

Wellenfeldsynthese im IEM CUBE

Seminararbeit aus:
Algorithmen in Computermusik und Akustik 2

Wolfgang Jäger
Sebastian Preis

Graz, 24. März 2009
Betreuer: Franz Zotter



institut für elektronische musik und akustik



Zusammenfassung

Die Wellenfeldsynthese ist der Versuch ein Wellenfeld, ausgehend von einer virtuellen Quelle, durch eine Lautsprecheranordnung zu simulieren. Diese Simulation basiert auf dem Huygens-Prinzip. Dieses besagt, dass jeder Punkt einer Wellenfront Ausgangspunkt einer Elementarwelle ist. Andererseits entsteht durch die Überlagerung der Elementarwellen eine geschlossene Wellenfront. Letzteres ist der Ausgangspunkt für die Wellenfeldsynthese. Auch jeder Lautsprecher ist Ausgangspunkt einer Kugelwelle durch welche in Summe die Wellenfront der virtuellen Quelle generiert wird. Natürlich kann das Huygens-Prinzip nur angenähert werden, da nicht an jedem Punkt einer Wellenfront ein Lautsprecher positioniert werden kann. Die theoretische Grundlage der WFS ist das Kirchhoff-Helmholtz-Integral. Dieses besagt, dass der Schalldruck, in jedem Punkt innerhalb eines Volumens, durch den Schalldruck und die Normalkomponente der Schallschnelle am Rand dieses Volumens beschrieben werden kann. Ausgehend von dieser Theorie wurde die WFS in PD umgesetzt und anschließend im CUBE des IEM getestet. Für eine objektive Bewertung wurde ein Hörversuch durchgeführt, der gut ausgefallen ist. In dieser Seminararbeit wird in weiterer Folge auf die Theorie, die Umsetzung, die Schwierigkeiten und Grenzen und auf den Hörversuch eingegangen werden. Als Basis diente unter anderem folgende Literatur: [Bla74], [Spo05], [SA08c].

Inhaltsverzeichnis

| | | |
|----------|--|-----------|
| 1 | Theoretische Grundlagen | 4 |
| 1.1 | Gaußscher Integralsatz | 4 |
| 1.2 | Kirchhoff-Helmholtz-Integral | 8 |
| 1.3 | Die Schallschnelle | 11 |
| 2 | Implementation | 13 |
| 2.1 | Filter | 13 |
| 2.1.1 | Repolarisierung des Filters | 14 |
| 2.2 | Fensterung | 17 |
| 2.3 | Amplitude | 17 |
| 2.4 | Delay | 19 |
| 2.5 | Flächengewichtung | 20 |
| 2.6 | Blockschaltbild | 23 |
| 3 | Mögliche Artefakte | 24 |
| 3.1 | Dopplereffekt | 24 |
| 3.2 | Truncation Error | 25 |
| 3.3 | Spatial Aliasing | 25 |
| 4 | Hörversuch im CUBE | 28 |
| 4.1 | Vorbereitungen und Überlegungen | 29 |
| 4.2 | Durchführung und Auswertung des Hörversuches | 30 |
| 4.3 | Ambisonics und Wellenfeldsynthese im Vergleich | 35 |

1 Theoretische Grundlagen

Im ersten Teil dieser Arbeit wird der Gaußsche Integralsatz und anschließend, basierend auf diesem Integralsatz, das Kirchhoff-Helmholtz-Integral hergeleitet. Dieses Kapitel bietet die Grundlagen für das Verständnis der praktischen Umsetzung.

1.1 Gaußscher Integralsatz

Für die Herleitung des Gaußschen Integralsatzes benötigt man zwei Ansätze. Gegeben sei ein allgemeines Vektorfeld, das noch nicht im direkten Zusammenhang mit den Größen des Schallfeldes stehen muss

$$\mathbf{F} = \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix}.$$

Dieses Vektorfeld durchdringt ein Volumen Ω_1 mit der Oberfläche Γ_1 . (siehe Abbildung 1)

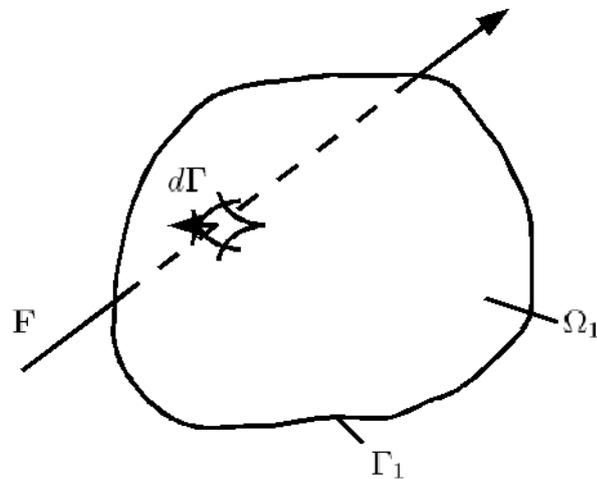


Abbildung 1: Vektorfeld durchdringt Volumen

a.) Durch Bildung eines geschlossenen Oberflächenintegrals über den differentiellen Fluss kann die Quellenstärke eines allgemein angenommenen Vektorfeldes berechnet werden.

Die Quellenstärke ist also die Summe aller differentiellen Flüsse dQ

$$Q = \int_{\Omega_1} dQ. \quad (1)$$

Der differentielle Fluss ist das Skalarprodukt des Vektorfeldes mit einem infinitesimal kleinen Flächenstück aus der Oberfläche Γ_1

$$dQ = \mathbf{F} d\mathbf{\Gamma}, \quad (2)$$

woraus folgt:

$$Q = \oint_{\Gamma_1} \mathbf{F} d\mathbf{\Gamma}. \quad (3)$$

b.) Durch Bildung eines Volumsintegrals über differentielle Hüllen kann die Quellenstärke eines allgemein angenommenen Vektorfeldes berechnet werden.

Die differentielle Hülle ist ein infinitesimal kleiner Quader mit den Ausmaßen $dx dy dz$. Das Ziel ist es den differentiellen Fluss durch die einzelnen Flächen der differentielle Hülle zu berechnen.

Wieder sei ein Vektorfeld gegeben

$$\mathbf{F}_x = F \mathbf{e}_x,$$

$$\mathbf{F}_y = F \mathbf{e}_y,$$

$$\mathbf{F}_z = F \mathbf{e}_z.$$

Das Vektorfeld tritt in die Hülle ein (siehe Abbildung 2). Die Flächenstücke der Hülle werden durch ihre Größe und ihrer Flächennormalen definiert

$$d\mathbf{\Gamma}_{-x} = dydz(-\mathbf{e}_x),$$

$$d\mathbf{\Gamma}_{-y} = dx dz(-\mathbf{e}_y),$$

$$d\mathbf{\Gamma}_{-z} = dx dy(-\mathbf{e}_z).$$

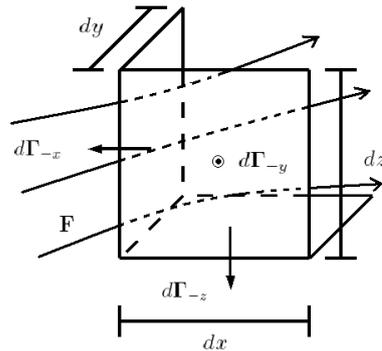


Abbildung 2: Vektorfeld dringt in die Hülle ein

Nun werden die einzelnen differentiellen Flüsse in allen drei Dimensionen berechnet

$$dQ_x = -dydzF_x, \quad (4)$$

$$dQ_y = -dxdzF_y, \quad (5)$$

$$dQ_z = -dxdyF_z. \quad (6)$$

Wenn das Vektorfeld aus der differentiellen Hülle austritt (siehe Abbildung 3), werden wieder die gleichen obigen Schritte durchgeführt. Diesmal sind die Normalvektoren der Flächen parallel zu den Einheitsvektoren des Vektorfeldes. Zusätzlich muss die Änderung des Vektorfeldes nach durchlaufener Strecke berücksichtigt werden

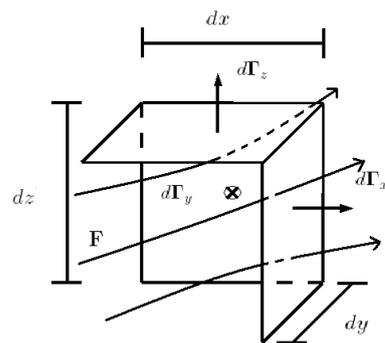


Abbildung 3: Vektorfeld dringt in die Hülle ein

$$\mathbf{F}_{x+dx} = \left(F_x + \frac{\partial F_x}{\partial x} dx \right) \mathbf{e}_x,$$

$$\mathbf{F}_{y+dy} = \left(F_y + \frac{\partial F_y}{\partial y} dy \right) \mathbf{e}_y,$$

$$\mathbf{F}_{z+dz} = \left(F_z + \frac{\partial F_z}{\partial z} dz \right) \mathbf{e}_z,$$

$$d\Gamma_x = dydz \mathbf{e}_x,$$

$$d\Gamma_y = dx dz \mathbf{e}_y,$$

$$d\Gamma_z = dx dy \mathbf{e}_z.$$

Aus dem austretenden Vektorfeld, werden die differentiellen Flüsse berechnet

$$dQ_{x+dx} = dydz F_x + \frac{\partial F_x}{\partial x} dx dy dz, \quad (7)$$

$$dQ_{y+dy} = dx dz F_y + \frac{\partial F_y}{\partial y} dx dy dz, \quad (8)$$

$$dQ_{z+dz} = dx dy F_z + \frac{\partial F_z}{\partial z} dx dy dz. \quad (9)$$

Die Summe aller differentiellen Flüsse ergibt den differentiellen Fluss durch die differentielle Hülle

$$dQ = dQ_x + dQ_y + dQ_z + dQ_{x+dx} + dQ_{y+dy} + dQ_{z+dz},$$

$$dQ = \left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \right) dx dy dz,$$

$$dQ = \left(\begin{array}{c} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{array} \right) \mathbf{F} d\Omega = \nabla \mathbf{F} d\Omega.$$

Die Quellenstärke ist das Integral über alle differentiellen Flüsse im Volumen Ω_1 (siehe Abbildung 1)

$$\Rightarrow Q = \int_{\Omega_1} \nabla \mathbf{F} d\Omega. \quad (10)$$

Setzt man die Quellenstärke des Volumsintegrals mit der des geschlossenen Oberflächenintegrals aus b.) gleich, ergibt sich der **Gaußsche Integralsatz**

$$Q = \oint_{\Gamma_1} \mathbf{F} d\Gamma = \int_{\Omega_1} \nabla \mathbf{F} d\Omega = \int_{\Omega_1} \text{div} \mathbf{F} d\Omega. \quad (11)$$

1.2 Kirchhoff-Helmholtz-Integral

In diesem Abschnitt wird auf das zuvor beschriebene mathematische Wissen zurück gegriffen, welches zu den Greenschen Identitäten führt. Mit der Definition der Greenschen Funktion ergibt sich der Gaußsche Integralsatz im Kirchhoff-Helmholtz-Integral.

2. Greensche Identität

Φ und Ψ sind zweifach differenzierbare Funktionen

$$\int_{\Omega_1} (\Psi \nabla^2 \Phi - \Phi \nabla^2 \Psi) d\Omega = \oint_{\Gamma_1} (\Psi \frac{\partial}{\partial n} \Phi - \Phi \frac{\partial}{\partial n} \Psi) d\Gamma. \quad (12)$$

3. Greensche Identität

Wenn in der 2. Greenschen Identität Ψ mit der Greenschen Funktion g ersetzt wird, erhält man die 3. Greensche Identität. Die Greensche Funktion ist eine Fundamentallösung der Laplace-Gleichung

$$\nabla^2 g = \delta(\mathbf{r} - \mathbf{r}_0).$$

Ausgehend von dem Gaußschen Integralsatz wird nun das Kirchhoff-Helmholtz-Integral hergeleitet.

Damit man die nun bekannten Greenschen Identitäten verwenden kann, wird das Vektorfeld \mathbf{F} mit einem speziell ausgewählten Vektorfeld ersetzt. p und g sind zweifach differenzierbare Funktionen. p ist der Schalldruck und g ist die Greensche Funktion

$$\mathbf{F} = g \nabla p - p \nabla g, \quad (13)$$

$$\begin{aligned}
\oint_{\Gamma_1} (g\nabla p - p\nabla g)d\Gamma &= \int_{\Omega_1} \nabla(g\nabla p - p\nabla g)d\Omega, \\
&= \int_{\Omega_1} (\nabla g\nabla p + g\nabla^2 p - \nabla p\nabla g - p\nabla^2 g)d\Omega, \\
&= \int_{\Omega_1} (g\nabla^2 p - p\nabla^2 g)d\Omega.
\end{aligned} \tag{14}$$

Nun werden die Helmholtzgleichungen für die Ausdrücke p und ρ genauer beschrieben und in die Gleichung 14 eingesetzt.

Helmholtzgleichung des Schalldrucks:

$$(\nabla^2 + k^2)p = 0 \Rightarrow \nabla^2 p = -k^2 p \dots \text{homogenes Schallfeld (Druck)} \tag{15}$$

Helmholtzgleichung der Greenschen Funktion:

$$(\nabla^2 + k^2)g = -\delta(\mathbf{r} - \mathbf{r}_0) \Rightarrow \nabla^2 g = -k^2 g - \delta(\mathbf{r} - \mathbf{r}_0) \dots \text{inhomogenes Feld} \tag{16}$$

Durch Einsetzen in Gleichung 14 verbleibt bei der Integration alleine der Schalldruck

$$\int_{\Omega_1} (g\nabla^2 p - p\nabla^2 g)d\Omega = \int_{\Omega_1} (-gk^2 p + pk^2 g + p\delta(\mathbf{r} - \mathbf{r}_0))d\Omega = p.$$

Das Kirchhoff-Helmholtz-Integral

Setzt man p wieder mit dem Gaußschen Integralsatz gleich, erhält man über die 2. Greensche Identität das Kirchhoff-Helmholtz-Integral

$$p = \oint_{\Gamma_1} (g\nabla p - p\nabla g)d\Gamma = \oint_{\Gamma_1} (g\frac{\partial}{\partial n} p - p\frac{\partial}{\partial n} g)d\Gamma. \tag{17}$$

$\frac{\partial}{\partial n} p$... Normalkomponente der Schallschnelle
 $\frac{\partial}{\partial n} g$... Dipolquelle

Das Kirchhoff-Helmholtz-Integral besagt, dass bei Kenntnis des Drucks p und der Normalkomponente der Schallschnelle $\frac{\partial p}{\partial n}$ an der Oberfläche eines Volumens, der Druck in jedem Punkt innerhalb dieses Volumens definiert ist. Jedes Schallfeld kann durch Monopol- und Dipolquellen auf der geschlossenen Oberfläche synthetisiert werden. Entlang dieser Oberfläche ist die Stärke der Monopolquelle durch die Normalkomponente der Schallschnelle und jene der Dipolquellen ist durch den Schalldruck gegeben (siehe Abbildung 4). [SA08c]

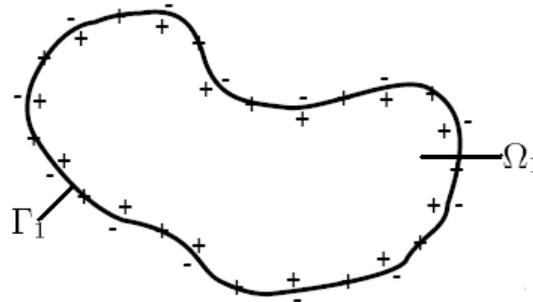


Abbildung 4: Oberfläche mit Monopol- und Dipolquellen

In der Elektroakustik gibt es keine kontinuierlich verteilten idealen Monopol- und Dipolquellen. Deshalb wird das Kirchhoff-Helmholtz-Integral bei der praktischen Umsetzung nur angenähert und vereinfacht. In der Wellenfeldsynthese verzichtet man auf die Dipolquellen indem man den Schalldruck auf der Oberfläche 0 setzt. Für die Monopollautsprecher werden Lautsprecher mit einem geschlossenen Gehäuse verwendet. Durch den Verzicht auf die Dipolquellen gibt es auch Folgen. Außerhalb des Volumens eliminieren sich nicht mehr die Schallfelder der Monopol- und Dipolquelle. Das Schallfeld innerhalb des Volumens wird nur noch von der Schallschnelle, also der Dipolquelle bestimmt. Außerdem wird eine künstliche Reflexionsstelle erzeugt, da der Schalldruck der Oberfläche mit 0 angenommen wird.

1.3 Die Schallschnelle

Wir wissen nun, dass man bei der Wellenfeldsynthese das Schallfeld innerhalb eines Volumens durch die Normalkomponente der Schallschnelle am Rand dieses Volumens synthetisieren kann. Wenn man den Schalldruck an der Oberfläche des Volumens 0 setzt, bleibt der Term $g \frac{\partial}{\partial n} p$ des Kirchhoff-Helmholtz-Integrals übrig. Zwischen dem Schalldruck und der Schallschnelle gibt es einen Zusammenhang. Die maximale Druckänderung und die Beschleunigung der Luftteilchen sind entgegengesetzt gerichtet

$$\rho_0 \dot{\mathbf{v}} = -\nabla p. \quad (18)$$

Formt man diese Gleichung nach der Beschleunigung der Luftteilchen um, integriert und transformiert sie in den Laplace-Bereich, resultiert die Schallschnelle

$$\mathbf{v} = -\frac{1}{\rho_0} \int \nabla p dt, \quad (19)$$

$$\mathbf{v}_{(s)} = -\frac{1}{\rho_0} \frac{\nabla p}{s}. \quad (20)$$

Zu beachten ist das dieses Resultat die Schallschnelle und noch nicht die Normalkomponente der dazugehörigen Begrenzungsfläche ist. Bei der Wellenfeldsynthese wird im vereinfachten Kirchhoff-Helmholtz-Integral folgender Zusammenhang verwendet

$$\frac{\partial}{\partial n} p = \langle \mathbf{n}, \nabla p \rangle = [\langle \mathbf{n}, \mathbf{v} \rangle (-s\rho_0)]. \quad (21)$$

Für die vollständige Definition der Schallschnelle $\mathbf{v}_{(s)}$ muss noch die Druckänderung ∇p bestimmt werden.

Der Schalldruck p einer Punktquelle (Kugelwelle) ist gegeben mit:

$$p = \frac{e^{ikr}}{r} \quad (22)$$

$$\dots k = \frac{\omega}{c}$$

r ist der Abstand der virtuellen Quelle zu dem jeweiligen Lautsprecher

$$r = \sqrt{(x_{LS} - x_Q)^2 + (y_{LS} - y_Q)^2 + (z_{LS} - z_Q)^2}.$$

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} = \begin{pmatrix} \frac{\partial r}{\partial x} \\ \frac{\partial r}{\partial y} \\ \frac{\partial r}{\partial z} \end{pmatrix} \frac{\partial}{\partial r}$$

$$\frac{\partial r}{\partial x} = \frac{x_{LS} - x_Q}{r}$$

$$\frac{\partial r}{\partial y} = \frac{y_{LS} - y_Q}{r}$$

$$\frac{\partial r}{\partial z} = \frac{z_{LS} - z_Q}{r}$$

$$\Rightarrow \nabla p = \frac{1}{r} \begin{pmatrix} (x_{LS} - x_Q) \\ (y_{LS} - y_Q) \\ (z_{LS} - z_Q) \end{pmatrix} \frac{\partial}{\partial r} \frac{e^{jkr}}{r} = \frac{1}{r} (\mathbf{x}_{LS} - \mathbf{x}_Q) \left(\frac{ikr - 1}{r^2} \right) e^{ikr} \quad (23)$$

$$= \frac{\mathbf{x}_{LS} - \mathbf{x}_Q}{r^3} (ikr - 1) e^{ikr}$$

Die Druckänderung wird in den Laplace-Bereich transformiert.

$$ikr = \frac{s}{c}r$$

$$\frac{1}{s}\nabla p = \frac{\mathbf{x}_{\text{LS}} - \mathbf{x}_{\text{Q}}}{r^3} \left(\frac{r}{c} - \frac{1}{s} \right) e^{\frac{s}{c}r} \quad (24)$$

2 Implementation

Aus den theoretischen Betrachtungen erhalten wir die Driving Function der einzelnen Lautsprecher. Es wird die Normalkomponente der Schallschnelle von den einzelnen Lautsprechern (stellen unsere Annäherungen an die Monopolquelle dar) wiedergegeben

$$\mathbf{v} = \frac{1}{s}\nabla p = \frac{(\mathbf{x}_{\text{LS}} - \mathbf{x}_{\text{Q}})}{r^3} \left(\frac{r}{c} - \frac{1}{s} \right) e^{\frac{s}{c}r}, \quad (25)$$

$$v_{\mathbf{n}} = \mathbf{n}^T \mathbf{v} = \frac{\mathbf{n}^T (\mathbf{x}_{\text{LS}} - \mathbf{x}_{\text{Q}})}{r^3} \left(\frac{r}{c} - \frac{1}{s} \right) e^{\frac{s}{c}r}. \quad (26)$$

Die Driving Function besteht aus drei Teilen:

1. einem Amplitudenteil
2. einem Filter
3. einem Delay

Diese drei Teile gilt es zu implementieren, zusätzlich kommen aufgrund der räumlichen Gegebenheiten und der theoretischen Vereinfachungen noch eine Flächenberücksichtigung und eine Fensterung hinzu (nur ein Teil der Lautsprecher sollte spielen).

2.1 Filter

$$\left(\frac{r}{c} - \frac{1}{s} \right) = \left(\frac{\frac{s}{c}r - 1}{s} \right) = \frac{1}{a} \left(1 - \frac{a}{s} \right) \text{ mit } a = \frac{c}{r} \quad (27)$$

Die vorliegende Formel liegt im Laplacebereich. Zur Implementierung muss sie allerdings über den Zeitbereich in den Z-Bereich transformiert werden. Hierfür wird der Ausdruck umgeformt um Standard Umformungen der Laplace bzw. Z-Transformation nutzen zu können (in entsprechenden Tabellen können die Ausdrücke nachgeschlagen werden) (siehe [OSB99]).

$$1 \bullet \text{---} \circ \text{---} \bullet \delta(0), \quad (28)$$

$$-\frac{a}{s} \bullet \text{---} \circ \text{---} \bullet - T \frac{a}{1 - z^{-1}}, \quad (29)$$

$$H(z) = \frac{r}{c} - \frac{T}{1 - z^{-1}}. \quad (30)$$

Nach der Transformation erhalten wir die Übertragungsfunktion für unser Filter. Probleme bereitet uns hier die Polstelle bei $z=1$, da das Filter durch die endliche Registerlänge der jeweilig verwendeten Recheneinheit instabil wird.

2.1.1 Repolarisierung des Filters

Um dem entgegenzuwirken nähern wir das Filter an. D.h. wir erweitern es um ein Notchfilter mit der Übertragungsfunktion:

$$H(z) = \frac{1 - z^{-1}}{1 - rz^{-1}} \text{ mit } r = ?. \quad (31)$$

Dabei muss mittels Annäherung ein Wert für die neue Polstelle abgeschätzt werden (siehe Abbildung 5). Die Polstelle bei $z=1$ wird nach innen verschoben und zwar wollen wir sie so

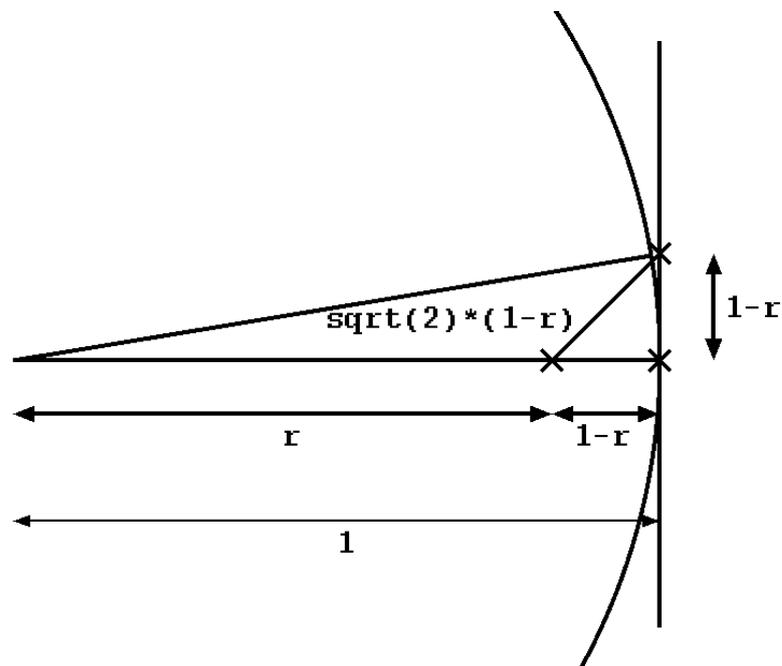


Abbildung 5: Implementierung des Notch Filters

weit nach innen verschoben, bis unsere Grenzfrequenz (bei gewünschten 60Hz) erreicht ist. Dies ist genau dann der Fall wenn die Seitenlängen des eingezeichneten Dreiecks folgende Werte erreichen:

- Katheten: $(1 - r)$
- Hypotenuse: $\sqrt{2}(1 - r)$

Warum können wir hier ein Dreieck einzeichnen, obwohl wir uns doch eigentlich auf einem Kreis bewegen? Die Polstelle befindet sich sehr nahe am Einheitskreis, deshalb kann der Kreisbogen durch eine Tangente angenähert werden. Nachdem wir die Seitenlängen des Dreiecks berechnet haben, können wir unsere Gleichung ansetzen (zu Erinnerung: die Grenzfrequenz sollte bei 60Hz liegen) und erhalten die angenäherte Polstelle

$$\Delta\omega = (1 - r), \quad (32)$$

$$\frac{2\pi f}{f_s} = (1 - r), \quad (33)$$

$$r = 1 - \frac{2\pi 60}{44100} = 0.99145. \quad (34)$$

In Abbildung 8 ist der Unterschied zwischen originaler und angenäherter Übertragungsfunktion dargestellt. Die implementierbare (da stabile) Übertragungsfunktion ergibt sich also zu:

$$H(z) = \frac{r}{c} - \frac{T}{1 - 0.992 \cdot z^{-1}}. \quad (35)$$

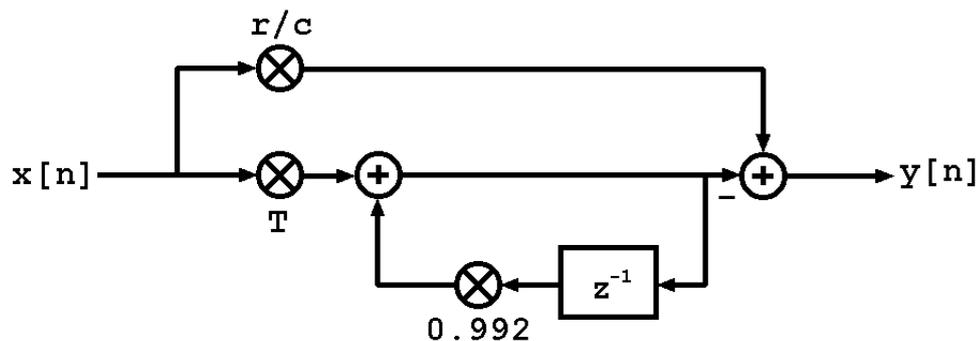


Abbildung 6: Signalflussplan

In der uns gewohnten Darstellung als Signalflussplan (Abbildung 6) und als Implementation in pd erhalten wir Abbildung 7. Wobei das Objekt „rpole“ eine reelle Polstelle bei 0.992 darstellt.

Damit wäre der erste Teil unserer Wellenfeldsynthese, das Filter, implementiert.

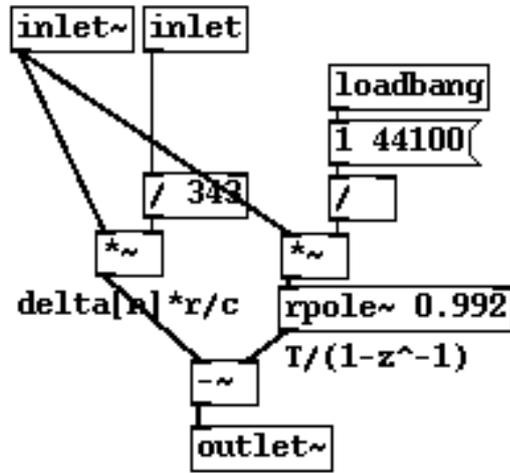


Abbildung 7: In pd implementiertes Filter

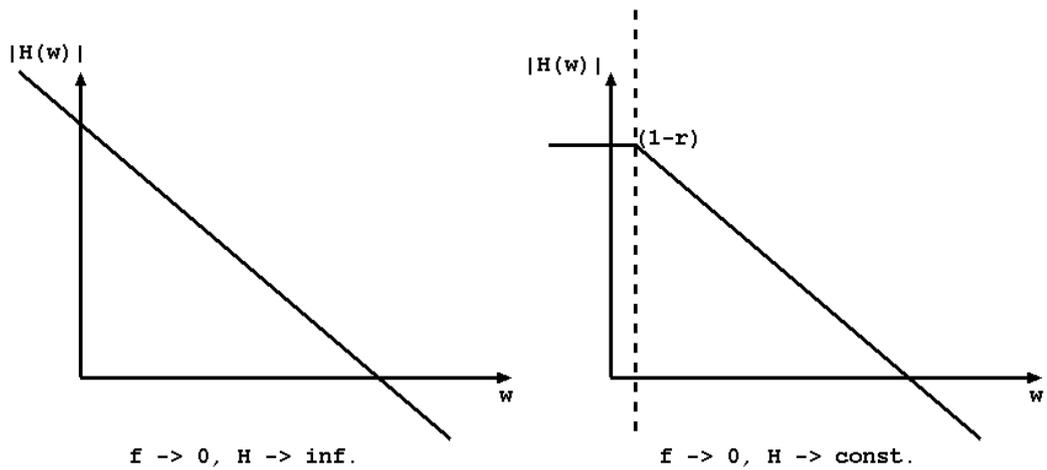


Abbildung 8: Frequenzdomänenarstellung des repolarisierten Filters

2.2 Fensterung

Um eine künstliche Reflexionsstelle zu vermeiden, spielen nicht alle vorhandenen Lautsprecher sondern nur die, die von der Quelle wegspielen (die der Quelle also abgewandt sind, siehe Abbildung 9 und Abbildung 10). Warum? Durch Vereinfachung des Kirchhoff-Helmholtz-Integrals, in welchem nur der Teil, der die Normalkomponente der Schallschnelle mit Monopolquellen reproduziert, verwendet wurde und der Schalldruck an den Begrenzungsflächen gleich NULL gesetzt wurde, ergeben sich an den der Quelle zugewandten Lautsprechern unerwünschte Reflexionsstellen. Um diesen entgegenzuwirken werden die jeweiligen Lautsprecher einfach gemutet (siehe [SA08c]). Erreicht kann das durch folgende Bedingung werden:

$$a_{SW}(\mathbf{x}_0) = \begin{cases} 1 & , \text{if } \langle \mathbf{x}_{LS} - \mathbf{x}_Q, \mathbf{n}(\mathbf{x}_{LS}) \rangle > 0, \\ 0 & , \text{otherwise.} \end{cases} \quad (36)$$

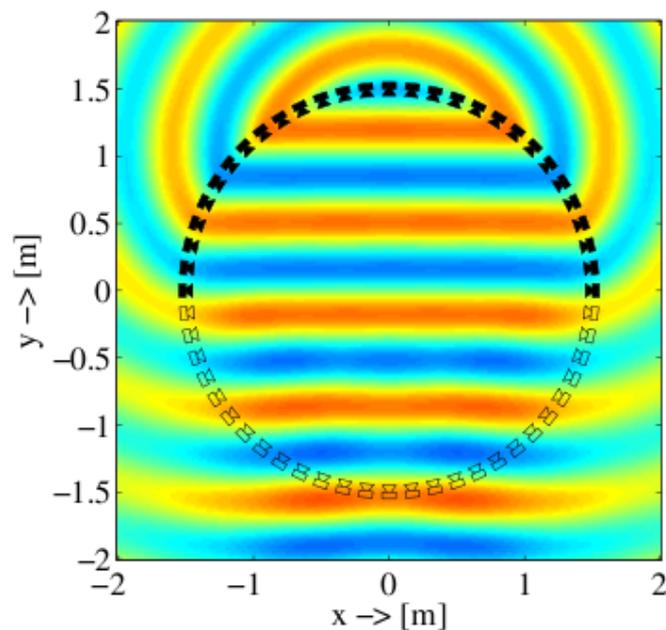


Abbildung 9: Fensterfunktion [SA08c]

2.3 Amplitude

$$Amplitude = \frac{\mathbf{n}^T(\mathbf{x}_{LS} - \mathbf{x}_Q)}{r^3} \quad (37)$$

Die Berechnung der Amplitude gestaltet sich sehr einfach. Der Abstand zwischen Lautsprecher und Quelle ist ausschlaggebend (siehe Abbildung 10).

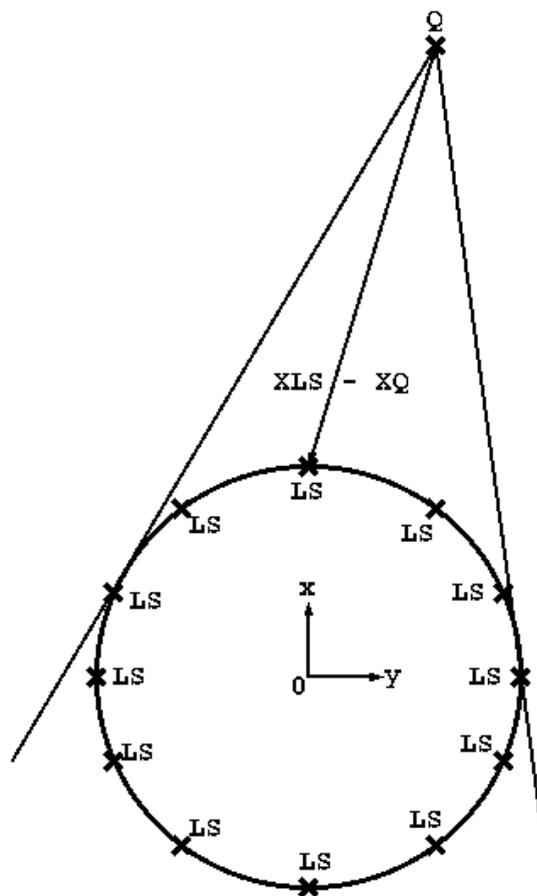


Abbildung 10: Fensterfunktion, Radiusberechnung

2.4 Delay

Es folgt die Implementierung des Delays. Durch die unterschiedlichen Entfernungen der einzelnen Lautsprecher von der Primärquelle ergeben sich für jeden Lautsprecher vom Radius r (geteilt durch die Schallgeschwindigkeit c) abhängige Delayzeiten.

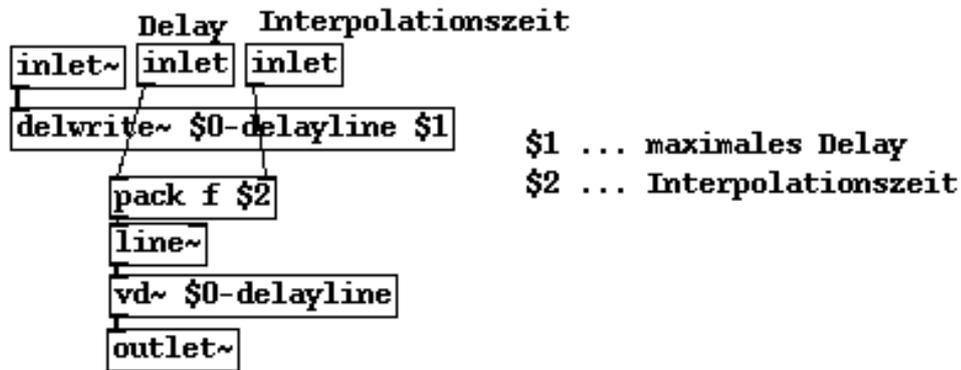


Abbildung 11: In pd implementiertes Delay

Die Implementierung in pd (siehe Abbildung 11) erfolgt mittels Delayline. Die Audiodaten werden mit dem Objekt „delwrite“ in eine Delayline (der maximaler Größe 100ms) geschrieben und mittels „vd“ Objekt verzögert wieder ausgelesen. Das Objekt „vd“ wird im Gegensatz zum Objekt „delread“ mit Audiodaten gesteuert. Zusätzlich erfolgt in diesem pd-subpatch eine Interpolation zwischen den Delaywerten. Sollte heißen: Bei bewegten Quellen ändert sich die Dauer des Delays, der Wert nimmt zu bzw. ab. Um eine sprunghafte Änderung zwischen zwei Werten zu verhindern (eine sprunghaft Änderung hätte Störgeräusche zur Folge), wird zwischen den Werten, mit dem Objekt „line“, interpoliert. Die Interpolationszeit beträgt in unserem Fall 20ms.

Ein kleines Beispiel zu Verdeutlichung: Die Primärquelle entfernt sich von 1m auf 20m. Das entspricht einer Delayänderung von $\Delta t = \frac{20}{c} - \frac{1}{c} = 55ms$ mit $c = 343m/s$. Dieser Wert von Δt wird nun nicht sprunghaft geändert sondern ändert sich kontinuierlich in 20ms von 3ms auf 58ms. Somit werden Störgeräusche bei bewegten Quellen verhindert.

2.5 Flächengewichtung

Voronoi, Delaunay: In unserem Fall wurde die Wellenfeldsynthese für den CUBE implementiert. Da die einzelnen Lautsprecher im CUBE nicht regelmäßig verteilt sind, versuchten wir mittels Gewichtung der einzelnen Lautsprecher eine energetisch gleichmäßige Verteilung zu erreichen

$$\int_S v_n(\mathbf{r}_q)g(\mathbf{r}_q, \mathbf{r})dS(\mathbf{r}_q) \rightarrow \sum_l v_n[l]g[\mathbf{r}_l, \mathbf{r}]S[l]. \quad (38)$$

Hierfür bedienten wir uns eines Tools von Jens Keiner der Universität Lübeck, in welchem die Berechnung von Delaunay Triangulationen und Voronoi Pakettierung auf der Kugel durchgeführt wird [Kei07].

Die Zuordnung der Punkte einer Kugeloberfläche erfolgt dabei nach folgendem Prinzip: Eine Voronoi Zelle enthält alle Punkte welche einem Lautsprecher am nächsten sind. Nachdem die Gewichtungen der einzelnen Zellen berechnet ist, tritt ein neues Problem auf. Da die Lautsprecher im CUBE nur auf einer Halbkugel angeordnet sind (in der unteren Halbkugel befinden sich keine Lautsprecher), sind diejenigen Voronoi Flächen die sich am Äquator der Kugel befinden nicht mit dem Äquator abgeschlossen (Halbkugel wäre jedoch erwünscht). Um dem entgegenzuwirken mussten die am Äquator liegenden Punkte um den Äquator in die untere Halbkugel gespiegelt werden. Somit sind die Flächen genau am Äquator begrenzt und uns stehen die einzelnen Flächengewichtungen nun korrekt (wenn auch nur, wie oben erwähnt, in einer Annäherung) zur Verfügung (siehe Abbildung 13).

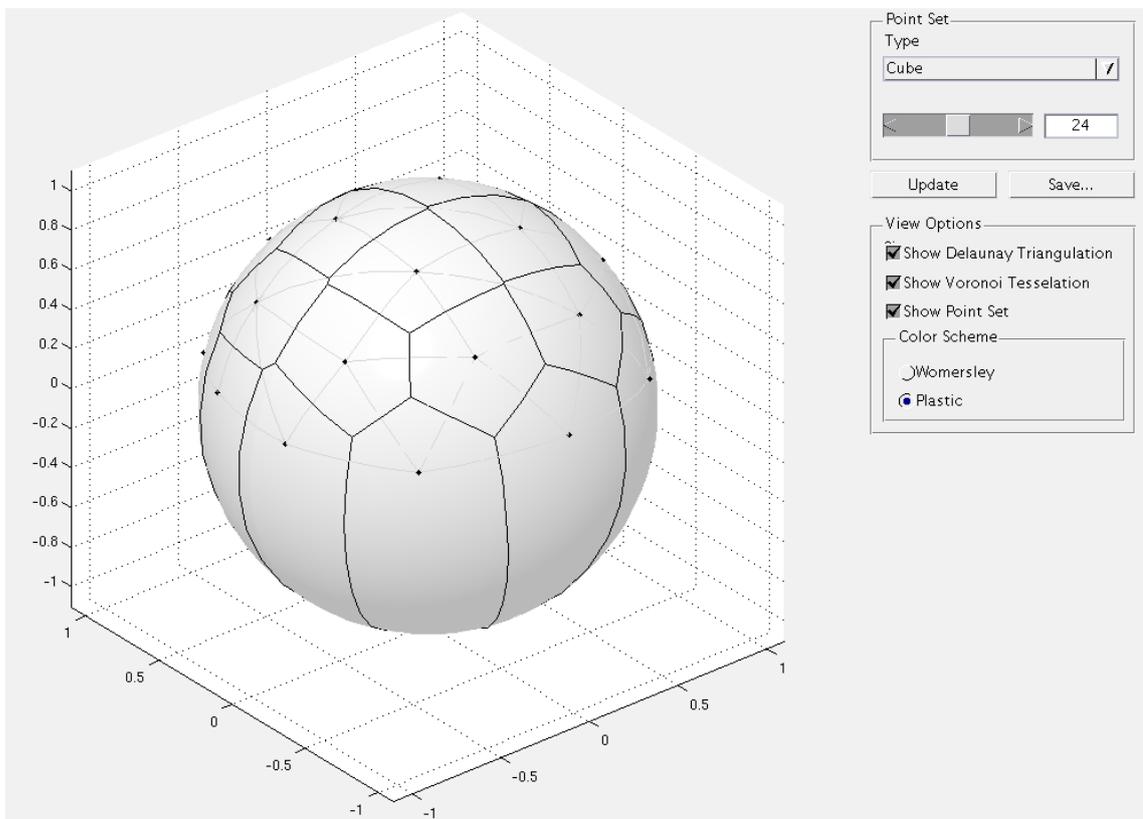


Abbildung 12: Delaunay Triangulationen und Voronoi Pakettierung bei 24 Lautsprechern [Kei07]

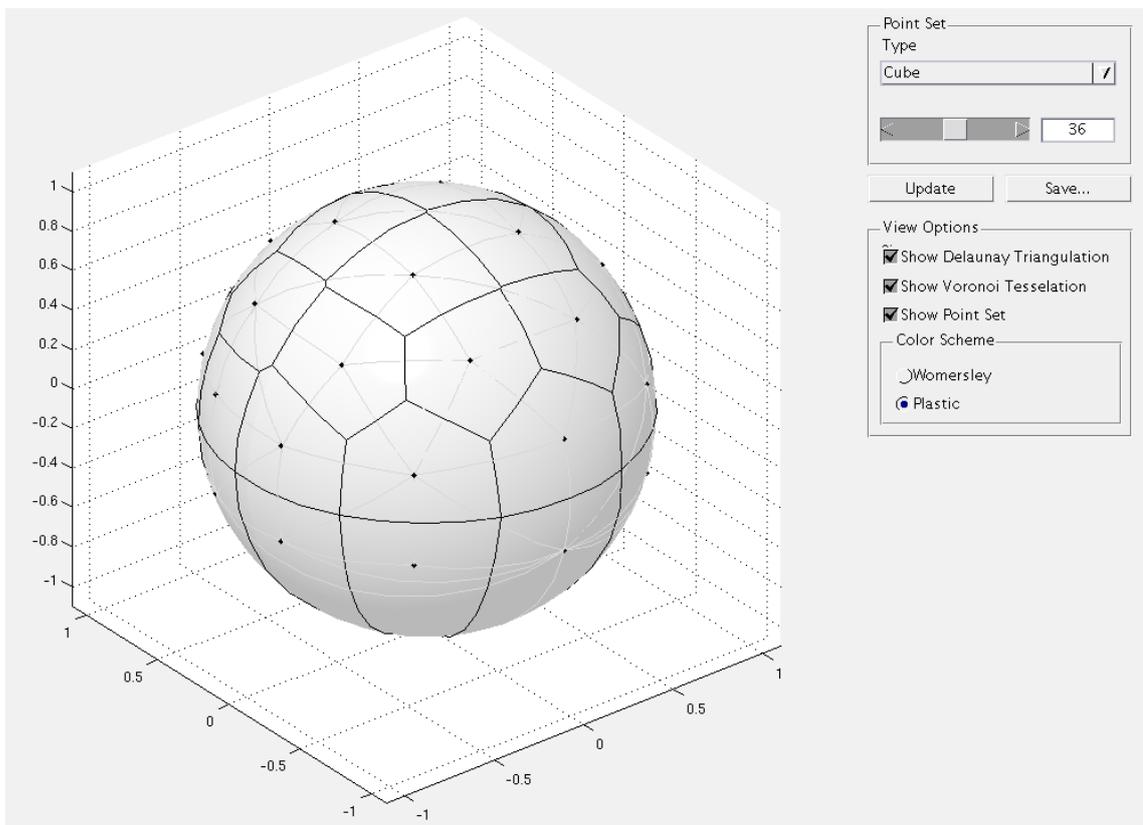


Abbildung 13: Delaunay Triangulationen und Voronoi Pakettierung mit gespiegelten Lautsprechern [Kei07]

2.6 Blockschaltbild

Nachdem wir die einzelnen Teile der Implementation besprochen haben, wollen wir uns das komplette Blockschaltbild in Abbildung 14 ansehen.

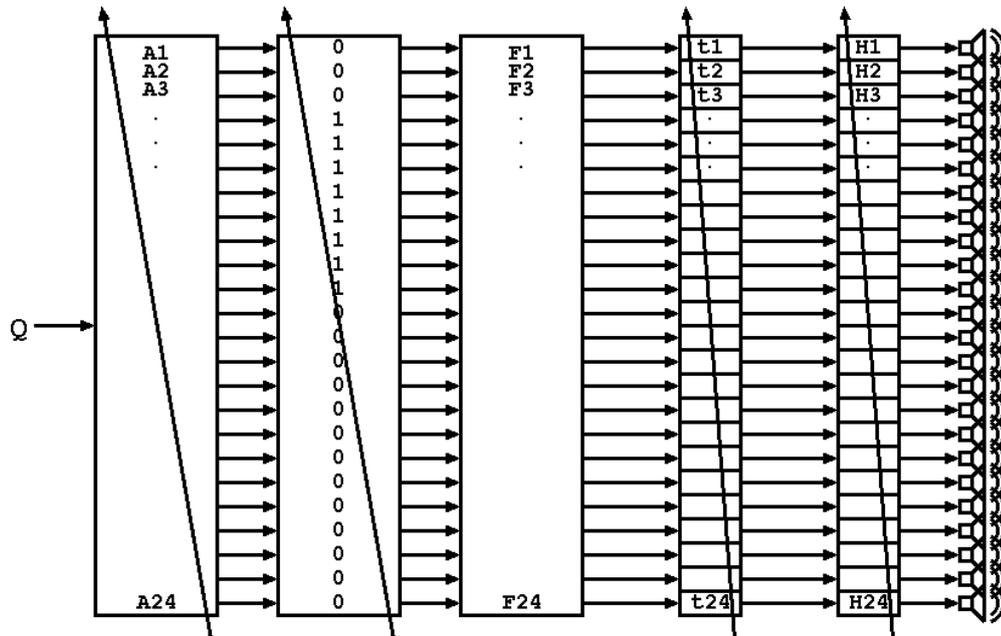


Abbildung 14: Blockschaltbild

Wir haben ein Quellsignal welches mit einem vom Abstand Primär- zu Sekundärquelle abhängigen Faktor belegt wird. Die in der Amplitude unterschiedlichen Signale werden des weitern je nach Position der Lautsprecher zur Primärquelle gefenstert, das heißt zu NULL gesetzt oder unverändert weitergeleitet. Es folgt die Gewichtung der einzelnen Lautsprechersignale laut zuständiger Fläche auf der Halbkugel. Die Gewichtungsfaktoren liegen hier innerhalb der Werte 0.2 - 0.35. Weiters erfahren die einzelnen Signale eine von der Distanz Primär- zu Sekundärquelle abhängige Verzögerung. Das nachgeschaltete Filter verfärbt unsere Signale wie gewünscht und über die einzelnen Lautsprecher wird das Wellenfeld synthetisiert.

3 Mögliche Artefakte

3.1 Dopplereffekt

Die fertig implementierte Wellenfeldsynthese wurde im CUBE getestet und dabei konnte bei bewegten Quellen eine Art Dopplereffekt (siehe Abbildung 15) wahrgenommen werden [SA08b]. Passend zu dieser Beobachtung gibt es ein Paper von Jens Arens und Sascha Spors von den Deutschen Telekom Laboratories. Sie untersuchten darin die Auswirkungen des Dopplereffekts auf ein Linienarray von Monopolquellen bei bewegten Sekundärquellen. Dabei stellte sich heraus, dass sich der natürliche Dopplereffekt von dem in der Wellenfeldsynthese unterscheidet:

1. natürlicher Dopplereffekt: -die Frequenz f der Quelle steigt schneller als die Geschwindigkeit v der Quelle und erreicht bei der Schallgeschwindigkeit den theoretischen Wert von unendlich.
2. Dopplereffekt in der WFS: -die Geschwindigkeit v der bewegten Quelle steigt linear mit der Frequenz f der Quelle (wird v unendlich, wird auch f unendlich);

In Abbildung 16 ist dies bei genauer Betrachtung erkennbar. Prinzipiell leidet das Diagramm am Kompromiss zwischen Frequenz- und Zeitauflösung der Short-Time-Fourier-Transformation, durch welche auch die niederenergetischen Artefakte des Diagramms erklärbar sind. Außerdem ist der Unterschied des natürlichen Dopplereffekts gegenüber dem der Wellenfeldsynthese bei moderaten Geschwindigkeiten (hier: $v = 40\text{m/s}$) äußerst gering.

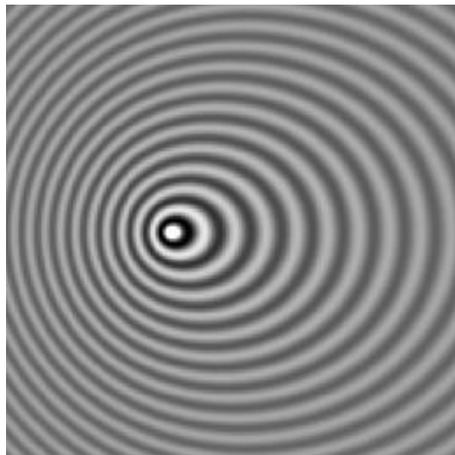
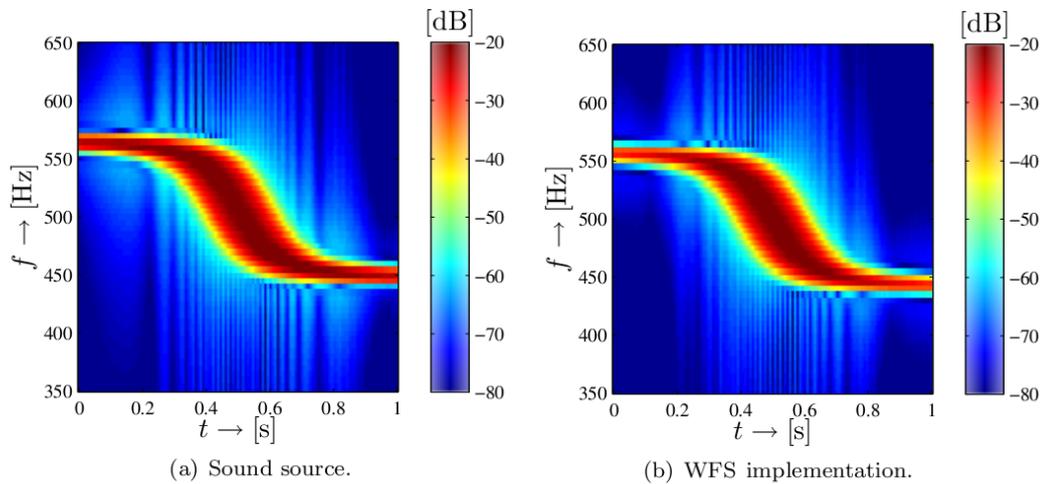


Abbildung 15: Der natürliche Dopplereffekt (Quelle bewegt sich nach links) [SA08b]

Beim natürlichen Dopplereffekt ist es nicht entscheidend, ob sich die Quelle bewegt oder die Abhörposition. Beim synthetisierten Wellenfeld ist dies hingegen anders. Während sich bei Bewegung der Abhörposition der natürliche Dopplereffekt einstellt, ist dies bei Bewegung der Quelle nicht der Fall. Hier wird entgegen dem natürlichen Fall nicht die



Spectrograms of a virtual source travelling with $v = 40$ m/s and $f_s = 500$ Hz observed at $\mathbf{x}_{\text{mic}} = [0 \ 4]^T$ m. The WFS array has an overall length of 50 m and is located at $y_0 = 1$ m.

Abbildung 16: Natürlicher Dopplereffekt vs. Dopplereffekt in der WFS [SA08b]

primäre Schallquelle auf einen zubewegt, sondern die Sekundärquellenanordnung Richtung primäre Quelle hinbewegt. Stellt dies bei nur einer primären Quelle noch keinen großen Unterschied zum natürlichen Dopplereffekt dar, ist es bei mehreren Primärschallquellen durchaus zu beachten, da sich die Anordnung dann natürlich nur in eine Richtung bewegen kann. Eine genauere Untersuchung dieses Phänomens steht allerdings noch aus.

3.2 Truncation Error

Bei nicht geschlossener Anordnung der Sekundärquellen treten truncation Errors auf. Wir wollen das Beispiel eines Linienarray betrachten. So müsste ein Linienarray laut Theorie unendliche Abmessungen haben. Da dies praktisch nicht erfüllbar ist, treten an den jeweiligen Enden des Linienarray Artefakte auf. Bei bewegten Quellen äußern sich diese Artefakte als Echos (siehe Abbildung 17). Verringert kann dieser Error durch Tapering werden. Dabei wird die Wiedergabe der Quelle an den Enden des Arrays räumlich ausgefadet. Dies kann wie in Abbildung 17 zum Beispiel mittels Cosine tapering durchgeführt werden [SA08b].

3.3 Spatial Aliasing

Der dritte hier besprochene Fehler tritt durch spatial Aliasing (räumliches Aliasing) auf. Uns stehen zur Implementation der Wellenfeldsynthese keine kontinuierlichen Schallquellen zur Verfügung, deshalb kommt es zu einer Diskretisierung. Die Sekundärquellen werden durch einzelne Lautsprecher, die in einem gewissen Abstand zueinander stehen realisiert. Günstigstenfalls betragen die Abmessungen zwischen den einzelnen Lautsprechern ca. 10-30cm was eine ungefähre Nyquistfrequenz von 1kHz (das spatial Aliasing

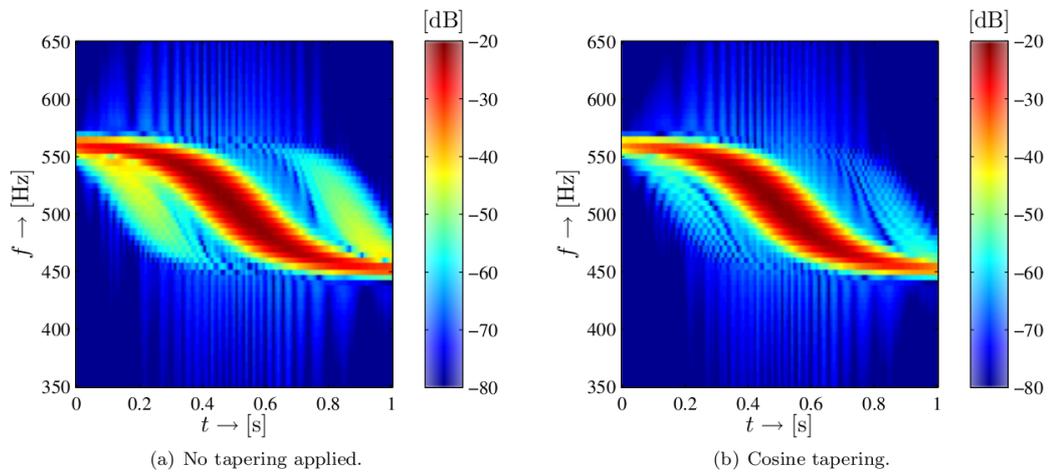


Abbildung 17: Truncation Error bei bewegten Quellen [SA08b]

betreffend) zur Folge hätte [SA08a]

$$f \leq \frac{c}{2\Delta x \sin \phi}. \tag{39}$$

Wie kommt diese Nyquistfrequenz zu stande? Betrachten wir hierfür mit Abbildung 18 ein adäquates jedoch leichter vorstellbares Problem. Ein Schallfeld wird durch Mikrofone

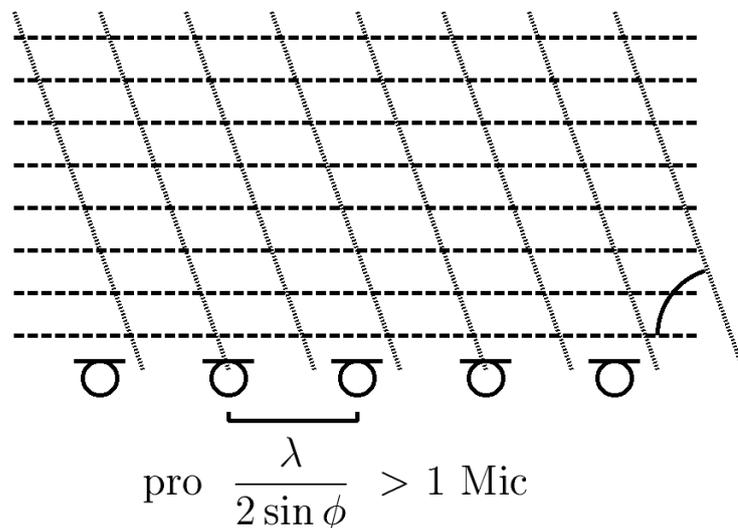


Abbildung 18: Nyquistfrequenz

abgetastet. Trifft die Wellenfront von vorne ($\phi = 0$) auf das Mikrofonarray ergibt sich räumlich betrachtet bei der Abtastung kein Problem. Trifft die Wellenfront hingegen von der Seite ein ($\phi = 90^\circ$), muss für jede halbe Wellenlänge ($\lambda/2$) mehr als ein Mikrofon zur Verfügung stehen um eindeutig reproduzieren zu können. In der vorliegenden Skizze

(Abbildung 18) würde das Abtasttheorem nicht erfüllt werden, es käme demnach zu räumlichem Aliasing.

In der Wellenfeldsynthese kann folgendes beobachtet werden:

1. Je mehr Bandbreite die Quelle hat, desto mehr Artefakte treten auf.
2. Das spatial Aliasing hängt von der Abhörposition ab.

Außerdem: Das Gehör scheint die Ortung betreffend nicht allzu sensibel auf spatial Aliasing zu sein. Die Artefakte äußern sich hingegen in Klangverfärbungen.

Wollen wir Aussage 1) anhand des Vergleiches von Higher Order Ambisonics (HOA) mit der Wellenfeldsynthese (WFS) betrachten (siehe Abbildung 19): Bei kontinuierlichen Sekundärquellen ist es bei HOA möglich durch Vorgabe der Ordnung (hier: 25. Ordnung) eine räumliche Bandbegrenzung zu erreichen. Durch die Diskretisierung kommt es zu räumlichen Spiegelfrequenzen, welche sich im Falle von HOA aufgrund der räumlichen Bandbegrenzung nicht überlagern. Bei der WFS kommt es zu keiner räumlichen Bandbegrenzung, deshalb überlagern sich die einzelnen Spiegelfrequenzen im diskretisierten Fall.

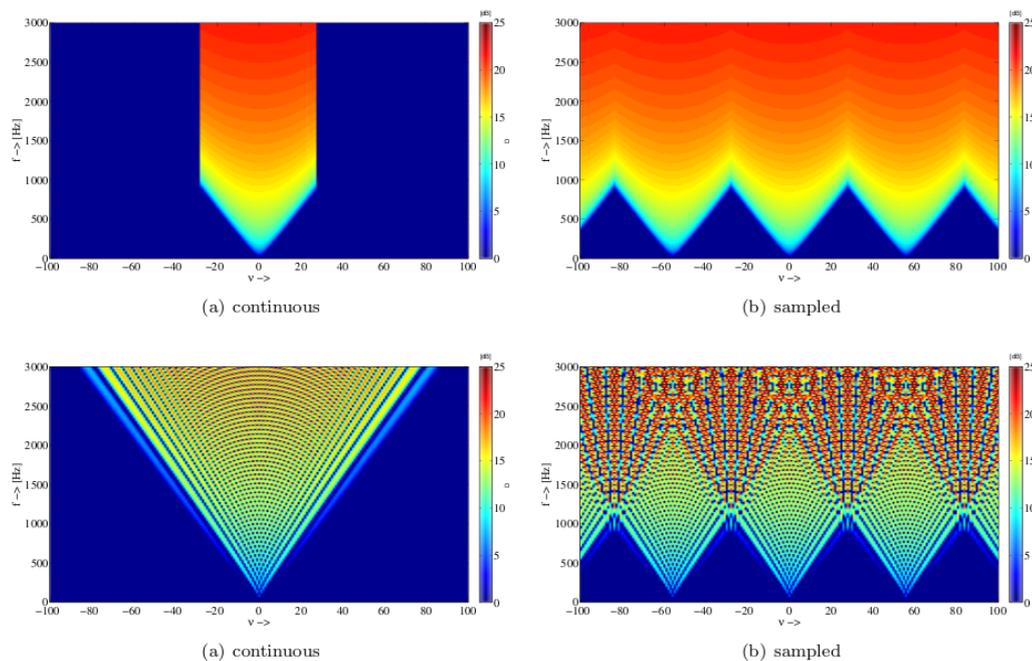


Abbildung 19: Vergleich HOA vs. WFS [SA08a]

Zu Aussage 2) (siehe Abbildung 20): Bei HOA wird der fehlerfrei reproduzierte Bereich unseres Wellenfeldes mit zunehmender Frequenz kleiner (Sweet Spot wird kleiner). Bei der WFS rutscht der fehlerfrei reproduzierte Bereich hingegen immer weiter nach hinten. Anmerkung: Im vorliegenden Diagramm (Abbildung 20) sollte ein ebenes Schallfeld reproduziert werden.

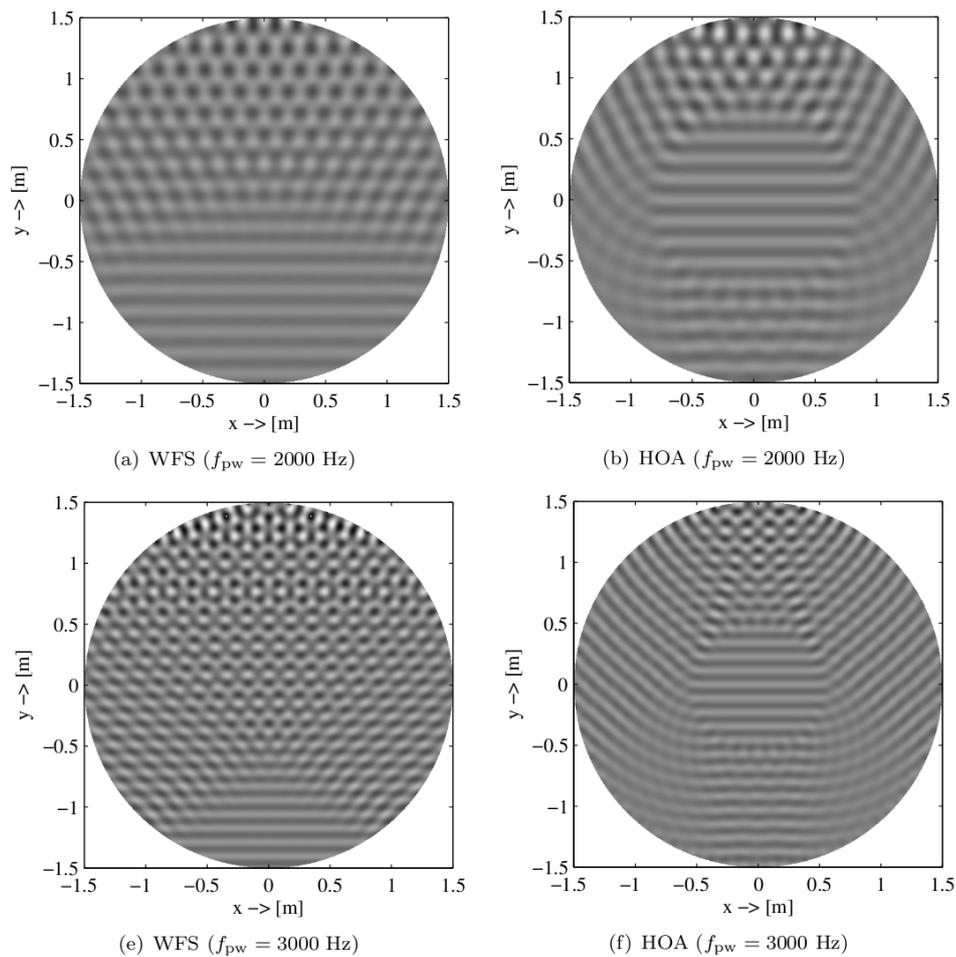


Abbildung 20: Ortsabhängigkeit [SA08a]

4 Hörversuch im CUBE

Um die Funktionalität der Implementierung der WFS im CUBE bewerten zu können, wurde ein Hörversuch durchgeführt. Als wesentliche Empfindungsgröße wurde die Genauigkeit der Lokalisation der Primärquellen gewählt. Obwohl in dieser Seminararbeit eine 3D Wellenfeldsynthese implementiert wurde, wurde für den Hörversuch nur eine vereinfachte 2D Implementation verwendet. Für eine vollständige 3D Umsetzung wäre eine kugelförmige Lautsprecheranordnung nötig. Im CUBE gibt es eine obere Halbkugel-

gelanordnung mit 24 Lautsprechern. Dies hatte bei ersten Hörversuchen (3D) eine Verschiebung, der in der Horizontalen positionierten Primärquellen, nach oben zur Folge. Ursache für diese Verschiebung ist der Ansatz der WFS, welche eine Wellenfront aus den einzelnen Wellen der Lautsprecher entlang einer geschlossenen Oberfläche zusammensetzt. Die Lautsprecher unterhalb der horizontalen Ebene fehlen. Dies verschiebt das synthetisierte Wellenfeld zwangsläufig zur oberen Halbkugel.

Die Anordnung in der Horizontalebene mit 12 Sekundärquellen ist in Abbildung 21 dargestellt.

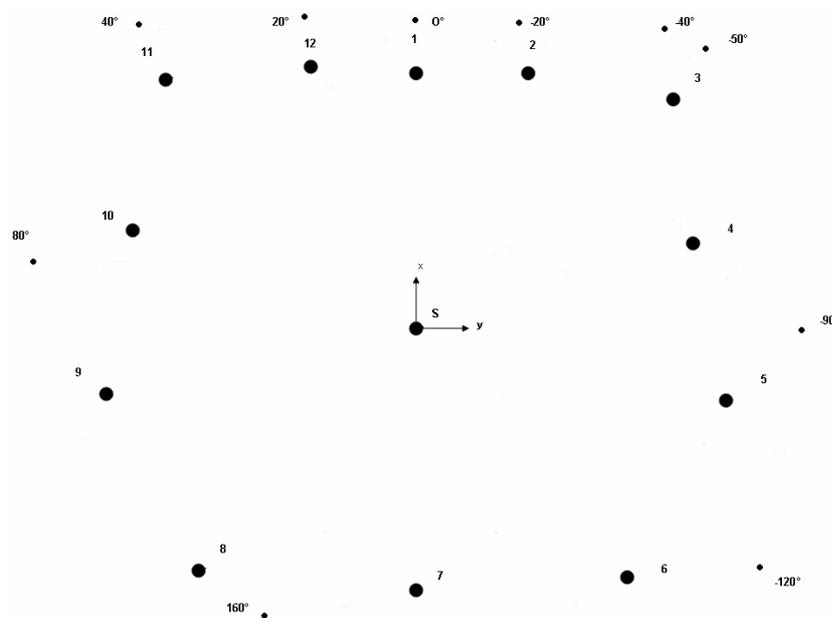


Abbildung 21: 2D Lautsprecheranordnung (große Punkte) mit den Positionen der Primärquellen (kleine Punkte)

4.1 Vorbereitungen und Überlegungen

Zur Durchführung des Hörversuches wurde das Infrarot Trackingsystem im CUBE verwendet. Dieses System kann mit Infrarotkameras die Position und Lage eines markierten Zielgerätes erkennen [Fra08]. Die Versuchsperson zielt mit dem Zielgerät auf die wahrgenommene Primärquelle und betätigt den Auslöser. Die Software berechnet ausgehend vom Ausgangspunkt den Azimut und Elevationswinkel des Durchtrittspunktes. Dieser ist der Schnittpunkt der vermuteten Richtung der Quelle mit der Fläche zwischen den Lautsprechern. Der Ausgangspunkt dieses Systems ist der Sweetspot, der genau der 0 Punkt der gesamten Versuchsanordnung ist (siehe Abbildung 21). Zu beachten ist, dass er auf einer Höhe von 1,34m liegt.

Bei diesem Hörversuch wird die Differenz zwischen eingestelltem Azimutwinkel der virtu-

ellen Quelle und des angezeigten Winkels dargestellt. Dazu muss sich die Versuchsperson entweder im Sweetspot oder die Primärquelle auf einer der Flächen befinden. Ist das nicht der Fall, ist der Durchtrittspunkt und der damit errechnete Winkel mit dem eingestellten Winkel der Primärquelle nicht vergleichbar. Die Versuchspersonen befanden sich bei den Messungen im Sweetspot und die virtuellen Quellen wurden in einer Entfernung von $10m$ positioniert.

Die Messdaten wurden mit Boxplots ausgewertet

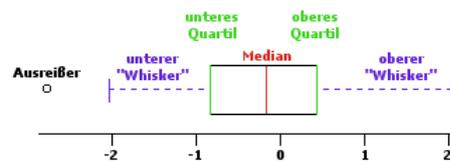


Abbildung 22: allgemeine Darstellung eines Boxplot

Die Box umfasst 50 Prozent der Messdaten. Innerhalb der Box ist der Medianwert markiert. Der Median ist das Messergebnis welches sich genau in der Mitte der sortierten Messdaten befindet, d.h. oberhalb und unterhalb des Medians befinden sich 50 Prozent der Messwerte. Durch den Median bekommt man einen Schiefen-Eindruck der Werte. Der Interquartilabstand ist die Länge der Box. In einem Whisker befinden sich, wenn keine Ausreißer vorkommen, 25 Prozent der Messwerte. Die maximale Länge des Whiskers beträgt das 1,5 fache des Interquartilabstandes. Die Messwerte die außerhalb dieser Länge liegen, werden als Ausreißer angezeigt.

4.2 Durchführung und Auswertung des Hörversuches

Vier Studienkollegen wurden als Versuchspersonen ausgewählt. Die Personen saßen während des Versuches auf einem Stuhl im Sweetspot, damit sich der Auslöser auf $1,34m$ befand. Es wurden 10 Primärquellen positioniert. 5 Quellen wurden für den Vergleich mit Ambisonics zwischen 40° und -40° , weitere 5 Quellen wurden siehe Abbildung 21, in der Horizontalebene positioniert. Der Versuch wurde für eine bessere Auswertung einmal wiederholt, ohne die Personen darüber aufzuklären.

Das in den folgenden Boxplots dargestellte Ergebnis ist die Differenz des Azimutwinkels der wahrgenommenen Primärquelle und der positionierten Primärquelle. Dieser Winkel ist in den Plots in Grad dargestellt. In den Plots 23 bis 26 ist der Absolutbetrag der Differenz der Absolutbeträge der Winkel dargestellt. In den Plots 27 und 28 ist die Differenz der Winkel dargestellt.

Für die richtige Interpretation der Absolutbeträge ist folgendes zu beachten. Je wei-

ter sich das Ergebnis von der 0° Abweichung entfernt, desto schlechter ist es. Für die Darstellung des Absolutbetrages werden die Boxplots in ihrer Form verändert und als Balkendiagramm dargestellt. Die vier Farben eines Balkens repräsentieren die vier Quartile des Boxplots. Innerhalb des schwarzen Bereiches befinden sich 25 Prozent der Messwerte mit der geringsten Abweichung (1.Quartil). Die Box ist nun der pinke und braune Bereich und enthält 50 Prozent der Messwerte (2. und 3.Quartil). Der Medianwert ist der Übergang zwischen pinken und braunen Bereich. Das 4.Quartil ist der weiße Bereich. Die wenigen Ausreißer werden nicht dargestellt.

Interpretation des Hörversuches

In Abbildung 23 sieht man die Absolutbeträge der Abweichungen aller 10 Primärquellen. Deutlich erkennbar ist, dass die Lokalisationsschärfe vor den Versuchspersonen deutlich zunimmt. Die Streuung der Ergebnisse nimmt ab. Das hat damit zu tun, dass das Richtungshören bei Menschen nach vorne besser ist als nach hinten.

Aus diesem Grund werden die Ergebnisse der Primärquellen, die zwischen 40° und -40° liegen, separat dargestellt (siehe Abbildung 24). Hier ist wiederum eine Zunahme, der Lokalisationsschärfe zum Zentrum hin, erkennbar. In den untersten Quartilen (also 25 Prozent der Messwerte) liegen die Abweichungen zwischen $0,1^\circ$ und $5,1^\circ$.

Abbildung 25 zeigt den Einfluss der einzelnen Versuchsperson auf das Endresultat. Liegen bei Person3 50 Prozent der Werte unterhalb von $4,8^\circ$ Abweichung, liegen bei Person1 50 Prozent der Werte unter $11,3^\circ$.

Abbildung 26 zeigt die Lokalisationsschärfe aller Messwerte. 25 Prozent der Messwerte liegen zwischen $0,1^\circ$ und $10,6^\circ$ Abweichung.

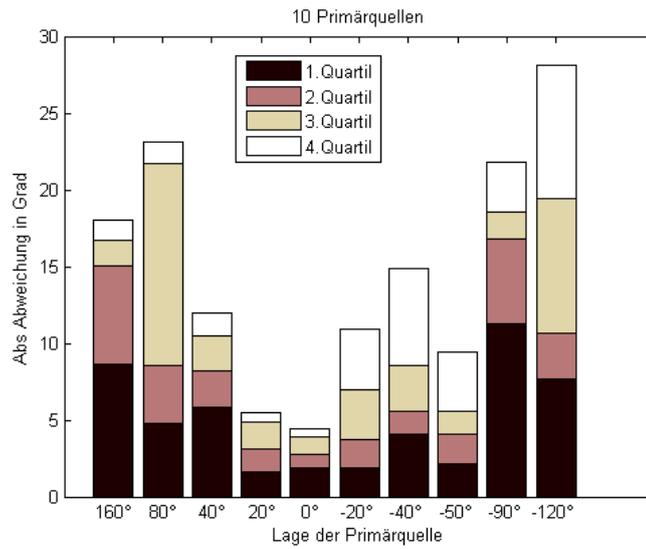


Abbildung 23: Absolutbetrag der Abweichung in Grad (10 Primärquellen)

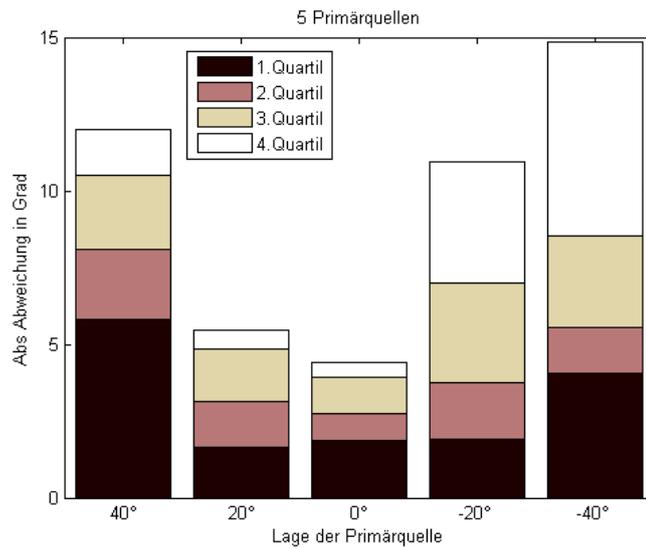


Abbildung 24: Absolutbetrag der Abweichung in Grad (5 Primärquellen)

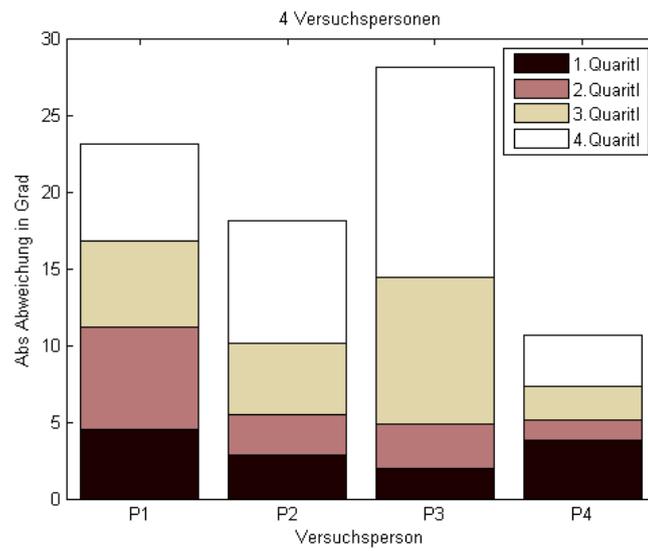


Abbildung 25: Absolutbetrag der Abweichung in Grad (4 Versuchspersonen)

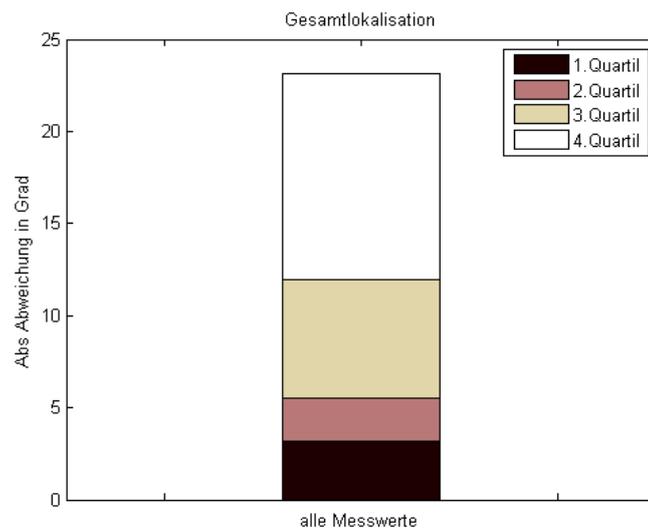


Abbildung 26: Absolutbetrag der Abweichung in Grad (aller Messwerte)

In den Abbildungen 27 und 28 ist die Abweichung ohne Absolutbetrag dargestellt. Deutlich zu erkennen ist, dass die Versuchspersonen die Primärquellen zentraler wahrnahmen als sie positioniert waren. D.h. die Versuchspersonen lokalisierten seitlich positionierte Quellen weiter vorne.

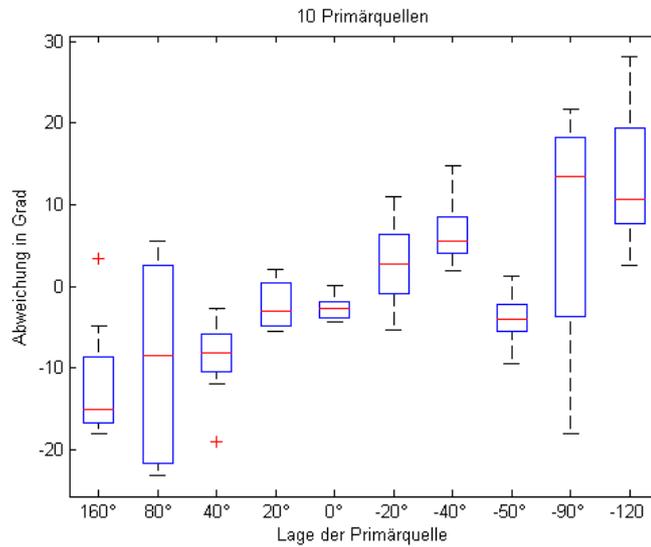


Abbildung 27: Abweichung in Grad (10 Primärquellen)

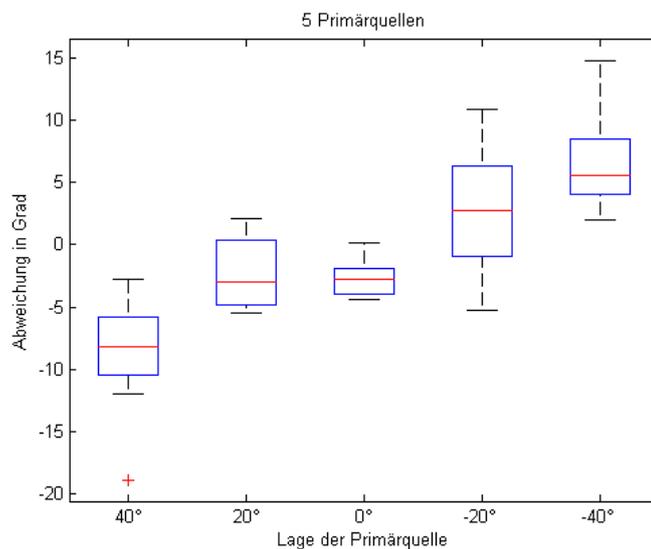


Abbildung 28: Abweichung in Grad (5 Primärquellen)

4.3 Ambisonics und Wellenfeldsynthese im Vergleich

Da bei Ambisonics viel mehr Messungen als bei der WFS durchgeführt wurden, ist dieser Vergleich nur eine Abschätzung. Zum Vergleich werden die 5 zwischen 40° und -40° positionierten Primärquellen herangezogen. Bei der Abbildung 30 stellt der grüne Bereich den unteren Whisker, der gelbe Bereich die Box bis zum Median und der rote Bereich die Box oberhalb des Medianes und den oberen Whisker da.

Beide Verfahren haben eine bessere Lokalisation in der 0° Richtung, was sich durch das bessere Richtungshören in dieser Richtung erklären lässt. Beide Grafiken zeigen ein ähnliches Bild. Insgesamt ist die Lokalisationsschärfe von Ambisonic 5. Ordnung ohne Delaykompensation besser, mit Delaykompensation etwas schlechter als bei der Wellenfeldsynthese.

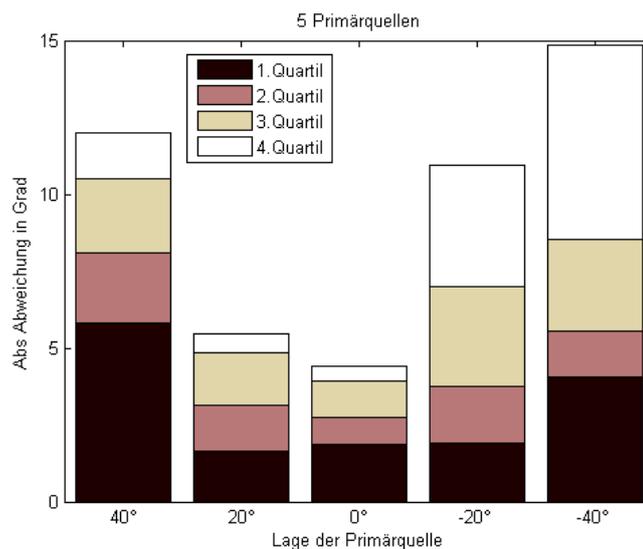


Abbildung 29: Absolutbetrag der Abweichung in Grad (Wellenfeldsynthese)

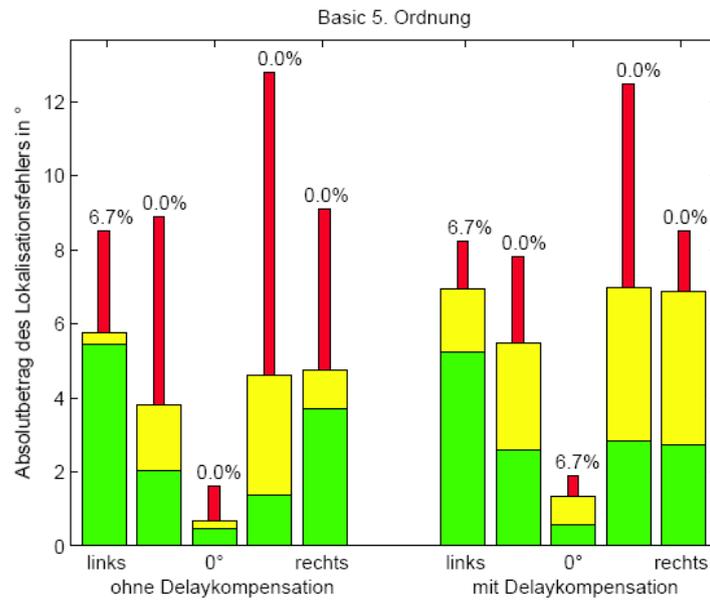


Abbildung 30: Absolutbetrag der Abweichung in Grad (Ambisonic 5.Ordnung) [Fra08]

Literatur

- [Bla74] J. Blauert, *Räumliches Hören*. Hirzel, 1974.
- [Fra08] M. Frank, "Hörversuche zur lokalisation bei verschiedenen ambisonics-dekodern," IEM, Tech. Rep., 2008.
- [Kei07] J. Keiner, "Tool for delaunay triangulations and voronoi tessellations on a sphere," 2007. [Online]. Available: <http://www.math.uni-luebeck.de/keiner/sonstiges.shtml>
- [OSB99] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-time Signal Processing*, 2nd ed. Prentice Hall, 1999.
- [SA08a] S. Spors and J. Ahrens, "A comparison of wave field synthesis and higher-order ambisonics with respect to physical properties and spatial sampling," in *125th AES Convention*, 2008.
- [SA08b] —, "Reproduction of moving virtual sound sources with special attention to the doppler effect," in *124th AES Convention*, 2008.
- [SA08c] —, "The theory of wave field synthesis revisited," in *124th AES Convention*, 2008.
- [Spo05] S. Spors, "Active listening room compensation for spatial sound reproduction systems," Ph.D. dissertation, Universität Erlangen, 2005.