehe Kap.2.3) verdeutlichen wie stark die Räumlichkeit durch die Geometrie der Mikrofonanordnung bestimmt wird. In Abb.2.7 ist dieses periodisches Muster von schwarzen Linien (siehe Abb.2.6) nicht mehr zu finden, obwohl beide Anordnungen dreidimensional sind. Ist ein besseres Ergebnis im hochfrequenten Bereich erwünscht, dann sollen kleinere Anordnungen eingesetzt werden, sodass die Mikrofone näher zum Ursprung positioniert sind. Ist das Ziel ein gutes Ergebnis für tiefere Frequenzen zu erreichen, dann soll dementsprechend eine größere Mikrofonanordnung verwendet werden.

Mit dieser Herangehensweise für eine *Surround Sound* Aufnahme kann eine Ambisonics Mikrofonanordnung 3.Ordnung unter Anwendung von 24 Mikrofonen mit einer kugelförmigen Richtcharakteristik erzeugt werden. Bei der Mikrofonanordnung in Abbildung 2.7 befinden sich alle Mikrofone innerhalb eines kugelförmigen Volumens mit einem Radius von 20 cm.

In Abbildung 2.8 und 2.9 ist der Unterschied zwischen einer Mikrofonanordnung 3.Ordnung und 1.Ordnung prägnant dargestellt. Die wichtige Rolle von Direktivitäten höherer Ordnung ist ersichtlich. In Abbildung 2.9 kann man die zwei Schallquellen und sogar die ersten Reflexionen vom Boden ganz genau erkennen. Die Achsen stellen die

Richtungen im Raum dar, wobei die *Y-Achse* die Elevation und die *X-Achse* den Azimutwinkel repräsentieren. Die Amplitude ist in Abbildung 2.10 dargestellt.

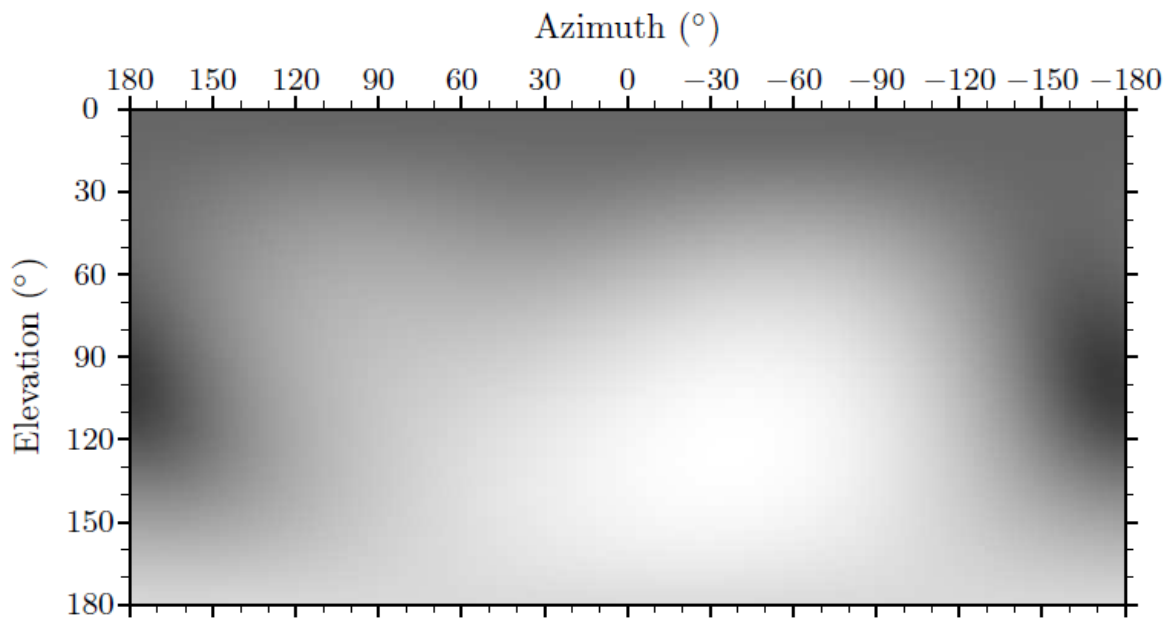


Abb.2.8: Darstellung der Amplituden einer Mikrofonanordnung 1.Ordnung

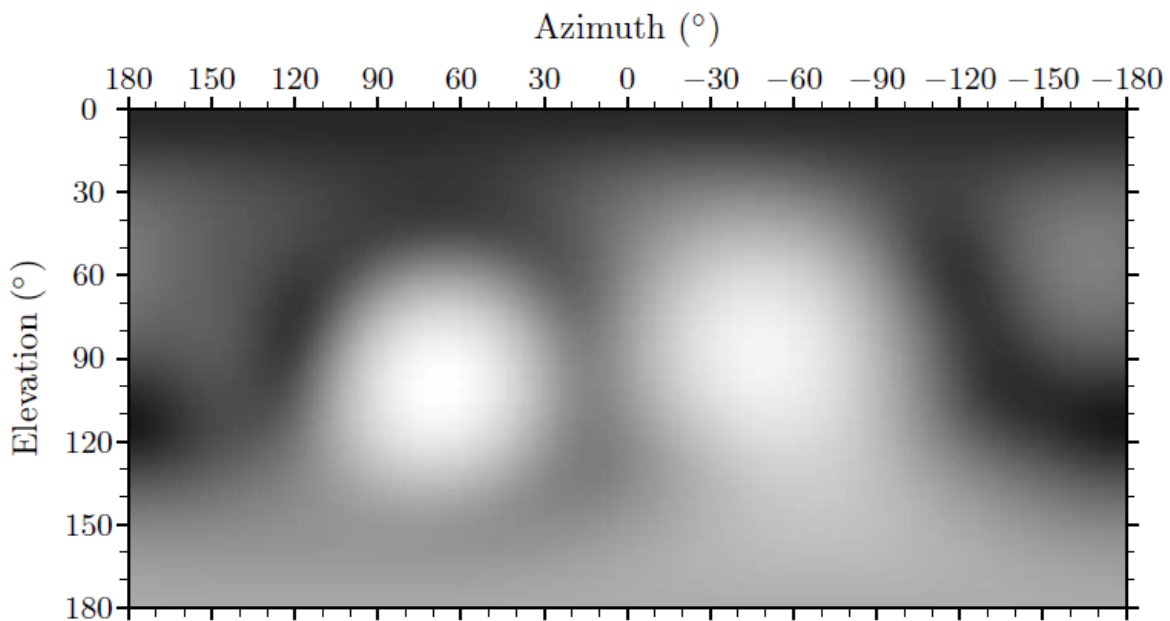


Abb.2.9: Darstellung der Amplituden einer Mikrofonanordnung 3.Ordnung

Die Mitte des Bildes entspricht der frontalen Richtung $(\phi, \theta) = (0^\circ, 90^\circ)$ und die Richtung $(\phi, \theta) = (-90^\circ, 90^\circ)$ entspricht der frontalen rechten Seite.

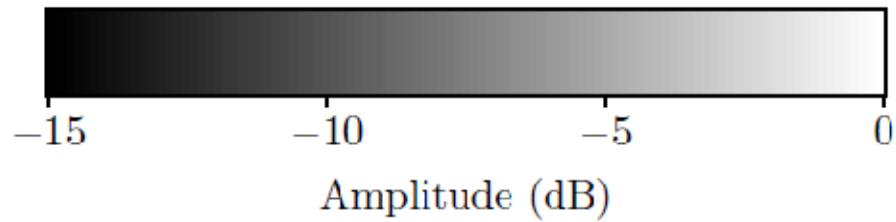


Abb.2.10: Amplitudendarstellung bei Abb.2.8 und 2.9 mittels einer Farbenskala

3. Räumliche Auflösung bei gängigen mehrkanaligen Aufnahmeverfahren

Die Mehrkanaltechnik wurde im Laufe der Jahre mehrmals erweitert²². Von Interesse für die in dieser Arbeit präsentierten Betrachtungen sind die Entwicklungen in Richtung Surround Sound. Abbildung 3.1 stellt eine Methode dar, bei der jedes Mikrofon einem Lautsprecher zugeordnet wird. Die akustische Umgebung wird durch die Laufzeit- und Intensitätsdifferenzen geschaffen. Man spricht hier von einer Laufzeits- bzw. einer Intensitätsstereofonie.

Diese Arten von Stereophonien zielen auf optimale Positionierungen der Mikrofonen zu finden sowie die adäquate Richtcharakteristiken einzusetzen. Ziel ist das dreidimensionale Schallfeld, das der Hörer durch interaurale Zeit- und Intensitätsdifferenzen²³ wahrnimmt, nachzubilden. Unter Berücksichtigung von psychoakustischen Modellen²⁴ kann ein Optimum gefunden werden, welche das gegebene Schallfeld bestmöglich abtastet. Trotzdem ergeben sich Ungenauigkeiten allein durch den Charakter dieses Verfahrens

- Abweichung zwischen realer und angenommener Position der Mikrofonen
- psychoakustische Modelle basieren auf Versuchsreihen, die mit einer begrenzten Anzahl an Versuchspersonen durchgeführt wurden. Sie sind also keine Gesetze die allgemein gelten (die Höreigenschaften sind vom Mensch zu Mensch immer unterschiedlich).
- jedes Mikrofon besitzt eine eigene Kennlinie. Beim einsetzen von Mikrofonpaaren (zwei Mikrofone die gleich abgestimmt sind bilden ein Mikrofonpaar) wird das Eingangssignal durch die Kennlinie beeinflusst. Am Ausgang eines Mikrofonpaares entstehen zwei verschiedene Ausgangssignale

²² Für eine detaillierte Beschreibung der verschiedenen Stereo Aufnahmeverfahren (siehe [1] bzw. [12])

²³ Interaurale Zeit- und Intensitätsdifferenzen sind die Zeit- und Intensitätsdifferenzen, die beim menschlichen Hörvorgang entstehen

²⁴ Ein psychoakustisches Modell beschreibt die Höreigenschaften der Menschen, z.B. Frequenzgruppenbildung, Hörschwelle, Maskierungseffekte, etc. (für eine genaue Beschreibung siehe [1])

In Abbildung 3.1 ist dieses Verfahren genau dargestellt. Man sieht fünf Mikrofone mit unterschiedlichen Richtcharakteristiken, die zu fünf Lautsprechern angeschlossen sind. Es entspricht einem 5.0 System (vgl. [1]), welche aus einem Centerkanal, einem Kanal links und rechts sowie zwei rückwärtigen Kanälen besteht. Die Mikrofone sind auch anders positioniert und wurden in unterschiedlichen Größen abgebildet.

Von großer Bedeutung ist die Größe der Mikrofonen in den Abbildungen sowie die Länge der Pfeile bei den Wiedergabesystemen. In beiden Abbildungen befindet sich die Schallquelle rechts. Trotzdem strahlen in Abb.3.1 beide hintere Lautsprecher (links und rechts) mit der gleichen Intensität ab. Bei diesen Systemen muss die Lücke im hinteren Bereich gefüllt werden und aus diesem Grund sind beide Mikrofone mit der Kugel Richtcharakteristik größer wie die gerichtete Mikrofone abgebildet.

Gesucht ist ein System wie in Abb.3.2 dargestellt, dabei ist die Anzahl der Lautsprecher variabel. Es werden trotzdem fünf Lautsprecher abgebildet um den Vergleich zum 5.0 System prägnant darzustellen.

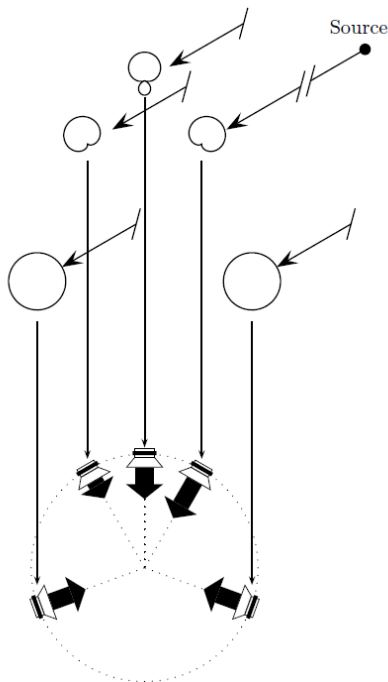


Abb.3.1: Übliche 5.0 mehrkanalige Mikrofonanordnung

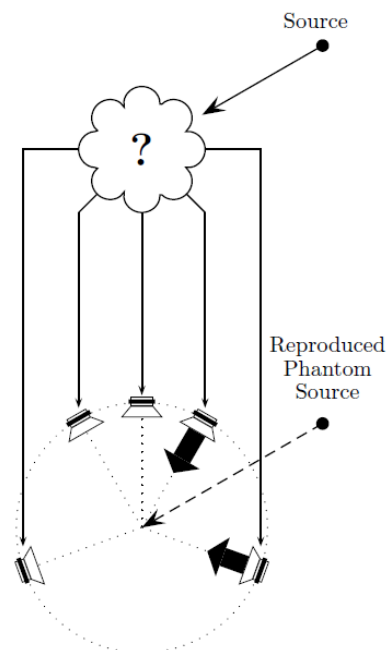


Abb.3.2: Ideale 5.0 mehrkanalige Mikrofonanordnung

3.1 Qualität

Viele Verfahren nutzen die Charakteristiken und die Positionierung von Mikrofonen aus, um eine räumliche Atmosphäre zu schaffen. Jedes Mikrofon wird im Endeffekt zu einem Kanal angelegt, welche wiederum zu einem Lautsprecher angeschlossen ist. Die *interaurale* Laufzeit- und Intensitätsdifferenzen, die beim natürlichen Vorgang des Hörens entstehen, werden durch Laufzeit- und Intensitätsdifferenzen an den Mikrofonen nachgebildet. Man kann zwischen zwei verschiedenen Verfahren unterscheiden

- Koinzidente Verfahren: Bei diesen Verfahren werden gerichtete Mikrofone verwendet, welche an einem gleichen Punkt im Raum platziert sind. Das Audio-Bild ist stabil und Schallquellen können sehr genau lokalisiert werden. Der Sweet-Spot ist breit und die Räumlichkeit wird nur mit Hilfe von Intensitätsunterschieden (ΔI) geschaffen. Es ist zu beachten, dass die Mikrofone physikalisch nicht am gleichen Punkt platziert werden können. Dadurch entstehen bei der Wiedergabe bereits erste Ungenauigkeiten²⁵.
- Nicht koinzidente Verfahren: Die Mikrofone befinden sich getrennt von einander im Raum. Die Räumlichkeit wird hauptsächlich durch Laufzeitunterschiede (ΔT) geschaffen. Das Audio-Bild ist weniger stabil und der Sweet-Spot ist kleiner, dafür erreicht man eine bessere Tiefe. Bei diesen Verfahren muss auch zwischen zwei Fällen unterschieden werden. Die Mikrofone können entweder auf einem Flächenvolumen²⁶ angeordnet oder arbiträr im Raum positioniert sein.

Für eine Mikrofonanordnung wie in Abbildung 3.1 dargestellt, treten beide Phänomene auf, sowohl Intensitätsunterschiede (ΔI) wie Laufzeitunterschiede (ΔT). Es entsteht also eine Unendliche Anzahl an möglichen Anordnungen. Die Aufgabe, eine ideale Anordnung zu finden ist dementsprechend komplizierter wie zuvor. Um Hilfe abzuschaffen werden Hörtests durchgeführt, die einen Vergleich zwischen den

²⁵ Schallquellen werden u.a. falsch abgebildet

²⁶ z.B. das Flächenvolumen einer Sphäre, in der englischen Literatur auch als "rigidsphere Microphonearrays" zu finden

verschiedensten Anordnungen ermöglichen. Eine qualitative Beurteilung dieser ist auf diese Weise deutlich leichter. Die Komplexität des psychoakustischen Modells für ein Wiedergabesystem mit N-Lautsprechern stellt ein weiteres Problem dar. Die Frage, nach welchen Kriterien soll eine Optimierung erfolgen bleibt noch offen.

3.2 Herangehensweise zu einer höheren räumlichen Auflösung

Die bis jetzt genannten Techniken arbeiten mit Richtcharakteristiken, die kommerziell verfügbar sind und versuchen eine optimale Mikrofonanordnung zu finden, mit der man brauchbare Surroundaufnahmen erstellen kann. Dieses Optimum nutzt aber nicht das ganze Potential eines 5.0 mehrkanaligen Systems.

Aber was ist unser Ziel überhaupt?

Ziel ist ein Verfahren zu finden, welche die Position von aufgenommenen Schallquellen bei der Wiedergabe²⁷ rekonstruieren kann. Folglich sollten bei der Wiedergabe Phantomschallquellen genau dort generiert werden, wo sich die realen Schallquellen befindet haben. Das heißt, eine Phantomschallquelle wird bei der Wiedergabe in der gleichen Richtung abgebildet, aus der die Mikrofonanordnung bestrahlt wurde. Für jede Raumrichtung ($0^\circ - 360^\circ$) können gezielte Amplituden für jeden Kanal festgelegt werden. Die Empfindlichkeit von jedem Kanal kann wie eine Art Panoramisierungsgesetz beschrieben werden. Abbildung 3.3 zeigt ein Beispiel für ein Intensitäts-Panoramisierungsgesetz 5.Ordnung.

Weil die Aufnahmeverfahren, die auf Intensitäts- und Laufzeitunterschiede basieren (siehe Kap.3.1), dieses Ziel nicht ganz für eine mehrkanalige Wiedergabe erreichen können, muss die Aufgabe mit einer anderen Herangehensweise gelöst werden.

²⁷ Die Wiedergabeanordnung ist für das Ziel von weniger Bedeutung. Eines ist klar, für eine Auflösung L.Ordnung (dreidimensionaler Fall) müssen mindestens $(L + 1)^2$ Lautsprecher eingesetzt werden

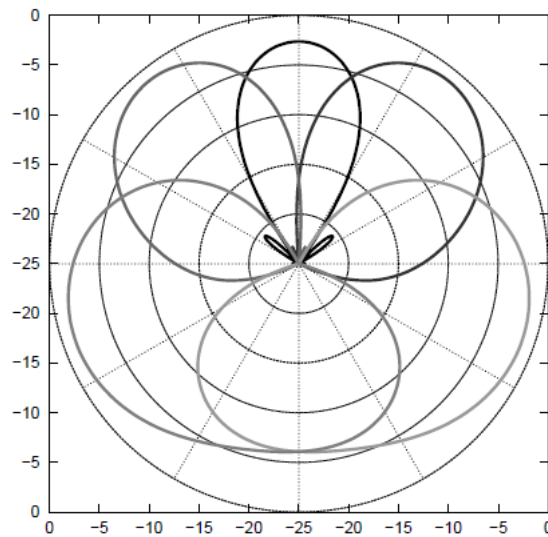


Abb.3.3: Intensitäts-Panoramisierungsgesetz 5.Ordnung

Mittels Direktivitäten wie im Panoramisierungsgesetz²⁸ in Abb.3.3 dargestellt, können Phantomschallquellen über (360°) sehr akkurat wiedergeben werden. In Abbildung 3.2 ist so ein System dargestellt. Es ist ersichtlich, dass die Anzahl der verwendeten Mikrofonen bei so einer Mikrofonanordnung keine wesentliche Rolle für die Wiedergabe spielt. Wenn eine Schallquelle das Array aus einer bestimmten Richtung bestrahlt, dann werden bei der Wiedergabe nur diejenigen Lautsprecher zum Einsatz kommen, welche dieser Richtung am besten entsprechen (vgl. Abb.3.1 und 3.2). Um das gewünschte räumliche Bild zu schaffen sollen trotzdem genug Mikrofone²⁹ zur Verfügung stehen. Der Hörer nimmt nun eine Phantomschallquelle aus der gleichen Richtung wahr, wo sich Schallquelle befand. Wenn ein System solche Direktivitäten besitzt, dann wird von einer hohen räumlichen Auflösung gesprochen.

Unter einer hohen räumlichen Auflösung versteht man bei der Wiedergabe eine präzise räumliche Darstellung von Phantomschallquellen. Die räumliche Auflösung wird analog zur zeitlichen Auflösung durch räumliche Frequenzen beschrieben. Ein Mikrofon

²⁸ Gesetz, welche die Raumrichtungen feststellt, aus denen Schall vorzuziehend aufgenommen wird

²⁹ Die Anzahl der nötigen Mikrofonen ist von der gewünschten Qualität bzw. Ordnung der Direktivitäten abhängig

0.Ordnung entspricht einer Kugel Richtcharakteristik, ein Mikrofon 1.Ordnung einer Achter Richtcharakteristik usw. Je höher die Ordnung eines Mikrofons, desto mehr werden bestimmte Raumrichtungen bei der Aufnahme bevorzugt (siehe Abb.3.4).

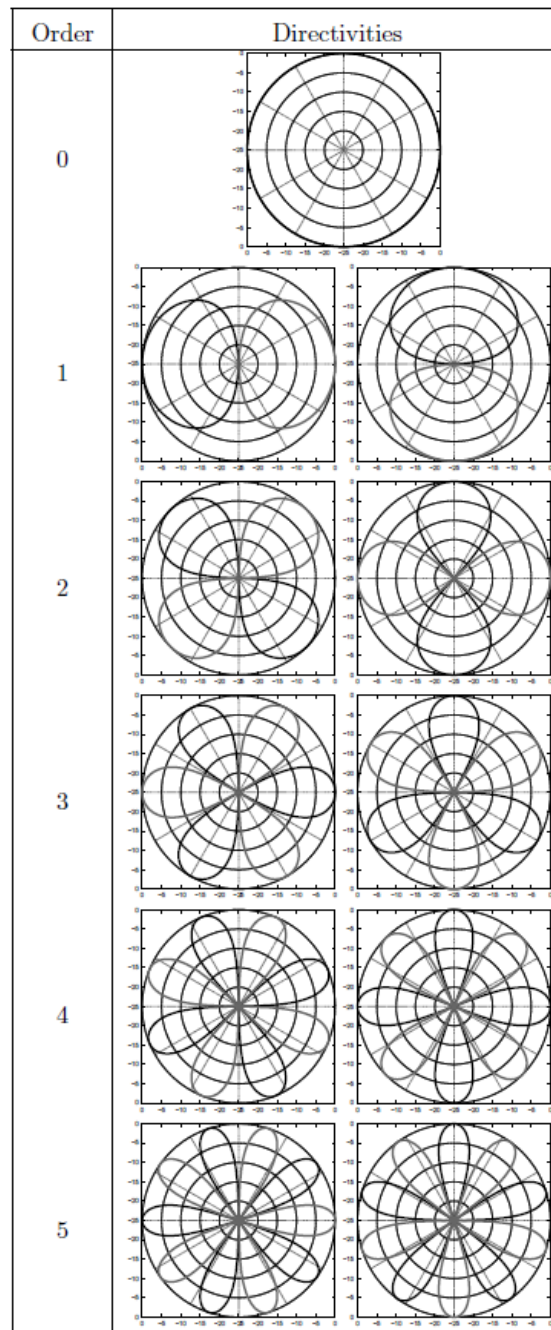


Abb.3.4: Räumliche Frequenzen bis zur 5.Ordnung (auch zylindrische harmonische genannt → 2D Fall)

Bis jetzt wurden nur die azimutale³⁰ Direktivitäten berücksichtigt. Da jedes Schallereignis dreidimensional ist, muss das vollständige dreidimensionale Modell mit einbezogen werden. Es handelt sich nicht mehr von den zweidimensionalen zylindrischen Harmonischen, sondern von den dreidimensionalen kugelförmigen Harmonischen.

Als nächstes sollen die Panoramisierungsgesetze auf ein dreidimensionales Modell erweitert werden. In Abbildung 3.5 ist eine mögliche Erweiterung eines solchen Panoramisierungsgesetz abgebildet.

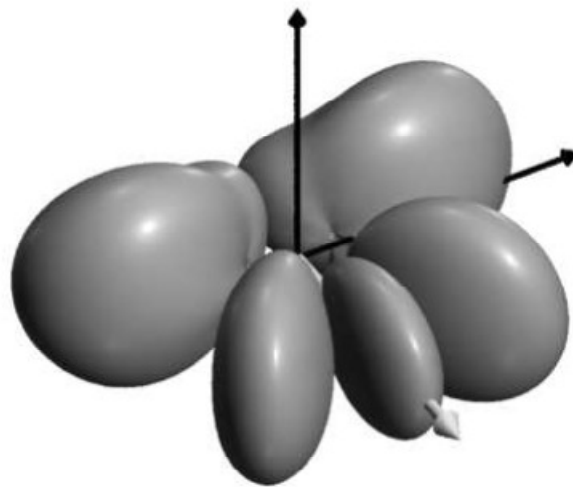


Abb.3.5: Mögliche dreidimensionale Erweiterung eines Panoramisierungsgesetz

3.3 Eigenschaften

- **Selektivität:** mit einer hohen Selektivität werden Schallquellen, die aus einer bestimmen Richtung die Mikrofonanordnung bestrahlen, nur durch die Lautsprecher, die diese Richtung am meisten entsprechen, wiedergegeben. Lautsprecher werden Raumrichtungen zugeordnet.

³⁰ Das Azimut ist ein Terminus, der einen Horizontalwinkel bezeichnet (siehe Abb.1.3)

- Asymmetrie: bis auf dem frontalen Kanal sind alle Direktivitäten asymmetrisch. Das heißt, die Achse einer Direktivität entspricht nicht der Richtung eines assoziierten Lautsprechers³¹. Somit ergibt sich eine gute Anpassung zwischen das Wiedergabesystem und die Mikrofonanordnung.
- Ergänzungsprinzip: die Summe aller Direktivitäten ergibt eine omnidirektionale Charakteristik. Die Leistung einer sich bewegenden Schallquelle hängt nicht von seiner Richtung ab. Um die Klangfarbe des Signals nicht zu verändern, muss diese Bedingung für alle Frequenzen eingehalten werden.

Das Problem ist dabei die Umsetzung der gewünschten Direktivitäten mit kommerziell verfügbaren Mikrofonen. Mit Hilfe von Singalverarbeitung und neuen Methoden ist es möglich diese Direktivitäten mit Mikrofonen 0. und 1.Ordnung zu erzeugen. Dieses Prinzip ist in Abbildung 3.6 dargestellt. Links ist eine Direktivität 1.Ordnung und rechts eine Direktivität 5.Ordnung abgebildet. Letztere lässt sich aus Direktivitäten 0. und 1 Ordnung zusammensetzen.

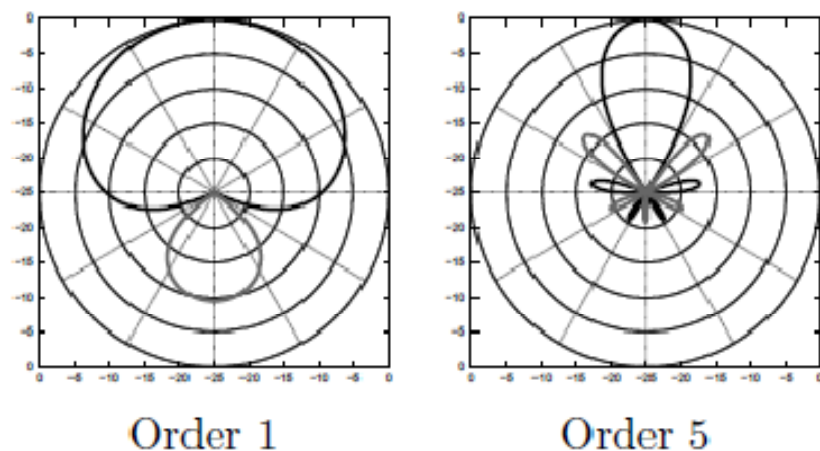


Abb.3.6: Direktivitäten 1. und 5.Ordnung

³¹ Es handelt sich nicht von einem System wie in Abb.3.1 abgebildet, die Lautsprecher stimmen nicht mit den Achsen der Diektivitäten überein

3.4 Implementierung

Durch Untersuchungen ist es bekannt, dass um eine hohe räumliche Auflösung zu erreichen, die Aufnahmepunkten verteilt werden müssen. Das Schallfeld wird nicht nur an einem Punkt im Raum oder nur aus einer Richtung abgetastet. Es können somit räumliche Frequenzen höherer Ordnung (vgl. Abb.3.4) unter Anwendung von Mikrofonen 0. und 1.Ordnung erstellt werden.

Diese Herangehensweise wurde mit zirkularen, sphärischen und irregulären dreidimensionalen Mikrofonanordnungen untersucht. Wie es bereits in Kap. 2.3 gezeigt wurde, ist die räumliche Auflösung bei irregulären Anordnungen über größere Abschnitte gleichmäßiger wie bei regulären (symmetrischen) Anordnungen.

In Kap.2.3 wurde eine irreguläre Mikrofonanordnung 3.Ordnung angesprochen, die aus 24 Mikrofonen mit einer Kugel Richtcharakteristik besteht. Aus den Mikrofonsignalen wurden 16 Fourier-Bessel Koeffizienten berechnet, dementsprechend die sphärische Harmonische 3.Ordnung ($(L + 1)^2 = 16$). Sollen sphärische Harmonische 5.Ordnung erzeugt werden, dann würde man mindestens 36 (vorzugsweise 48) Mikrofonen verwenden müssen. Dies ist schwer realisierbar und aus diesem Grund soll das Verfahren für eine bekannte Wiedergabeordnung optimiert werden. Es ist sinnvoll die Wiedergabeordnung festzulegen und auf diese Weise die Anzahl der nötigen Mikrofonen zu minimieren.

Für eine 5.0 Wiedergabeordnung genügen 8 Mikrofone um Direktivitäten 5.Ordnung zu erzeugen. Bei dieser Wiedergabeordnung ist man auf den zweidimensionalen Fall begrenzt (die Lautsprecher befinden sich alle auf einer Ebene, siehe Abb.3.2). Für eine Ordnung $L = 5$ sind mindestens $(L + 1)$ Signale erforderlich. Um die Lautsprechersignale zu erhalten werden die Mikrofonsignale durch die Filtermatrix geschickt (siehe Abb.3.7). Um die beste Ergebnisse zu bekommen muss für jede Wiedergabeordnung eine optimale Filtermatrix gefunden werden. Sobald die

Wiedergabeordnung festgelegt wurde, kann die Filtermatrix berechnet werden. Der Vorteil besteht darin, dass die Matrix nicht jedes mal berechnet werden muss.

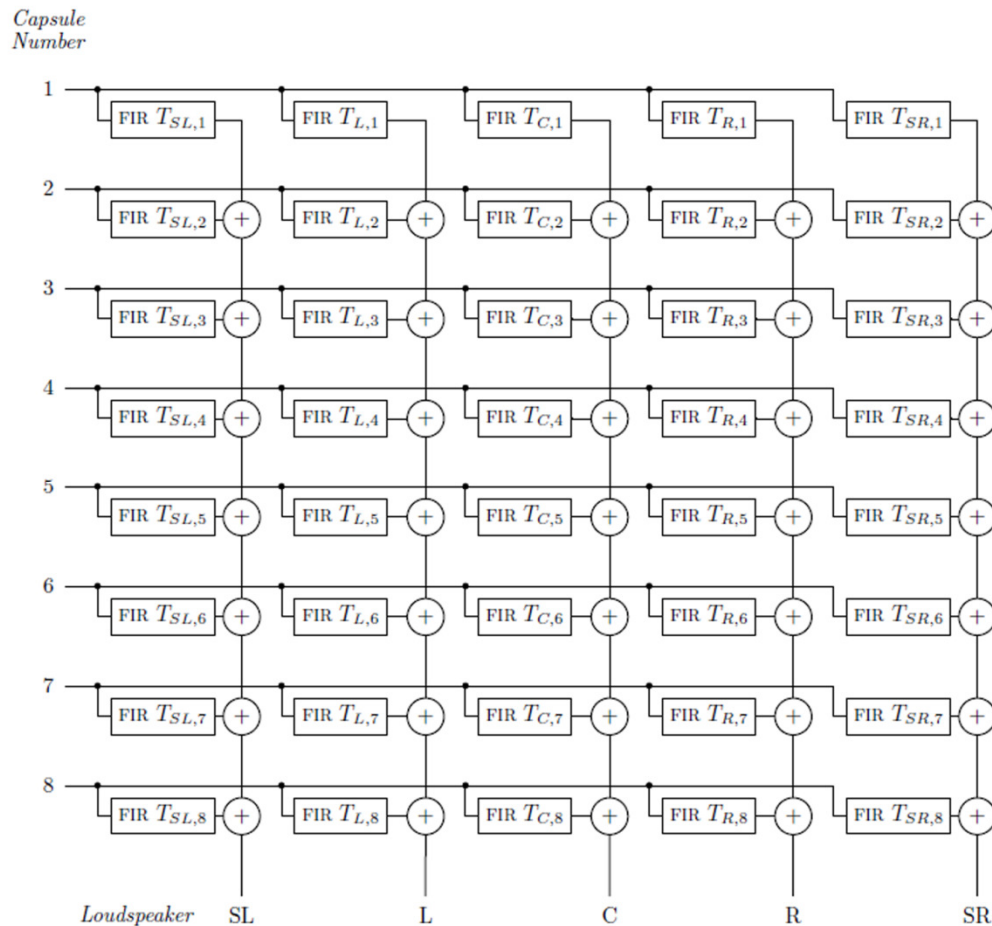


Abb.3.7: Matrix Filterung

3.5 Optimierung

Die Optimierung besteht darin, Filter zu entwerfen, die unmittelbar die gewünschte 5.0 Direktivitäten liefern. Somit erspart man sich die Berechnung dieser Direktivitäten aus den Ambisonic Signalen 5.Ordnung. Mit Hilfe der Fourier-Bessel Erweiterung (siehe

Gl.3) kann ein dreidimensionales Schallfeld aus einer endlichen Summe von Signalen dargestellt werden.

In diesem Fall ist unser Ziel nicht mehr das akustische Feld (p) zu berechnen (siehe Gl.6), sondern die Eingänge der Lautsprecher (v), welche eine Approximation ($\tilde{p}$) des akustischen Schallfeldes (p) darstellen. Wir erhalten die lineare Matrixbeziehung zwischen (c) und (v).

$$v = T * c \quad (\text{Gl.11})$$

Als nächstes muss die Transcoding Matrix (T) definiert werden. (T) ist aus den gewünschten Eigenschaften unter Anwendung eines akustischen Modells zu berechnen. Ihre Koeffizienten sind die Frequenzantworten der Filter. Jede gewünschte Eigenschaft wird allerdings mit einem elementaren Fehler assoziiert, welche noch minimiert werden muss. Der globale Fehler $e^2(T)$ ist als eine gewichtete Summe aller elementaren Fehlern definiert.

Der elementar Fehler $e_p^2(T)$ beschreibt wie genau die Direktivitäten mit der Transcoding Matrix (T) berechnet werden. Darüberhinaus sollen Direktivitäten in allen drei Raumrichtungen (dreidimensional) betrachtet werden. Die Gewichtungsmatrix (W) ist folglich im elementaren Fehler enthalten. Mit dieser Matrix können die drei Raumrichtungen gewichtet werden³².

Ein weiterer Fehler wird durch Rauschen erzeugt. Alle Mikrofone besitzen ein Eigenrauschen und dadurch ergeben sich Modellierfehler³³. Die Verstärkung von Rauschen $e_n^2(T)$ muss im globalen Fehler mit einbezogen werden, sodass Rauschen nicht durch den transcoding Prozess verstärkt wird.

³² Sollen alle drei Raumrichtungen gleich gewichtet werden, dann entspricht die Matrix (W) einer Identitätsmatrix

³³ z.B. Omnidirektionale Mikrofone messen nicht den Schalldruck in einem Punkt im Raum, sondern den durchschnittlichen Betrag des Schalldruckes an der Mikrofonmembran

Der globale Fehler ergibt sich zu

$$e^2(T) = e_p^2(T) + \mu e_n^2(T) \quad (\text{Gl.12})$$

mit

$$e_p^2(T) = \text{trace} (TB - U_o)W(TB - U_o)^T \quad (\text{Gl.13})$$

und

$$e_n^2(T) = \text{trace} TT^T \quad (\text{Gl.14})$$

Wobei $\text{trace}(X)$ in der linearen Algebra die Summe der Diagonalelemente der quadratischen Matrix (X) ist. Die Matrix (U_o) enthält die dreidimensionalen Direktivitäten des mehrkanaligen Mikrofons. Diese Direktivitäten sind meistens durch das gewählte Panoramisierungsgesetz bestimmt (vgl. Abb.3.3)

Die gesuchte Transcoding Matrix (T) ist diejenige Matrix, die den globalen Fehler $e^2(T)$ am meisten minimiert.

$$T = U_oWB^T(BWB^T + \mu I)^{-1} \quad (\text{Gl.15})$$

Die Transcoding Matrix (T) enthält ein optimales Verhältnis zwischen Laufzeit- und Intensitätsdifferenzen. In den Mikrofonsignalen können nur Laufzeitunterschiede entstehen³⁴. Bei den Ausgangsignalen des mehrkanaligen Systems können allerdings nur Intensitätsunterschiede erzeugt werden. Aus der Sicht der Mikrofone handelt es sich um ein nicht koinzidentes Verfahren (die Mikrofone sind im Raum verteilt). Trotzdem ist dieses ein koinzidentes Verfahren, weil der Schalldruck auf einer Kugeloberfläche betrachtet wird. Es können also keine Laufzeitunterschiede auftreten.

Mit der Definition von (c) (Gl.5) und der Kenntnis über die Koeffizienten von (T) ist es nun möglich die Lautsprecher signale (v) des mehrkanaligen Systems zu berechnen

³⁴ Alle Mikrofone besitzen eine Kugel Richtcharakteristik

$$v = T * B * p \quad (\text{Gl.16})$$

Wenn (p) eine ebene Welle in einer bestimmten Richtung beschreibt, dann gibt (v) die Signale an, die den Lautsprechern zugeführt werden müsste, um die gewünschte Wiedergabe zu erlangen (vgl. Kap.3.2). Das Prinzip ist in Abbildung 3.8 genau dargestellt (vgl. Gl.16).

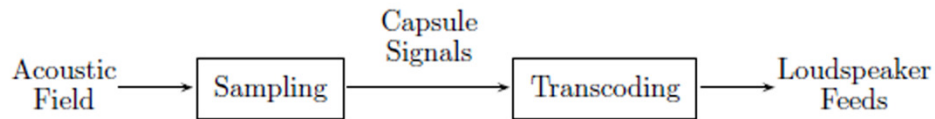


Abb.3.8: Transcoding Prinzip

Um die beste Mikrofonanordnung zu finden wurden viele Simulationen von *Trinnov* durchgeführt. Diese Simulationen bestanden aus der Positionierung der einzelnen Mikrofonen, Berechnung der Matrix (T) (siehe Gl.15) und die Auswertung der Ergebnisse. Die Anordnung mit den besten Ergebnissen und deren erzielten Direktivitäten sind in Abbildung 3.9 bzw. 3.11 abgebildet. Diese Mikrofonanordnung ist auf dem zweidimensionalen Fall begrenzt weil sie für eine 5.0 Wiedergabeanordnung optimiert wurde. Für den dreidimensionalen Fall würde man deutlich mehrere Lautsprecher benötigen.

Im tieffrequenten Bereich (ab ca. 5 kHz, siehe Abb. 3.11) konnten Direktivitäten 5.Ordnung nicht erreicht werden. Die Wellenlänge (λ) ist zu groß im Vergleich zur Mikrofonanordnung (ungefähr 20 cm mal 20 cm). Für hohe Frequenzen treten ausgeprägte Nebenkeulen auf (vgl. Abb.3.11 bei einer $f = 5 \text{ kHz}$). Dies kann wiederum durch die Wellenlänge (λ) erklärt werden, die bereits kleiner ist, als der kleinste Abstand zwischen den Mikrofonen. Eine eindeutige Ortung der Schallquelle ist nicht mehr möglich.

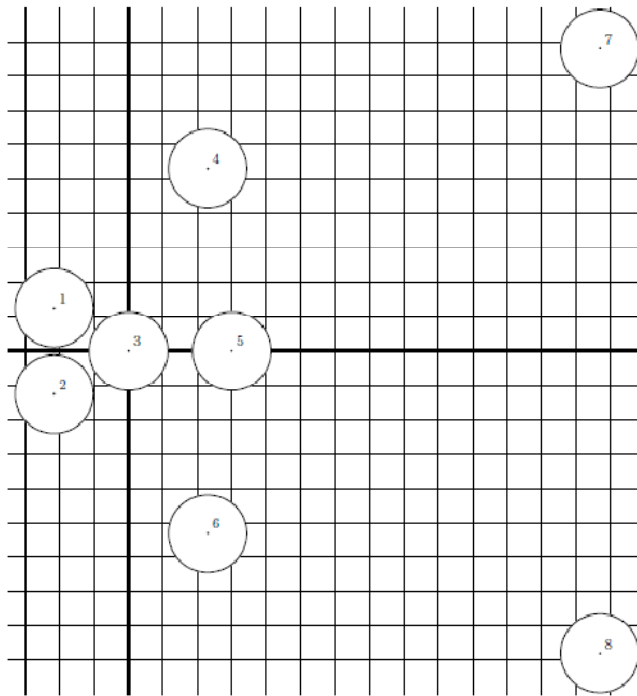


Abb.3.9: HSR- Mikrofonanordnung (High Spatial Resolution)

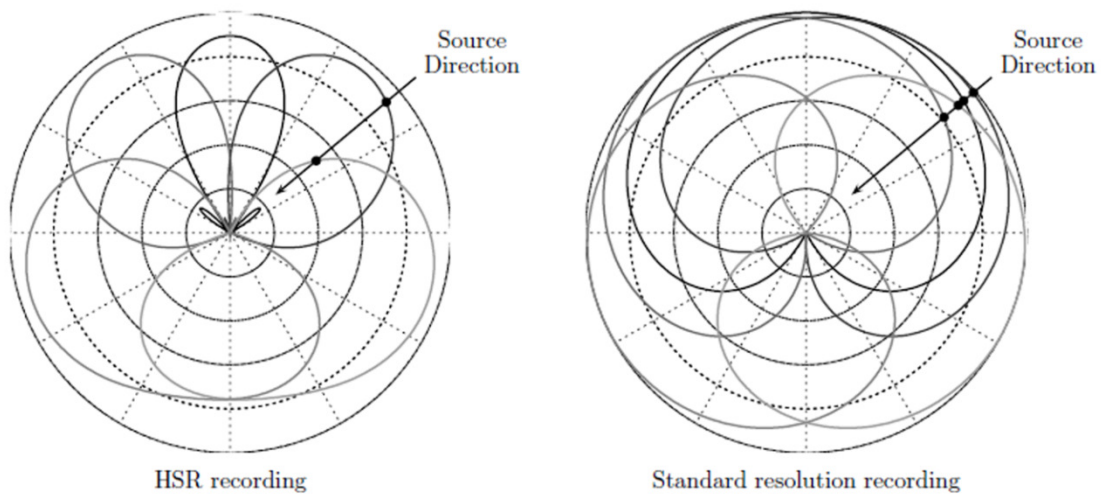


Abb.3.10: Aufnahme einer Schallquelle unter Anwendung von HSR und normaler Auflösung

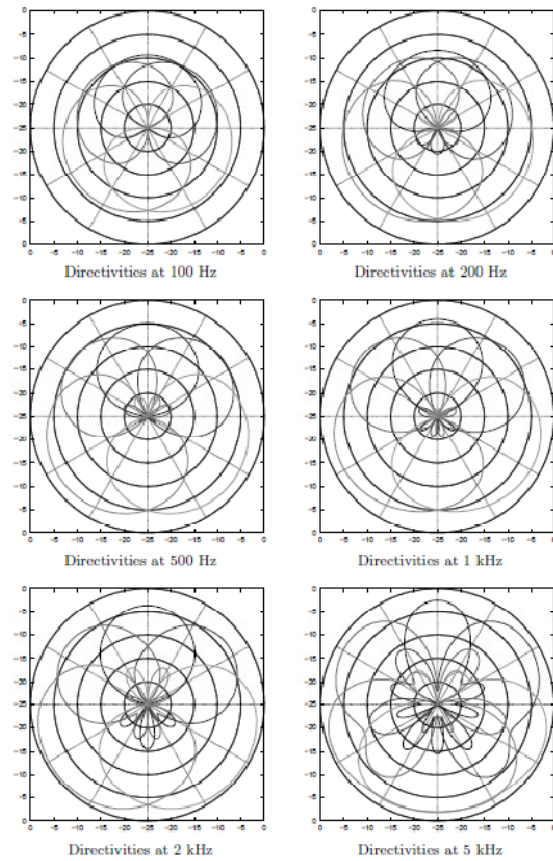


Abb.3.11: Erzielte Direktivitäten der Mikrofonanordnung in Abb.3.9

4. Zusammenfassung

Die Anwendung eines Verfahrens, welche ein hohe räumliche Auflösung erzeugen kann, hat wegen der erweiterten Selektivität (siehe Abb. 3.10) viele Folgen für die Aufnahmetechnik. Es handelt sich von einer kompletten neuen Herangehensweise und erlaubt eine Vielfalt an Gestaltungsmöglichkeiten der Räumlichkeit und des Aufnahmewinkels unter Verwendung von möglichst wenigen Mikrofonen.

Vorteile dieses Verfahrens sind unter anderen, die Trennung der Kanälen. Dies kann mit der Kreuzkorrelation³⁵ zwischen den einzelnen Kanälen beschrieben werden. Wenn die Kreuzkorrelation zu groß ist, dann sind die Signale zu ähnlich zu einander und die räumlichen Kapazitäten nehmen ab, vor allem was die Größe des Sweet-Spots betrifft. Ist sie im Gegenteil zu klein, dann haben die Signale keinen statistischen Zusammenhang mehr und die Abbildung von Phantomschallquellen verschlechtert sich erheblich. Diese Technik ermöglicht es die gewünschte Abtrennung zwischen den Lautsprechern mit Hilfe von Panoramisierungsgesetzen zu definieren. Dadurch können Phantomschallquellen bestmöglich abgebildet werden (im zweidimensionalen Fall werden nicht mehr wie 2 Lautsprecher zur Darstellung von Phantomschallquellen eingesetzt, vgl. Abb.3.2). Außerdem ist dieses Verfahren sehr robust bei der Behandlung durch einen Equalizer oder Kodierer. Veränderungen die nur bei einem bestimmten Kanal vorgenommen wurden, beeinflussen auch nur die Phantomschallquellen, die durch diesen Kanal abgebildet werden.

Ein weiterer Vorteil ist die Stabilität der Klangfarbe. Jedes Filter der Matrix in Abbildung 3.7 ist für alle Frequenzen optimiert worden, die Klangfarbe wird also nicht beeinflusst. Die Summe aller Direktivitäten ergibt eine definierte Direktivität, meistens eine Kugel Richtcharakteristik (siehe Abb.3.3). Somit bleibt der Raumanteil konstant und verändert sich nicht durch die gewählten Direktivitäten.

Falls eine Schallquelle sich um die Mikrofonanordnung bewegen sollte, dann bewegt sich die Phantomschallquelle bei der Wiedergabe genau so. Sie verschwindet nicht vom

³⁵ Die Korrelation beschreibt die lineare Beziehung zwischen zwei oder mehreren statistischen Variablen. Die Kreuzkorrelationfunktion beschreibt die Korrelation zwischen zwei verschiedenen Signalen

Bild oder wird genau dort abgebildet, wo sich der Lautsprecher befindet. Die Amplitude und das Spektrum des Signals werden nicht dadurch verändert.

Die Direktivitäten der Mikrofonanordnung werden von der Zielsetzung abhängig entworfen. Durch Verändern der Größe und Richtung der Keulen der Direktivitäten kann den Aufnahmewinkel leicht auf die Zielsetzung optimiert werden.

Fazit

Eine hohe räumliche Auflösung ermöglicht es, Surround Aufnahmen aus einer anderen Perspektive zu betrachten. Da die Kanäle besser von einander abgetrennt sind und die Direktivitäten besser (auch dreidimensional) festgelegt werden können, ist es nun möglich Schall im Raum akkurater wiederzugeben. Die Direktivitäten sind auch steuerbar, das heißt sie können in unterschiedlichen Raumrichtungen gewichtet werden und bieten somit eine große Flexibilität.

Eine hohe räumliche Auflösung (Direktivitäten 5.Ordnung) ist unter Anwendung von acht Mikrofonen mit einer Kugel Richtcharakteristik erreichbar. Die Signale aus der Mikrofonanordnung werden dem Computer zur Berechnung der HSR mehrkanaligen Signalen übergeben. Es können dann andere Mikrofone, Richtcharakteristiken und Positionierungen ausprobiert werden ohne die Filtermatrix wieder berechnen zu müssen. Dies stellt den riesen Vorteil dieser Methode dar.

Das Verfahren kann auch dafür verwendet werden, jedes Set von Direktivitäten unter Anwendung von variablen Mikrofonanordnungen zu entwerfen. Dabei kann mit beliebigen Richtcharakteristiken experimentiert werden.

Ob das Ergebnis brauchbar ist, hängt wiederum von der Zielsetzung ab und vor allem ob der Toningenieur dieses Werkzeug richtig anwenden kann.

Quellenverzeichnis

[1] Thomas Görne: „Tontechnik“, Fachbuchverlag Leipzig, 2006

[2] Ambisoncs Website. URL: <http://www.ambisonic.net>

[3] Alois Sontacchi, Robert Höldrich: “Schallfeldreproduktion durch ein verbessertes Holophonie – Ambisonic System”, Institut für Elektronische Musik und Akustik, Universität für Musik und Darstellende Kunst Graz

[4] Immersive Sound Project: “A New Comprehensive Approach of Surround Sound Recording”, 114th AES Convention in Amsterdam, Preprint 5717, März 2003

[5] Trinnov Audio: “High Spatial Resolution Multichannel Recording”, 116th AES Convention in Berlin, Preprint 6116, May 2004

[6] Michael A. Gerzon: “Periphony: With-height sound reproduction”, J. Audio Eng. Soc., 21:2-10, Januar 1973

[7] Angelo Farina and Emanuele Ugolotti: “Software implementation of B-format encoding and decoding” J. Audio Eng. Soc., 34:716-744, September 1986

[8] Michael A. Gerzon: “The optimum choice of surround sound specification”, AES Preprint 1199, März 1977

[9] Peter Craven, Michael A. Gerzon: “Coincident microphone simulation covering three dimensional space and yielding various directional outputs”, US Patent 4,042,779. Aug. 16, 1977

[10] Michael A. Gerzon: "Sound reproduction system with non-square lay-out", US Patent 4,086,433. Apr. 25, 1978

[11] Trinnov Website. URL: <http://www.trinnovaudio.com>

[12] Michael Dickreiter: "Handbuch der Tonstudioteknik", 7.völlig neu bearbeitete und erweiterte Auflage, Nürnberg, 2 Bände, Verlag: K G Saur, 2008

[13] Alan V. Oppenheim, Ronald W. Schafer, John R. Buck: "Discrete-Time Signal Processing"

[14] Gerhard Graber, Werner Weselak: "Raumakustik Skryptum", Version 5.0, 2007

[15] Earl G. Williams: "Fourier Acoustics, Sound Radiation and Nearfield Acoustical Holography", 1999.

[16] Hannes Pomberger: Diplomarbeit "Angular and Radial Directivity Control for Spherical Loudspeaker Arrays", Institut für Elektronische Musik und Akustik, Universität für Musik und Darstellende Kunst Graz, 2008

Anhang

Sphärische Wellen

Die Beschreibung von Wellenausbreitung in zylindrischen und ebenflächigen Geometrien besteht aus Wellenfronten, die sich für mindestens eine Dimension ins Unendliche ausdehnen. Die sphärische Geometrie liefert eine endliche und kompakte Entfaltung der Wellenfronten und ermöglicht einen Verstand von sich ausbreitenden Wellen.

Das Kugel Koordinatensystem (siehe Abb.1.3) wird so in Verbindung mit dem Kartesischen Koordinatensystem gesetzt

$$x = r \sin\theta \cos\phi \quad [1]$$

$$y = r \sin\theta \sin\phi \quad [2]$$

$$z = r \cos\theta \quad [3]$$

dabei gilt

$$\begin{aligned} r &= \frac{x}{\sin\theta \cos\phi} & \frac{x}{\sin\theta \cos\phi} &= \frac{y}{\sin\theta \sin\phi} \\ r &= \frac{y}{\sin\theta \sin\phi} & y &= x \frac{\sin\phi}{\cos\phi} = x \tan\phi & \phi &= \tan^{-1}\left(\frac{y}{x}\right) \end{aligned} \quad [4]$$

$$\begin{aligned} r^2 &= x^2 + y^2 + z^2 = r^2 \sin^2\theta \cos^2\phi + r^2 \sin^2\theta \sin^2\phi + r^2 \cos^2\theta = \\ &= r^2 \cos^2\theta \left(\frac{\sin^2\theta}{\cos^2\theta} \cos^2\phi + \frac{\sin^2\theta}{\cos^2\theta} \sin^2\phi + 1 \right) = r^2 \cos^2\theta \tan^2\theta \left(\cos^2\phi + \sin^2\phi + \frac{1}{\tan^2\theta} \right) \\ &= r^2 \cos^2\theta (\tan^2\theta + 1) = r^2 \left(\frac{\cos^2\theta \sin^2\theta}{\cos^2\theta} + \cos^2\theta \right) = r^2 (\sin^2\theta + \cos^2\theta) = r^2 \end{aligned}$$

$$r = \sqrt{x^2 + y^2 + z^2} \quad [5]$$

Ausgehend von Gleichung [3]

$$r = \frac{z}{\cos\theta}$$

$$r^2 \sin^2\theta = \frac{z^2}{\cos^2\theta} \sin^2\theta$$

$$r^2 \sin^2\theta (\cos^2\phi + \sin^2\phi) = z^2 \tan^2\theta$$

$$r^2 \sin^2\theta \cos^2\phi + r^2 \sin^2\theta \sin^2\phi = z^2 \tan^2\theta$$

und mit den Gleichungen [1], [2] und [3] erhält man letztendlich

$$x^2 + y^2 = z^2 \tan^2\theta$$

$$\sqrt{x^2 + y^2} = z \tan\theta$$

$$\theta = \tan^{-1}\left(\frac{\sqrt{x^2 + y^2}}{z}\right) \quad [6]$$

Die zeitabhängige Wellengleichung in Kugel Koordinaten ist so gegeben

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial p}{\partial r} \right) + \frac{1}{r^2 \sin\theta} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial p}{\partial \theta} \right) + \frac{1}{r^2 \sin^2\theta} \frac{\partial^2 p}{\partial \phi^2} - \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = 0$$

[7]

dabei ist

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial p}{\partial r} \right) = \frac{1}{r} \frac{\partial^2}{\partial r^2} (rp)$$

Der Gradient des Schalldruckes in Kugel Koordinaten ist so definiert,

$$\nabla p = \hat{e}_r \frac{\partial p}{\partial r} + \hat{e}_\theta \frac{1}{r} \frac{\partial p}{\partial \theta} + \hat{e}_\phi \frac{1}{r \sin \theta} \frac{\partial p}{\partial \phi} \quad [8]$$

wobei $\hat{e}_r, \hat{e}_\theta, \hat{e}_\phi$ die Einheitsvektoren der einzelnen Koordinatenrichtungen darstellen.

Die Euler'sche Gleichung³⁶

$$-\nabla p = \rho_0 \frac{\partial v}{\partial t} \quad [9]$$

und dessen Fourier-Transformiert

$$\nabla p(r, \theta, \phi) = \rho_0 i k c v(r, \theta, \phi) \quad [10]$$

liefern den Schalldruck in Abhängigkeit der Schallschnelle. Dabei sind die Kreisfrequenz und der Schnellevektor so definiert

$$w = kc$$

$$v \equiv \dot{u} \hat{e}_\theta + \dot{v} \hat{e}_\phi + \dot{w} \hat{e}_r$$

Die Lösungen zur Gleichung [7] ergeben sich durch Trennung der Variablen zu

$$p(r, \theta, \phi, t) = R(r) \Theta(\theta) \Phi(\phi) T(t) \quad [11]$$

Auf diese Weise kann nun jedes Ergebnis der Wellengleichung (siehe [7]), implizit $e^{-i\omega t}$ aus Gleichung [12] berechnet werden.

³⁶ Für die genaue Herleitung siehe Quellenverzeichnis [14] Kap.2.2.2

$$p(r, \theta, \phi, w) = \sum_{l=0}^{\infty} \sum_{m=-l}^l (A_{l,m}(w) j_l(kr) + B_{l,m}(w) y_l(kr)) Y_l^m(\theta, \phi) \quad [12]$$

Bei dieser Lösung werden stehende Wellen betrachtet. Da dieser Fall von Interesse ist (vgl. Kap 1.3, Gl.3), werden sich ausbreitende Wellen nicht weiters untersucht. Der zu messenden Schalldruck setzt sich aus allen durch die Schallquellen abgestrahlten Wellenfronten und Reflexionen zusammen. Der andere Fall eignet sich gut zur Untersuchung von einer hinlaufenden bzw. rücklaufenden Welle.

Wie in Gleichung [12] ersichtlich ist, wurden die Winkelfunktionen zu einer Funktion zusammengefasst, die sogenannten Kugelflächenfunktionen.

$$Y_l^m(\theta, \Phi) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi (l+m)!}} P_l^m(\cos \theta) e^{im\phi} \quad [13]$$

Die Winkelfunktionen

Die Legendre Polynome sind die Lösungen zur Gleichung [14] wenn $m = 0$ ist. Diese Lösungen beschreiben Felder, bei denen keine Änderung des azimutalen Winkels feststellbar ist. Sie sind aus diesem Grund von großer Bedeutung.

$$\frac{\partial}{\partial \eta} \left[(1 - \eta^2) \frac{\partial \theta}{\partial \eta} \right] + \left[l(l+1) - \frac{m^2}{1 - \eta^2} \right] \theta = 0 \quad [14]$$

mit $\eta = \cos \theta$ folgt $-1 \leq \eta \leq 1$

Für den Fall dass $m = 0$ ergibt sich

$$P_l^{m=0}(x) = P_l(x) \quad [15]$$

und mit Rodrigues Formel erhält man letztendlich

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l \quad [16]$$

Mit $x = \cos \theta$ ist die Periodizität der Legendre Polynome ersichtlich.

Weiters werden die assoziierten Legendre Funktionen betrachtet. Diese Funktionen sind durch zwei ganzzahlige Indizes definiert (also $m \neq 0$).

für $m > 0$ sind diese wie folgt definiert

$$P_l^m(x) = (-1)^m (1 - x^2)^{m/2} \frac{d^m}{dx^m} P_l(x) \quad [17]$$

Der Normalisierungsfaktor $(-1)^m$ wird in der Literatur nicht immer angegeben. In den von Trinnov³⁷ präsentierten Papers ist dieser Faktor auch nicht zu finden. Für den weiteren Fall von $m < 0$ ergibt sich eine andere Definition von $P_l^m(x)$, die hier nicht mehr von großer Interesse ist³⁸.

Das sphärische Wellenspektrum $P_{l,m}(r)$ an der Stelle $r = r_0$ ist so definiert (siehe auch Quellenverzeichnis [15] Kap.6.7.1)

$$P_{l,m}(r_0, w) \equiv \int p(r_0, \theta, \phi, w) Y_l^m(\theta, \phi)^* d\Omega \quad [18]$$

In diesem spezifischen Fall wird das Problem nur für den Innenbereich eines Volumens betrachtet. Für diese Fälle befinden sich alle Schallquellen außerhalb des Radius $r = b$. Das sphärische Volumen ist also Quellenfrei und die Lösungen sind nur im Bereich $r < b$ gültig. Die Lösungen bestehen aus radialen Funktionen, die am Ursprung endlich sind, so dass sich der innere Schalldruck aus Gleichung [12] zur

³⁷ Siehe Quellenverzeichnis [4] und [5]

³⁸ Siehe Quellenverzeichnis [15] Kap.6.3.2

$$p(r, \theta, \phi, w) = \sum_{l=0}^{\infty} \sum_{m=-l}^l A_{l,m}(w) j_l(kr) Y_l^m(\theta, \phi) \quad [19]$$

aufschreiben lässt.

Gleichung [18] ist wie eine Fourier Transformation mit der Basisfunktion $Y_l^m(\theta, \phi)$ zu verstehen. Gleichung [19] ist die dazugehörige inverse Fourier Transformation, eine zweifache Fourier Serie.

Streuung von Sphären

Für unsere Betrachtungen muss den gesamten Schalldruck in zwei aufgeteilt werden. Der einen Anteil beschreibt die einfallende Welle und der anderen die gestreute Welle. Streuung entsteht wenn die einfallende Welle auf die Sphäre trifft.

$$p = p_s + p_i \quad [20]$$

(p_i) beschreibt den Schalldruck, der man messen würde, wenn die Sphäre gar nicht vorhanden wäre. (p_s) beschreibt also den Schalldruck, der durch das Einbringen der Sphäre ins Schallfeld entsteht. Ist $(p_s) = 0$, dann hat die Sphäre keinen Einfluss auf die einfallende Welle bzw. auf den gemessenen Schalldruck.

Weiters wird angenommen, dass die einfallende Welle eine ebene Welle ist. Das heißt, die Schallquelle befindet sich unendlich weit weg von der Sphäre. Und letztens wird angenommen, dass $(p_s) \ll 1$ wäre.

Mit Gleichung [3] und diesen letzten Annahmen kann der Schalldruck (Fall für die Ortung einer Schallquelle im Unendlichen auf der z-Achse, also $z = -\infty$) so angeschrieben werden

$$p(r, \theta, \phi, w) = P_0 e^{ikz} = P_0 e^{ikr \cos \theta} \quad [21]$$

wobei (k) die Wellenzahl ist und so definiert wird $(k = \frac{w}{c})$.

Analog zur Gleichung [19] erhält man den Schalldruck in dieser Form

$$p_i(r, \theta, \phi = 0, w) = P_0 \sum_{l=0}^{\infty} \sum_{m=-l}^l A_{l,m}(w) j_l(kr) Y_l^m(\theta, \phi) \quad [22]$$

Der allgemeiner Fall soll hier näher betrachtet werden, das heißt, man weiß nicht aus welcher Raumrichtung die einfallende Welle auf die Sphäre trifft.

$$p_i(r, t) = P_0 e^{i(k_i r - wt)} \quad [23]$$

$$e^{ik_i r} = 4\pi \sum_{l=0}^{\infty} i^l j_l(kr) \sum_{m=-l}^l Y_l^m(\theta, \phi) Y_l^m(\theta_i, \phi_i)^* \quad [24]$$

Mit diesem Ergebnis und mit der Definition des sphärischen Wellenspektrums (Gleichung [18]) kann auf das kompakte Ergebnis in Gl.3 (Kap.1.3) zurückgeführt werden. Es verbleiben nun die sphärischen Bessel Funktionen, die als nächstes betrachtet werden.

Sphärische Bessel Funktionen

Bessel'sche Gleichung

$$\left(\frac{d^2}{dr^2} + \frac{1}{r} \frac{d}{dr} + k^2 - \frac{\left(l + \frac{1}{2}\right)^2}{r^2} \right) u_l(r) = 0 \quad [25]$$

$$u_l(r) = R_l(r) r^{\frac{1}{2}} \quad [26]$$

Die Lösungen der Gleichung [25] sind die Bessel Funktionen $J_{l+\frac{1}{2}}(kr)$ (1.Ordnung) und $Y_{l+\frac{1}{2}}(kr)$ (2.Ordnung).

Die Sphärische Bessel Funktionen können rekursiv berechnet werden

$$j_l(x) \equiv \left(\frac{\pi}{2x}\right)^{\frac{1}{2}} J_{l+\frac{1}{2}}(kr) \quad [27]$$

$$y_l(x) \equiv \left(\frac{\pi}{2x}\right)^{\frac{1}{2}} Y_{l+\frac{1}{2}}(kr)$$

$$h_l^{(1)} = j_l(x) + i y_l(x)$$

$$h_l^{(2)} = j_l(x) - i y_l(x)$$

sowie durch Ableiten der trigonometrischen Funktionen

$$j_l(x) = (-x)^l \left(\frac{1}{x} \frac{d}{dx}\right)^l \left(\frac{\sin x}{x}\right) \quad [28]$$

$$y_l(x) = -(-x)^l \left(\frac{1}{x} \frac{d}{dx}\right)^l \left(\frac{\cos x}{x}\right)$$

Für unsere Betrachtungen sind nur die Bessel Funktionen 1.Ordnung (Gleichung [27] bzw. [28]) von Interesse. Die Hankel Funktionen sind also rein reel.

In dieser Arbeit wurden folgende Notationen verwendet:

- (p) ist ein Vektor, der die Fourier Transformierte der richtungsunabhängigen Koeffizienten enthält und ist so angeordnet

$$p = \left(P_{0,0}(f) \ P_{1,-1}(f) \ P_{1,0}(f) \ P_{1,1}(f) \ P_{2,-2}(f) \ \dots \ P_{L,-L}(f) \ \dots \ P_{L,L}(f) \right)^t$$

- (c) ist ein Vektor, der die Fourier Transformierte der N-Mikrofonsignalen enthält und ist so angeordnet

$$c = \left(c_1(f) \ c_2(f) \ \dots \ c_N(f) \right)^t$$

- (v) ist ein Vektor, der die Fourier Transformierte der Lautsprechersignalen enthält und ist so angeordnet

$$v = \left(V_1(f) \ V_2(f) \ \dots \ V_N(f) \right)^t$$

- (B) ist die Sampling Matrix und ist so angeordnet

$$B = \begin{pmatrix} B_{1,0,0}(f) & \dots & B_{1,L,L}(f) \\ \vdots & \ddots & \vdots \\ B_{N,0,0}(f) & \dots & B_{N,L,L}(f) \end{pmatrix}$$

Werden Mikrofone mit einer Kugel Richtcharakteristik an den Kugel Koordinaten (r_n, θ_n, ϕ_n) positioniert, dann lassen sich die Elemente der Sampling Matrix (B) so berechnen

$$B_{n,l,m}(f) = 4\pi \ i^l \ j_l(kr_n) \ Y_l^m(\theta_n \phi_n)$$

- (T) ist die Transcoding Matrix, die die Parameter der Filter enthält und ist so angeordnet

$$T = \begin{pmatrix} T_{1,1}(f) & \dots & T_{1,N}(f) \\ \vdots & \ddots & \vdots \\ T_{N,1}(f) & \dots & T_{N,N}(f) \end{pmatrix}$$