

Soundfield Microphone (B-Format)

Bachelorarbeit aus Aufnahmetechnik 1, SE

Florian Iglisch
Jakob Spötl

Betreuung: Dr Alois Sontacchi
Graz, Juli 2011



institut für elektronische musik und akustik



Inhaltsverzeichnis

1. Einleitung	3
1.1. Aufgabenstellung	3
1.2. Theoretische Grundlagen	3
1.2.1. Ambisonics und das B-Format	3
1.2.2. Soundfieldmikrofon SPS200	4
1.2.3. Impulsantwort	6
1.2.4. Sweep	7
1.2.5. Richtungsinformation im B-Format	9
2. Messung	10
2.1. Rahmenbedingungen	10
2.2. Equipment	12
2.2.1. Trackingsystem	12
2.2.2. Messlautsprecher	13
2.2.3. Weitere Gerätschaften	14
2.3. Durchführung	15
2.3.1. Messaufbau	15
2.3.2. Automatisierung	16
2.3.3. Messpunkteverteilung	16
2.4. Störeinflüsse/Fehlerquellen	17
2.4.1. Reflexionen	17
2.4.2. Rauschen	18
2.4.3. Positionsfehlerquellen	19
2.4.4. Grobe Genauigkeitsabschätzung	21
2.4.5. Bias	22
3. Datenverarbeitung	22
3.1. Entfaltung	23
3.2. B-Format	25
3.3. Empfindlichkeit	26
3.4. Richtungsermittlung	26
3.5. Winkelabweichung	27
4. Datenanalyse	27
4.1. Entzerrung	27
4.2. Empfindlichkeit	29

4.2.1. Gesamtsignalbetrachtung	29
4.2.2. Einzelne Frequenzbereiche	30
4.3. Durchschnittliche Winkelabweichung	32
4.4. Winkelabweichung pro Position	33
4.4.1. Einzelne Frequenzbereiche	34
4.4.2. Drift	36
Literatur	38
A. Anhang	39
A.1. MATLAB-Funktionen	39
A.1.1. Datenerzeugung	39
A.1.2. Impulsantwort	40
A.1.3. Richtung und Richtungsfehler	41
A.1.4. Darstellung	44
A.1.5. Sonstige	45

1. Einleitung

1.1. Aufgabenstellung

Ziel war herauszufinden, wie genau die Aufnahme im B-Format mit dem Soundfield Mikrofon funktioniert. Dabei ging es primär um die räumliche Abbildung. Ferner sollte herausgefunden werden, ob die Genauigkeit abhängig von der Position und der Frequenz ist. Um das zu ermitteln war eine Theorieeinarbeitung, anschließend eine Vermessung des Mikrofons und schlussendlich eine Datenverarbeitung/-analyse notwendig.

1.2. Theoretische Grundlagen

1.2.1. Ambisonics und das B-Format

Das Ambisonics-Format basiert auf dem holophonen Ansatz und ist ein mehrkanaliges Aufnahme- und Wiedergabesystem [1]. Genauer gesagt ist es ein Spezialfall der Hologophonie indem das Schallfeld aus einer Überlagerung *ebener Wellen* angenommen wird. Damit diese Annahme erfüllt ist müssen folgende Randbedingungen getroffen werden:

1. Bei der Aufnahme des realen, bzw. Kodierung des synthetisierten Schallfeldes muss sich die Schallquelle weit genug entfernt von der Abhörposition befinden.
2. Bei der Wiedergabe über Lautsprecher muss die Entfernung zwischen Lautsprecher und Zuhörer groß genug sein.

Sind die Randbedingungen erfüllt, kann von ebener Wellenausbreitung gesprochen werden.

Für eine genaue mathematische Herleitung von Ambisonics soll an dieser Stelle auf die weiterführende Literatur [1, 3, 9] verwiesen werden. Das Ambisonics-Format basiert auf einer Bessel-Fourier-Reihenentwicklung, welche ab einer bestimmten Ordnung abgebrochen wird und der Ordnung des Ambisonics Systems entspricht. Nur für eine unendliche Ordnung ist die ideale Schallfeldreproduktion möglich. Daraus ergibt sich die Anzahl der Kanäle, welche im dreidimensionalen Fall quadratisch zunimmt:

$$N_{Kanal} = (M + 1)^2 \quad M \dots \text{Ordnung des Systems} \quad (1)$$

Aus der Reihenentwicklung ergeben sich Koeffizienten, die den Schall je nach Einfallsrichtung unterschiedlich gewichten. Je nach Ordnung entsprechen diese Koeffizienten

mehr oder weniger komplexen sphärischen Figuren, mit denen der Schall kodiert wird.

Während im synthetischen Fall der Kodierung aufgrund des reinen Berechnungsvorganges theoretisch keine Grenzen gesetzt sind, steht man bei der Kodierung des realen Schallfeldes schnell vor einem technischen Problem, da die Umsetzung mittels geeigneter Mikrofone erfolgt. Diese benötigen allerdings eine Richtcharakteristik, welche den Koeffizienten entspricht, damit der Schall in der richtigen Weise gewichtet wird. Für Ambisonics 0. Ordnung entspricht die benötigte Richtcharakteristik einer Kugel, welche mit einem Druckempfänger realisiert werden kann. Für Ambisonics 1. Ordnung benötigt man drei Schnellemikrofone, da die gewünschte Richtcharakteristiken jeweils einer Acht entsprechen. Für Ambisonics höherer Ordnung sind die Figuren allerdings komplexer und daher mittels Mikrofonen mit dem derzeitigen Stand der Technik nicht mehr zu realisieren.

Um den gesamten Frequenzbereich kodieren zu können, müssen alle Mikrofone im Koordinatenursprung übereinander liegen und unendlich klein sein. Da dem ebenfalls technische Grenzen gesetzt sind, muss neben der räumlichen Ungenauigkeit der Schallfeldreproduktion auch eine Einschränkung im oberen Frequenzbereich in Kauf genommen werden.

Für Ambisonics 1. Ordnung, auch als *B-Format* bezeichnet, ergeben sich vier Kanäle, wobei der w -Kanal den Einfluss des Schalldrucks p beschreibt und die restlichen drei Kanäle (x,y,z) die drei Komponenten der Schallschnelle v in kartesischen Koordinaten [1]. Die folgende Formel gibt eine Übersicht der Faktoren, mit denen p und v in diesem Fall gewichtet werden müssen ($\varphi \dots$ Azimut, $\vartheta \dots$ Elevation):

$$\begin{aligned} w &= \frac{1}{\sqrt{2}} && 0. \text{ Ordnung} \\ x &= \cos(\varphi) \cos(\vartheta) \\ y &= \sin(\varphi) \cos(\vartheta) && 1. \text{ Ordnung} \\ z &= \sin(\vartheta) \end{aligned} \tag{2}$$

1.2.2. Soundfieldmikrofon SPS200¹

Das Soundfieldmikrofon wurde für das B-Format entwickelt und besteht aus vier Mikrophonkapseln mit Nierencharakteristik, welche in Form eines Tetraeders angeordnet sind [2]. Jeder Membranmittelpunkt entspricht dabei genau einer Spitze des geometrischen

¹SPS200 Software Controlled Microphone, SoundField Ltd., Wakefield, West Yorkshire, England; User Guide: <http://www.soundfield.com/downloads/sps200-manual.pdf>, 2011

Körpers. Die Ausrichtung der Kapseln erfolgt jeweils vom Mittelpunkt des Tetraeders zu der entsprechenden Spitze. Aufgenommen werden die Kanäle getrennt, im sogenannten A-Format, welche über eine XLR-Steckverbindung mit einer entsprechenden Soundkarte oder einem Mischpult verbunden werden können.



Abbildung 1: Soundfieldmikrofon SPS200

Durch die geometrischen Eigenschaften eines Tetraeders lassen sich die Daten leicht in die vier Kanäle w, x, y, z des B-Formates umwandeln. Am leichtesten lässt sich das veranschaulichen, wenn man sich den Tetraeder in einem Würfel denkt, wobei die Kanten des Tetraeders den Diagonalen der Würfelflächen entsprechen. Die Flächennormalen des Würfels entsprechen dabei jeweils einer Achse eines kartesischen Koordinatensystems (siehe Abbildung 2).

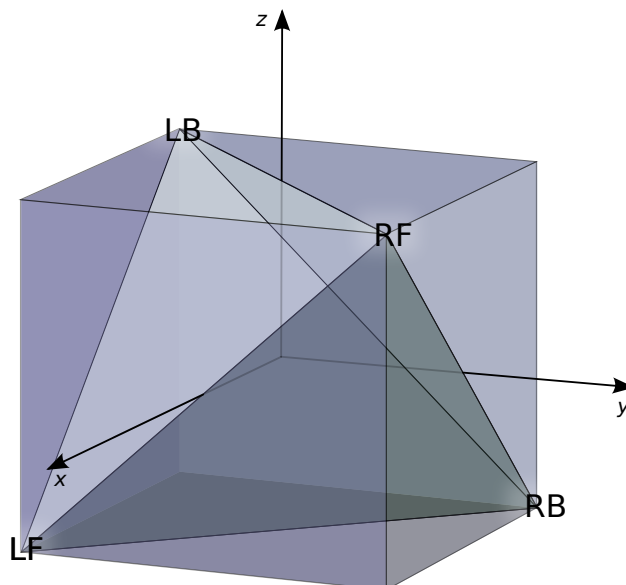


Abbildung 2: Geometrisches Modell des Mikrofons: das Mikrofon wird durch das Tetraeder symbolisiert, wobei die Mikrofonkapseln jeweils eine Spitze des Tetraeders darstellen. Der Würfel dient zur besseren Vorstellung der Lage des Koordinatensystems

Um nun das B-Format zu erzeugen, müssen die einzelnen Kanäle nur noch in entsprechender Weise addiert bzw. subtrahiert werden. Dabei erfolgt die Kanalbenennung nach folgendem Schema:

LF ... *LeftFront*
RF ... *RightFront*
LB ... *LeftBack*
RB ... *RightBack*

Für das B-Format gilt:

$$\begin{aligned} w &= LF + RF + LB + RB \\ x &= LF + RF - (LB + RB) \\ y &= RF + RB - (LF + LB) \\ z &= RF + LB - (LF + RB) \end{aligned} \tag{3}$$

Die linke vordere und die rechte hintere Kapsel schauen dabei jeweils leicht nach unten, die Anderen entsprechend nach oben.

1.2.3. Impulsantwort

Jedes System kann mathematisch durch seine systemeigene Impulsantwort beschrieben werden. Der Zusammenhang sieht wie folgt aus:

$$u'(t) = u(t) * h(t) \tag{4}$$

Wenn ein Signal $u(t)$ durch ein beliebiges System geschickt wird, so wird es mit der Systemimpulsantwort $h(t)$ gefaltet. Als Ergebnis ergibt sich die Antwort des Systems auf dieses Eingangssignal $u'(t)$.

Will man nun die Impulsantwort eines Systems ermitteln, muss eine Entfaltung durchgeführt werden. In den Fourier-Bereich transformiert sieht Gleichung 4 folgendermaßen aus:

$$U'(f) = U(f) \cdot H(f) \tag{5}$$

Wird das gemessene Spektrum $U'(f)$ durch das Spektrum der Ursprungsfunktion $U(f)$ dividiert, so erhält man die Übertragungsfunktion $H(f)$ (6), welche nur noch in den Zeitbereich zurück transformiert werden muss, um die systembeschreibende Impulsantwort

in Abhängigkeit von der Zeit zu erhalten.

$$H(f) = \frac{U'(f)}{U(f)} \quad (6)$$

1.2.4. Sweep

Der Begriff Sweep steht als Abkürzung für Sinussweep und bezeichnet ein in der Frequenz veränderliches ($\omega = \omega(t)$) Sinussignal, welches mit einer Anfangsfrequenz $\omega_1 = \omega(0)$ beginnt und zu einer Endfrequenz $\omega_2 = \omega(T)$ (T = Sweepdauer) verändert wird.

Charakteristisch für unterschiedliche Sweeps sind Startfrequenz, Endfrequenz, Dauer und auch die Art mit der der Frequenzbereich durchquert wird. Üblicherweise gibt es zwei Möglichkeiten:

- Lineare Sweeps, dabei ist die Frequenz linear abhängig von der Zeit.

$$\omega(t) = k \cdot t + d \quad \text{mit } d = \omega_1 \text{ und } k = \frac{(\omega_2 - \omega_1)}{T} \quad (7)$$

- Exponentielle Sweeps, dabei ändert sich die Frequenz exponentiell mit der Zeit.

$$\omega(t) = a \cdot e^{\left(\frac{t}{\tau}\right)} \quad \text{mit } a = \omega_1 \text{ und } \tau = \frac{T}{\ln\left(\frac{\omega_2}{\omega_1}\right)} \quad (8)$$

Die Vorteile der jeweiligen Methoden liegen auf der Hand. Lineare Sweeps sind mathematisch einfacher zu beschreiben. Hingegen entsprechen exponentielle Sweeps viel besser unserem natürlichen Hören. Während bei linearen Sweeps tiefere Frequenzen verhältnismäßig viel zu schnell durchschritten werden (weniger Anregungsenergie), bleibt die Zeitspanne pro Oktave beim exponentiellen Ansatz über den gesamten Frequenzbereich gleich.

Wichtige Vorteile von exponentiellen Sinussweeps als Messsignale gegenüber anderen Anregungsarten sind:

- Guter SNR², da das System lange Zeit einem konstant hohen Eingangspegel ausgesetzt ist (vor allem im Gegensatz zur Impulsanregung) .
- Harmonische Verzerrungen können sehr gut vom eigentlichen Signal getrennt werden.
- Leicht anzupassen an den relevanten Frequenzbereich welcher damit besser aus-

²Sound Noise Ratio: Verhältnis von Signalleistung zu Rauschleistung

genutzt werden kann.

- Frequenzauflösung theoretisch unendlich. Diese wird nur noch von der Fouriertransformation beschränkt (vor allem im Gegensatz zur Messung mit reinen Sinustönen). Die Abhängigkeitsparameter sind dabei die Messdauer und die Abtastfrequenz.

Die harmonischen Verzerrungen werden bei Messungen mit exponentiellen Sweeps nicht wie bei anderen als Teil der Anregung interpretiert (vgl. z.B. MLS³), sondern aufgrund des konstanten relativen Frequenzabstands (z.B. bei der ersten Oberschwingung, eine Oktave/doppelte Frequenz), als "separate", zeitlich versetzte Anregung.

Zur Veranschaulichung sei auf Abbildung 3 verwiesen, welche das Ergebnis einer Sweepmessung in Zeit-/Frequenzabhängigkeit exemplarisch darstellt.

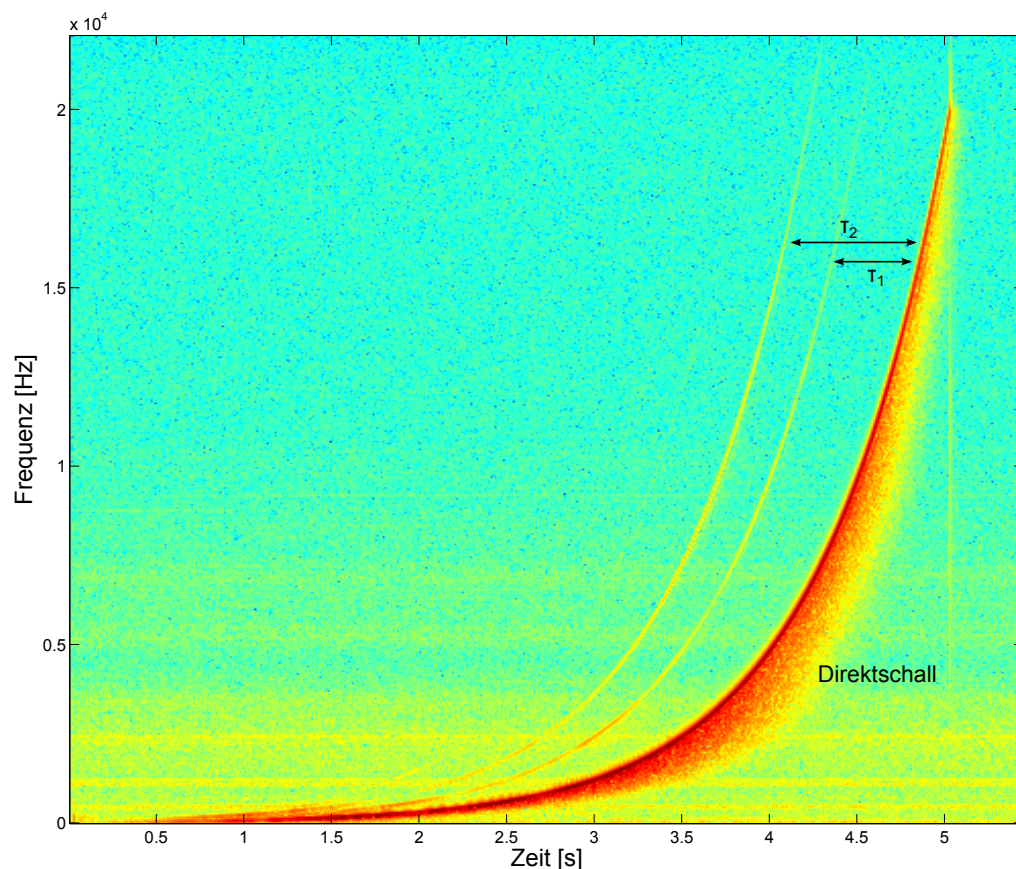


Abbildung 3: Spektrum einer Messung mittels Exponentialsweep: die durch die harmonischen Verzerrungen entstandenen zusätzlichen Sweeps besitzen den zeitlichen Abstand τ_1 und τ_2 zum Direktschall

³Maximum Length Sequence: Verfahren zur Ermittlung der Impulsantwort mittels einer Pseudozufallsfolge

Das mit der Impulsantwort gefaltete Signal besteht nun aus drei Sweeps, welche zeitlich parallel durch den oben genannten Frequenzabstand voneinander getrennt sind. Die zwei zuerst auftretenden Sweeps entstehen dabei durch die harmonischen Verzerrungen wodurch unterschiedliche Startpunkte der Impulsantworten für die jeweiligen Sweeps entstehen. Dies ermöglicht wiederum eine verbesserte Trennung der Verzerrungen vom eigentlichen Signal. Bei Bedarf kann man auch die Abstände vergrößern um die jeweiligen Frequenzgänge besser betrachten zu können. Der entstehende zeitliche Abstand lässt sich folgendermaßen berechnen:

$$\tau_i = \tau \cdot \ln(i) \quad \text{mit} \quad \tau = \frac{T}{\ln(\frac{\omega_2}{\omega_1})} \quad (9)$$

[7], siehe oben.

Dabei ist i die Ordnung der Harmonischen, $i = 1$ entspricht der Grundschwingung, dem "Grundsweep". Die Abstände werden mit zunehmendem Grad immer geringer.

Harmonische Verzerrungen treten bei Sweepmessungen grundsätzlich im negativen Zeitraum auf, wenn für den Impuls des Direktschalls $t = 0$ gilt. Dadurch lassen sie sich leicht von Reflexionen unterscheiden und von der eigentlichen Impulsantwort trennen. Für eine genauere Behandlung des Themas sei der Leser auf die Arbeit von *Farina A.* [6] und *Majdak P. et al* [7] verwiesen.

1.2.5. Richtungsinformation im B-Format

Ausgehend von einem Signal im B-Format kann theoretisch die exakte Richtung der Schallquelle bestimmt werden, jedoch ohne Entfernungsinformation. Dies gilt allerdings nur wenn diese Richtung die einzige ist, aus welcher Schall auf das Mikrofon trifft. Bei Überlagerung mehrerer Schallquellen ergibt sich die Richtung mit der höchsten Summenschnelle. Die Richtung kann mit Hilfe der Intensität folgendermaßen berechnet werden ($p \dots$ Schalldruck, $\mathbf{v} \dots$ Schallschnelle):

$$\mathbf{I} = p \cdot \mathbf{v}^* \quad (10)$$

Auch im B-Format lässt sich dieses Verhältnis anwenden, da die einzelnen Kanäle in direkter Beziehung zu Schalldruck und -schnelle stehen (s. 1.2.1). \mathbf{I}_R beschreibt dabei die Richtung, aus der das Signal kommt. Da die Richtungsinformation rein reell ist, muss

der Realteil als Ergebnis genommen werden:

$$\mathbf{I_R} = \text{Re} \left\{ w \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}^* \right\} \quad (11)$$

Der sich dadurch ergebende Abstand zum Mikrofon hängt rein von der Signalstärke ab und enthält damit keine allgemein gültige Entfernungsinformation. Eine Umrechnung in Polarkoordinaten erscheint hier für die weiteren Berechnungen allgemein sinnvoll, da ϑ und φ unabhängig vom Radius betrachtet werden können.

Für weiterführende Literatur und das DirAC-Verfahren⁴, welches mit diesem Ansatz arbeitet, sei auf [4, 5] verwiesen.

2. Messung

Für die Datenerzeugung wurde eine Reihe von vier exponentiellen Sweeps generiert, wobei jeder Sweep eine Länge von 5 s besitzt und die Abstände zwischen den Sweeps 2 s betragen. Die Grenzfrequenzen bilden mit $\omega_u = 20$ Hz und $\omega_o = 20$ kHz den gesamten Hörbereich ab. Die gemessenen Daten wurden mit einer Samplerate von 44,1 kHz und einer Bittiefe von 16 Bit aufgezeichnet.

Im Folgenden sollen die Rahmenbedingungen kurz erörtert und das verwendete Equipment beschrieben werden. Ebenso wird ein Überblick über die Durchführung der Messung gegeben, sowie eine grobe Fehlerabschätzung durchgeführt, mit der die Genauigkeit der Messung eingegrenzt und die Daten gegebenenfalls korrigiert werden können.

2.1. Rahmenbedingungen

Um die Messpunktdichte möglichst hoch halten zu können, ohne dass die Messdauer überhand nimmt, oder dass die Anzahl der Messungen auf Kosten mehrfacher Messungen pro Punkt geht, gilt es zu überlegen, ob die Symmetrie des Mikrofons zu einer kleineren Messfläche führt.

Ein Tetraeder besteht aus vier gleich großen, gleichseitigen Dreiecken, weshalb sich im Idealfall für das Soundfieldmikrofon die Bedingungen über jeder gedachten Dreiecksfläche zwischen den Mikrofonkapseln wiederholen. Mit dieser Überlegung lässt sich die

⁴Directional Audio Coding

Messfläche von einer vollständigen Kugeloberfläche auf ein gutes Viertel reduzieren. In unserem Fall wurde die Messung im folgenden Winkelbereich realisiert: $42^\circ \leq \varphi \leq 162^\circ$, $30^\circ \leq \vartheta \leq 150^\circ$, wobei besonders bei ϑ die äußeren Gegebenheiten die Grenzen beeinflussten. Die Fläche über den gemessenen Bereich ergibt:

$$A = \iint_{\varphi \vartheta} 1 \cdot r^2 \sin(\vartheta) d\vartheta d\varphi = \frac{2\pi}{\sqrt{3}} r^2 \quad (12)$$

Bezieht man diese Fläche auf die Gesamtoberfläche einer Kugel ($A_{ges} = 4\pi r^2$), so erhält man:

$$\frac{A}{A_{ges}} = 28,87\% \quad (13)$$

Es wird also gut ein Viertel der Kugeloberfläche vermessen. In Abbildung 4 sind die geplanten Messpunkte (braun) in Bezug zu dem Tetraeder über das Mikrofon dargestellt. Die schwarzen/grauen Punkte symbolisieren die Mikrofonkapseln (Bezeichnung siehe 1.2.2). Für weitere Abbildungen siehe auch Abbildung 10.

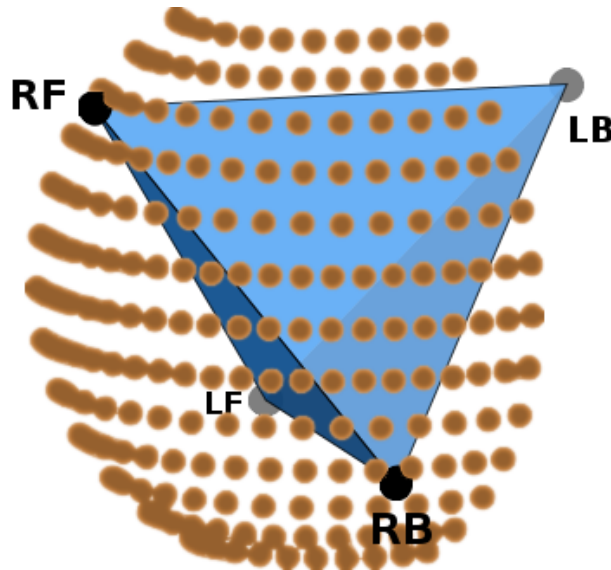


Abbildung 4: Die braunen Punkte symbolisieren die einzelnen Positionen, an denen eine Messung durchgeführt wurde. Das Tetraeder, dessen Spitzen durch die schwarzen/grauen Punkte die Mikrofonkapseln symbolisieren, stellt den Bezug zur Lage des Mikrofons gegenüber den Messpunkten her.

2.2. Equipment

2.2.1. Trackingsystem

Um die Positionen möglichst genau verifizieren zu können, sollte ein Trackingsystem zu Hilfe genommen werden. Da das im Cube fix installierte System zu der Zeit der Messung nicht zur Verfügung stand, musste auf ein mobiles Trackingsystem (OptiTrack⁵) ausgewichen werden. Ausgestattet mit sechs Kameras, einem Switch und einem Laptop auf welchem die Software (*Tracking Tools*⁶) lief, konnte man die Position markierter Gegenstände in Echtzeit beobachten.

Die Kalibrierung des Systems erfolgte durch Bewegung von vier Kugeln im Raum, der von den Kameras eingeschlossen wurde. Alle Kugeln lagen in einem der Software bekannten konstanten und unregelmäßigen Abstand auf einer Linie. Je nach Verteilung der aufgezeichneten Punkte (abhängig von der Dauer der Kalibrierung) konnte die Software die Abstände der Kameras zueinander und damit den eingeschlossenen Raum bis zu einer bestimmten Genauigkeit bestimmen.

Mittels kleiner Kügelchen, welche auf die zu verfolgenden Objekte befestigt wurden, konnte man selbige als ein zusammen gehöriges Objekt definieren und dessen aktuelle Position ablesen. Die Schwierigkeit dabei war, dass jedes Objekt auf einen Punkt reduziert wurde. Die Position dieses "Summenpunktes" ist automatisch in der Mitte aller Punkte eines Objektes, welche für eine sinnvolle Winkelbestimmung zwischen dem Lautsprecher und dem Mikrofon nicht zu gebrauchen war. Der Summenpunkt kann zwar verschoben werden, jedoch gestaltete sich auch das eher schwierig, da die Orientierung anhand von wenigen am Bildschirm dargestellten Punkten nicht sehr einfach ist. Abhilfe schafften temporäre Trackingpunkte, welche für diesen Einstellvorgang auf die Membranen geklebt wurden. Beim Lautsprecher musste nun einfach der Summenpunkt auf den Membranmittelpunkt (nun auch optisch am Bildschirm erkennbar) verschoben werden. Beim Mikrofon wurden die Kapseln als Punkte definiert und zu einem Objekt zusammen gefasst. Der Mittelpunkt dieses Objektes lag nun exakt zwischen den Kapseln (= Mikrofonmittelpunkt), hier wurde der vorherige Summenpunkt auf den neuen/temporären verschoben. Nach dieser Justierung wurden diese Tracking-Kügelchen für die Messung wieder entfernt.

Das Trackingsystem selbst kann nur absolute Koordinaten ausgeben, allerdings können

⁵OptiTrack, NaturalPoint Inc./OptiTrack, Oregon, USA

<http://www.naturalpoint.com/optitrack/>

⁶Tracking Tools, NaturalPoint Inc./OptiTrack, Oregon, USA

<http://www.naturalpoint.com/optitrack/products/tracking-tools/>

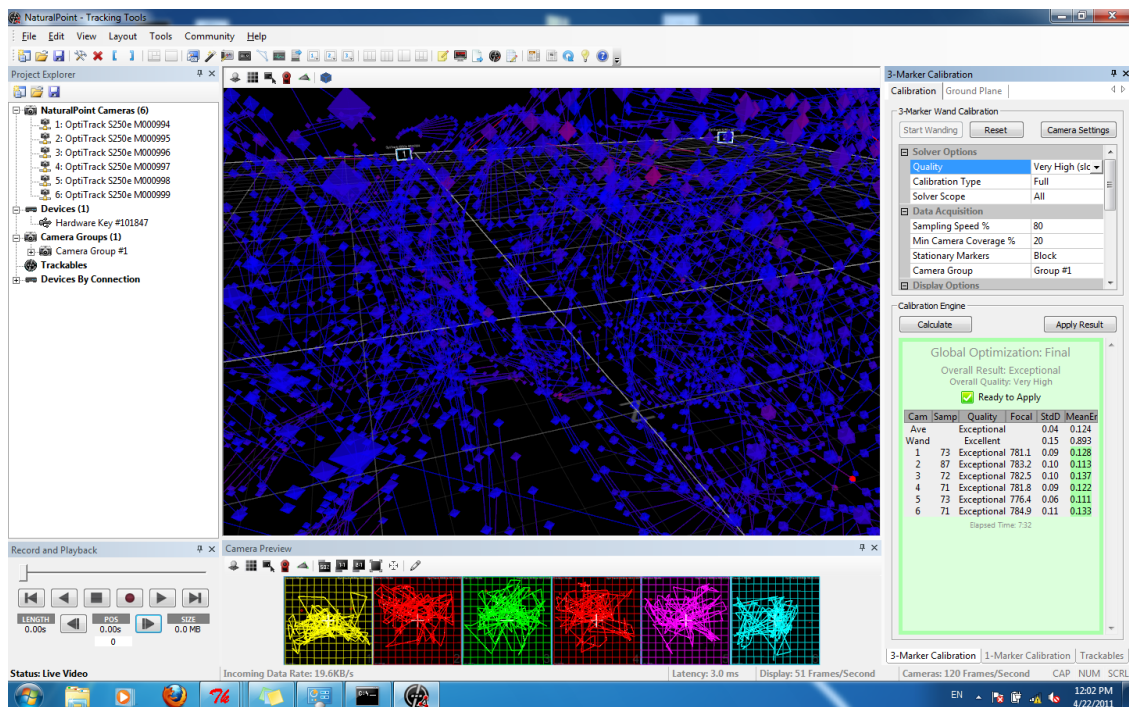


Abbildung 5: Trackingsystem Kalibrierung: die Karos in dem großen Fenster markieren die aufgenommenen Punkte, anhand derer die Software die Position der Kameras und der Raum zwischen ihnen bestimmen konnte. Der grün markierte Bereich (rechts unten) gibt die erreichte Genauigkeit wieder.

diese Daten an PD⁷ weitergegeben werden, sodass der Koordinatenstream in Echtzeit in die Relativposition zwischen Mikrofon und Lautsprecher (Polarkoordinaten) umgerechnet und ausgegeben werden kann.

2.2.2. Messlautsprecher

Um ein möglichst über den gesamten Frequenzbereich konstantes akustisches Zentrum gewährleisten zu können, ist ein Einweglautsprecher (oder ein Mehrweglautsprecher mit konzentrischen Membranen) notwendig, wobei ein Kompromiss bei der Abstrahlung in bestimmten Frequenzbereichen (in unserem Fall bei tiefen Frequenzen) eingegangen wird. Ein linearer Frequenzgang ist hingegen vernachlässigbar, da sich die Richtungsinformation (um welche es sich in dieser Arbeit primär handelt) rein durch Relativverhältnisse zwischen den Kanälen ergibt. Für derartige Messungen werden am IEM⁸ unter anderem Tropfenlautsprecher verwendet. Ein Vorteil ist auch das geringe Gewicht und die unkomplizierte Befestigung an jedem (stabileren) Mikrofonstativ. Abbildung 6 zeigt den Frequenzgang des verwendeten Lautsprechers.

⁷Pure Data: grafische Echtzeitprogrammierungssprache für Audio-, Video- und Grafikanwendungen

⁸Institut für Elektronische Musik, Kunstuniversität Graz

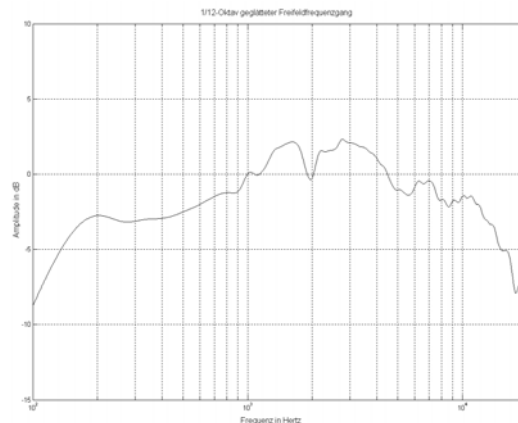


Abbildung 6: Frequenzgang des Tropfenlautsprechers [8]

2.2.3. Weitere Gerätschaften

- *Soundfield Mikrofon SPS200*: das zu vermessende Mikrofon.
- *Soundfield SMP200*⁹: Preamplifier, welcher sicherstellt, dass alle Mikrofonsignale gleich verstärkt werden.¹⁰
- *Motu 986HD*: Audio Interface (eigentlich hätte dafür das RME Fireface 400 verwendet werden sollen, jedoch ergaben sich hierbei softwaretechnische Probleme mit dem zur Verfügung stehenden Aufnahmerechner, so dass auf das angeführte Gerät ausgewichen werden musste)
- *Macbook Pro (2008)*: Aufnahmerechner (da der uns zur Verfügung gestellte Laptop nur für das Trackingsystem verwendet werden konnte, mussten wir auch hier ausweichen)
- *Macbook Pro (2011)* (dieser Rechner war für das Trackingsystem zuständig, mit Hilfe eines einfachen PD-Patches konnten wir in Echtzeit die Relativposition zwischen Lautsprecher und Mikrofon überprüfen)
- *Drehteller, Outline ET250-3D*: Dies ist eine motorgesteuerter Drehvorrichtung, welche sowohl von Hand direkt bedient werden kann, als auch via Netzwerkverbindung (in unserem Fall, wurde der Drehteller von einem PD-Patch halbautomatisch gesteuert. Die kleinsten Schrittgrößen waren 0.5° was bei der Messpunktwahl berücksichtigt werden musste).

⁹SMP 200 Gain Matched Mic-Pre, SoundField Ltd., Wakefield, West Yorkshire, England, User Guide unter <http://www.soundfield.com/downloads/smp200.pdf>, 2011

¹⁰Bei dem zur Messung verwendeten Mikrofon stellte sich bei Überprüfung heraus, dass die Beschriftung der Kabel nicht mit den Kanälen übereinstimmt



Abbildung 7: Mikrophon und Lautsprecher beim Kalibrieren, Kamera (kleines Bild), Aufbau (v.l.n.r.)

2.3. Durchführung

2.3.1. Messaufbau

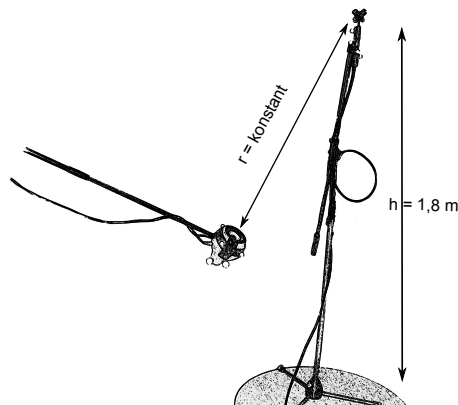


Abbildung 8: Messaufbau: Das Mikrophon wurde in einer Höhe von 1,8 m zum Boden auf einem Drehteller positioniert. Der Lautsprecher wurde je nach ϑ mit konstantem Abstand und Blickrichtung zum Mikrophon platziert.

Die Vermessung des Soundfield Mikrofons erfolgte im Cube des IEM. Das Mikrophon wurde ca. in der Mitte des Raumes auf dem Drehteller positioniert und konzentrisch zu diesem ausgerichtet. Die Orientierung war standardmäßig, mit XLR-Anschluss nach unten. Die Höhe des Mikrophonzentrums war ca. 1,8 m. Der Lautsprecher wurde für jede Messreihe in Höhe und Neigungswinkel (Membranflächennormale in Richtung Mikrophon) eingestellt. Sein Abstand zum Mikrophon sollte allerdings möglichst konstant bleiben wie auch seine horizontale Richtung. Mit Hilfe des Trackingsystems konnte der Lautsprecher auf $\varphi, \vartheta = \pm 0,4^\circ$ und $r = \pm 0,02 \text{ m}$ genau eingestellt werden. Diese Ungenauigkeit wurde mitprotokolliert und bei der Datenverarbeitung mitberücksichtigt.

2.3.2. Automatisierung

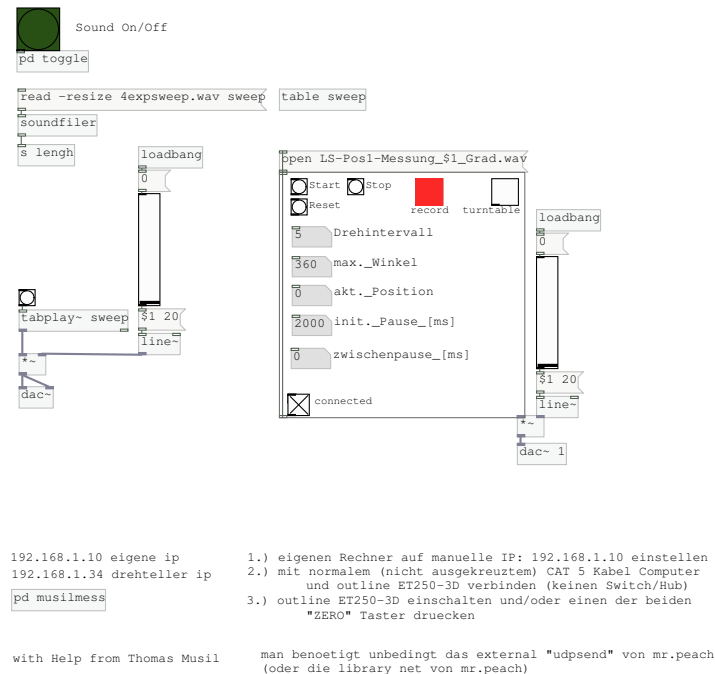


Abbildung 9: Aufnahme patch.pd, PD-Oberfläche für die automatisierte Messung

Mit Hilfe des Drehtellers und eines PD-Patches (Abbildung 9) war es möglich, die Messungen, bei denen der Lautsprecher in der Position nicht verändert werden musste, zu automatisieren. Eingabeparameter des PD-Patches sind der Zielwinkel und die Schrittweite jeweils in Grad. Die Audiodaten werden automatisch mit der Lautsprecherposition und dem aktuellen Winkel benannt und im Wave-Format abgespeichert. Für eine Wiederholung der Messung werden die Daten einfach überschrieben.

2.3.3. Messpunkteverteilung

Um eine möglichst konstante Messpunktverteilung zu erreichen wurde der Horizontalwinkel/Azimut durch oben erwähnten Drehteller verändert. So war es möglich einzelne Bahnen auf einer Kreisoberfläche möglichst genau mit Messpunkten zu versehen. Der Lautsprecher wurde für jede Bahn aufs Neue positioniert und ausgerichtet. Der Abstand dieser Bahnen war konstant 10° Vertikalwinkel/Elongation bei gleichem Radius. Die Anzahl der Messpunkte auf den einzelnen Bahnen variierte, bei 90° (horizontal) wurden am meisten Punkte vermessen und zu den Rändern jeweils symmetrisch weniger. Gemessen wurde über einen 120° × 120° Ausschnitt, von 30° zur Vertikalen/Boden bis 30° Vertikal/Decke \implies Azimut: $42^\circ \leq \varphi \leq 162^\circ$, Elevation: $-60^\circ \leq \vartheta \leq 60^\circ$.

Zusätzlich zu den normalen Punkten im zu vermessenden Bereich wurden hin und wieder weitere Positionen vermessen (zusätzliche Messpunkte). Dabei handelte es sich um jeweils 3 Punkte pro Bahn mit denen die restlichen 240° abgedeckt wurden (Abbildung 10).

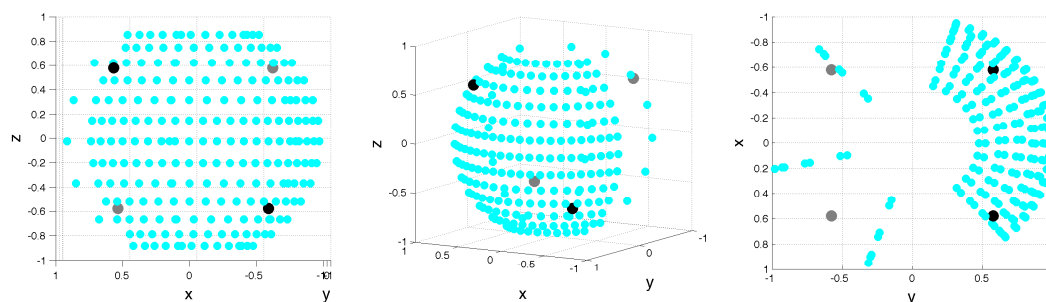


Abbildung 10: Darstellung der aus dem Trackingsystem herausgelesenen Messpunkte. Die zusätzlichen Messpunkte mit großem φ -Abstand zueinander wurden aufgrund der hohen Aussagekraft der eigentlichen Messung nicht weiter beachtet.

Die zusätzlich gemessenen Messpunkte wurden, auf Grund der hohen Aussagekraft der eigentlichen Messpunkte, nicht für die Analyse benötigt. Daher wurde auf eine weitere Betrachtung verzichtet.

ϑ	#Messpunkte	$\Delta\varphi$	zusätzliche Messpunkte
± 60	10	12	x
± 50	12	10	
± 40	12	10	x
± 30	15	8	
± 20	15	8	x
± 10	20	6	
0	20	6	x

Gestartet wurde jeweils bei $\varphi = 42^\circ$ und geendet bei $\varphi = 162^\circ$. Für eine externe Geräuschbewertung wurden zusätzlich vier Messungen ohne Raumbeschallung durchgeführt.

2.4. Störeinflüsse/Fehlerquellen

2.4.1. Reflexionen

Um die Reflexionen möglichst schwach und zeitlich weit entfernt vom Nutzsignal zu halten, wurde das Mikrophon auf einer Höhe von ca. 1,8 m befestigt. Dadurch war

es dann allerdings auch nicht mehr möglich Lautsprecherpositionen höher als die oben angegebene ($\vartheta = 60^\circ$) zu realisieren.

Um den Einfluss der Bodenreflexion als stärkste und v.a. zeitlich früheste Reflexion möglichst gering zu halten, wurde Molton (dicker Stoff) verwendet, welcher, über Stühle gehängt, eine brauchbare Dämpfung ergab.

In Folge der Auswertung der Ergebnisse stellte ein Analysezeitfenster von 100 Samples den besten Kompromiss zwischen Frequenzauflösung und Vermeidung störender Reflexionen dar.

2.4.2. Rauschen

Die größten Störquellen während der Messung waren zum einen die Kameras, welche für das Trackingsystem verwendet wurden und zum anderen der dazu benötigte Laptop, welcher aus Arbeitskomfortgründen im Aufnahmerraum stand (der Direktschall des Laptop-Geräusches wurde mit mobilen Absorberwänden minimiert).

Für das Verhältnis zwischen Nutzsignal und Rauschen wird normalerweise eine klassische SNR Angabe verwendet, in diesem Fall ist der SNR sehr stark frequenzabhängig (in erster Linie durch den Frequenzgang des Lautsprechers), weshalb eine Angabe hier nicht sinnvoll erscheint. Interessant ist dagegen die in Abbildung 11 (links) dargestellte Gegenüberstellung des Energieverlaufs des Signals (w -Kanal) und des Rauschens über die Frequenz. Man kann erkennen, dass der Frequenzgang des Lautsprechers erst ab ca. 200 Hz annähernd linear wird. Zu tieferen Frequenzen sinkt der Pegel drastisch, was in Abbildung 11 (rechts) noch deutlicher zu sehen ist. Um einen ausreichenden SNR gewährleisten zu können, dürfen Frequenzen unterhalb von min. 50 Hz eigentlich nicht betrachtet werden. Darüber jedoch wird ein konstant sehr hoher SNR von über 60 dB gehalten.

Interessant dabei ist bei Betrachtung des Rauschens vor allem das Maximum bei 471,5 Hz, welches ziemlich sicher von den Kameras verursacht wurde. Dieser Ton ($\hat{=}$ b') ist aus dem Rauschspektrum heraus durchaus hörbar.

Alles in allem kann man das Rauschen für die Messung als irrelevant bezeichnen. Lediglich bei tiefen Frequenzen sinkt der SNR auf immer noch über 10 dB. Dort wächst zum einen die Leistung des Rauschens, zum anderen sinkt die Abstrahlleistung des Lautsprechers (Fehlanpassung).

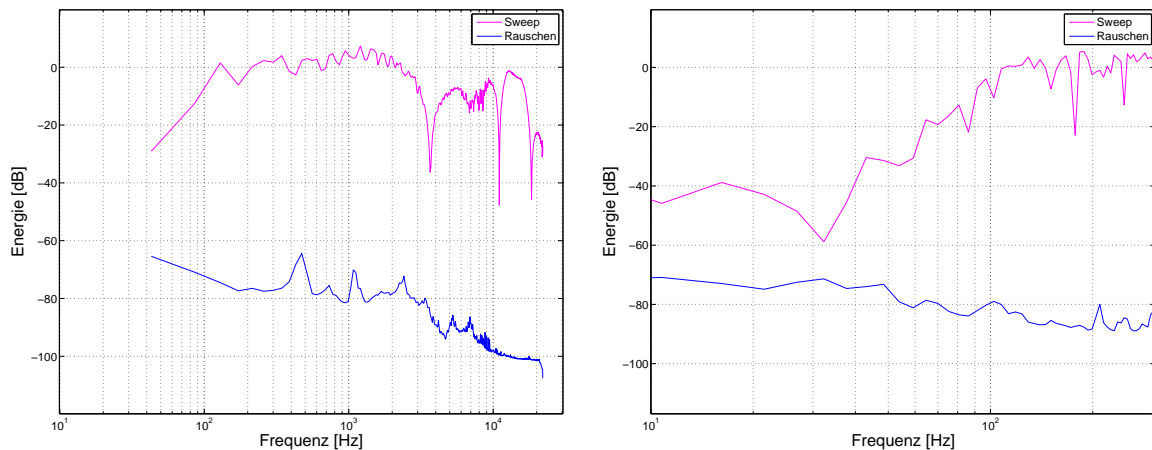


Abbildung 11: Gegenüberstellung der Signal- zur Rauschleistung über die Frequenz. Links: gesamter Frequenzbereich (1024 FFT), rechts: Ausschnitt bis 200 Hz (8192 FFT). Gut zu sehen ist der erhöhte Rauschanteil im unteren Frequenzbereich (bis ca. 100 Hz) und der starke Peak bei 471,5 Hz. Ebenso deutlich sichtbar sind die schlechte Anpassung des Lautsprechers (besonders um 30 Hz) bis etwa 200 Hz und die starken Auslöschungen im oberen Frequenzbereich

2.4.3. Positionsfehlerquellen

Im Folgenden werden mögliche Fehler-, beziehungsweise Ungenauigkeitsquellen aufgelistet.

- Die Ungenauigkeit des Trackingsystems gilt es sowohl beim Lautsprecher, als auch beim Mikrofon zu berücksichtigen. Die Messunsicherheit des Trackingsystems beläuft sich auf 0,875 mm. Die generelle Messunsicherheit lässt sich aus den Ergebnissen der Kalibration herauslesen (Abbildung 12, Overall Wand Error).
- Die Positionierung des Mikrofons auf dem Drehteller (damit verbunden die Konzentrität) wurde teils optisch (Lot) und teils mit Hilfe des Trackingsystems durchgeführt und überprüft. Maximale Fehler beliefen sich auf $\pm 0,1^\circ$ für φ und ϑ , sowie ± 1 mm für den Radius pro Umdrehung.
- Die horizontale Ausrichtung des Mikrofons konnte leider nur optisch erfolgen und ist damit sicher die größte Ungenauigkeitsquelle mit $\pm 2^\circ$. Allerdings konnte dieser Fehler durch Berechnen einer Durchschnittsabweichung (Bias), wie es im nächsten Punkt 2.4.4 angegeben ist, größtenteils korrigiert werden.
- Die vertikale Ausrichtung erfolgte optisch anhand von umgebenen vertikalen Gegebenheiten (Türstock, Fenster, aber auch dem Faden des Lotes). Die Ungenauigkeiten müssten in diesem Bereich bei ca. $\pm 1^\circ$ liegen.

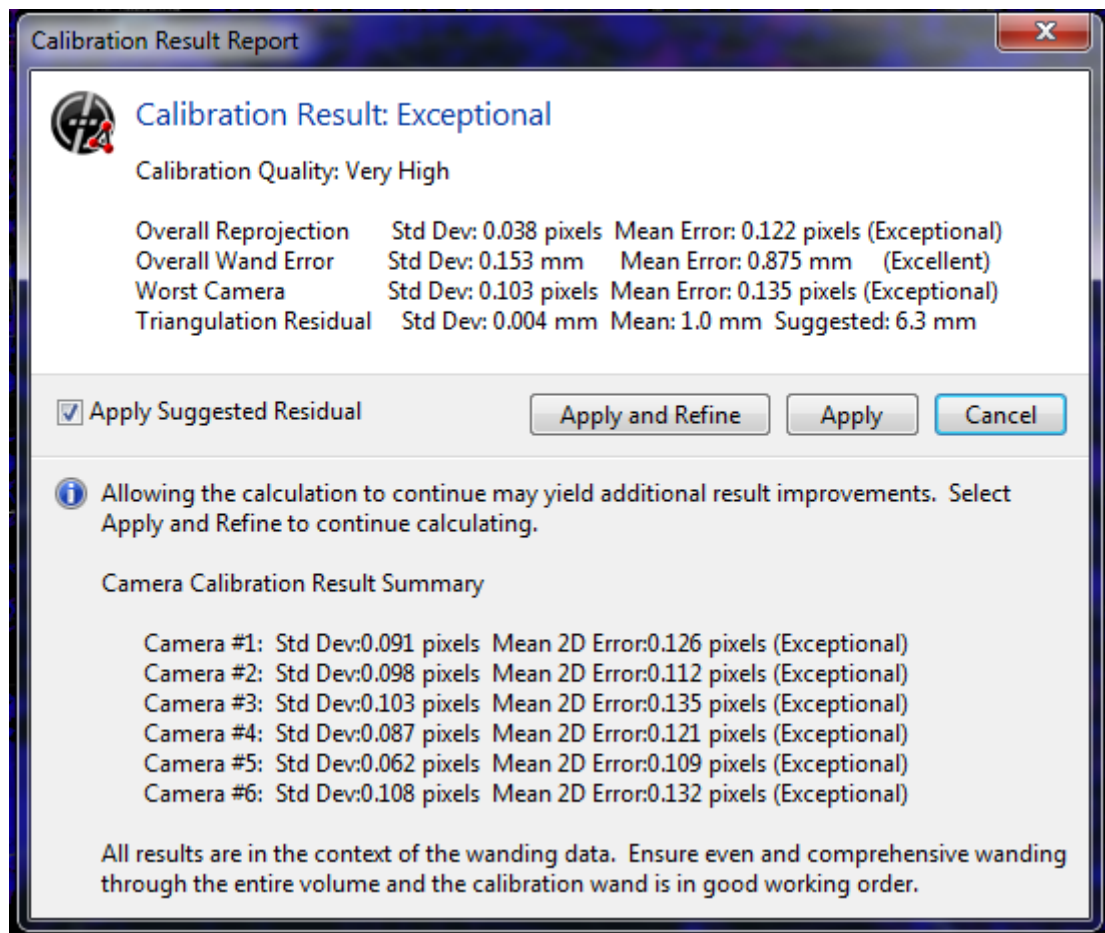


Abbildung 12: Trackingsystem Kalibrationsgenauigkeit: relevant für die Genauigkeit der Messpositionen, welche die Tracking-Software ausgibt, ist der *Overall Wand Error* mit 0,875 mm.

- Eine weitere Fehlerquelle stellt auch das Definieren eines Objektes in der Trackingsystemsoftware dar. Wenn man mehrere Punkte zusammenfasst errechnet sich die Software automatisch einen gemeinsamen Mittelpunkt, dessen Koordinaten dann für das Objekt angezeigt werden. Dieser "Mittelpunkt" kann sinnvollerweise verschoben werden, was auch erfolgte (siehe 2.2.1). Dadurch ergeben sich auch hier geringe Ungenauigkeiten, welche eine Auswirkung haben können. Es ergibt sich eine empirische Abweichung < 1 cm.
- Der Drehteller wurde für das Weiterdrehen bzw. das Retourdrehen des Mikrofons verwendet. Die Genauigkeit des Drehtellers wird vom Hersteller mit $0,05^\circ$ ¹¹ angegeben.

¹¹ET 250-3D, 3D Electronic Turntable for Polar Measurements, Outline s.r.l., Flero, Brescia, Italien; User Guide: http://89.96.202.198/documents/Z1BR_NUL0008R00.pdf, 2011

Die folgende Tabelle zeigt eine Zusammenfassung der oben angeführten Fehlereinflüsse:

Messunsicherheit des Trackingsystems:	0,875 mm
Positionierung des Mikrofons auf dem Drehteller: $\varphi, \vartheta = \pm 1^\circ, r = \pm 1 \text{ mm}$	
Horizontale Ausrichtung des Mikrofons:	$\pm 2^\circ$
Vertikale Ausrichtung des Mikrofons:	$\pm 1^\circ$
Objektdefinition in der Tracking-Software:	max. 1 cm
Genauigkeit des Drehtellers:	0,05° (Herstellerangabe)

2.4.4. Grobe Genauigkeitsabschätzung

Die Summe der oben angegebenen Ungenauigkeiten (Trackingsystem, Objektzuweisung, Drehteller sowie Exzentrizitätsfehler und Ausrichtungsfehler des Mikrofons) ergeben einen Gesamtgenauigkeitsbereich. Im schlimmsten Fall addieren sich diese Abweichungen. In den nächsten Schritten betrachten wir die Fehler in X , Y und Z als unabhängig voneinander, sodass sich als Fehlerbereich eines Punktes keine Kugel sondern ein Quader ergibt. Diese Annahme ist zwar nicht sehr wahrscheinlich und geht von einem zu großen Fehler aus, ist jedoch als grobe Fehlereingrenzung genauso geeignet. Ferner sind die nun durchgerechneten Positionen bei diesen Vereinfachungen nicht wirklich das Worstcase-Szenario, jedoch vereinfachen sich die Überlegungen und die Abweichungen liegen innerhalb jener, welche bei der Unabhängigkeitsannahme zuviel gemacht wurden.

ϑ : Eine gute Näherung für den maximalen Winkelfehler bei ϑ ergibt sich in der Horizontalebene. Da die Ungenauigkeiten in X und Y (sowie in φ und ϑ) Richtung gleich groß sind, kann der generelle Maximalfehler in X und Y Koordinaten wie folgt angegeben werden (Abbildung 13 stellt eine Skizze der ungünstigsten Fehlerüberlagerung dar).

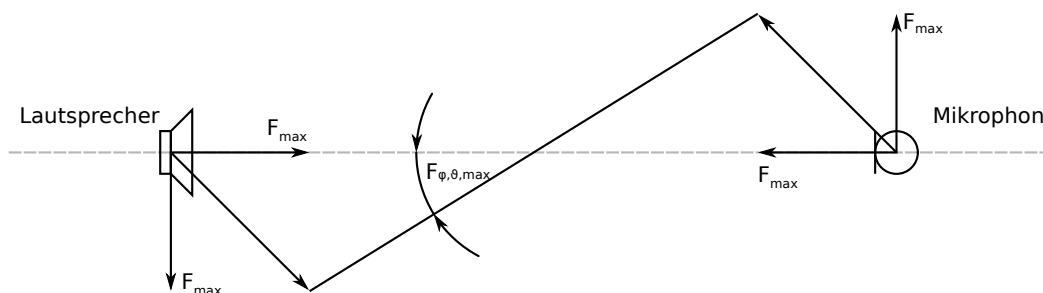


Abbildung 13: Skizze der ungünstigsten Fehlerüberlagerung bei $\varphi, \vartheta = 0^\circ$ (F_{max} = maximaler Fehler in x - y - und z -Richtung). $F_{\varphi, \vartheta, max}$ ergibt sich dabei aus der Fehlpositionierung von Lautsprecher und Mikrophon.

$$F_{max} = F_{Trackingsystem} + F_{Punktzuweisung} + r \cdot \tan\left(\frac{F_{Exzentrizität}}{2}\right) = 11,9 \text{ mm} \quad (14)$$

Angenommen diese Fehler würden auch von den Richtungen her maximal ungünstig zusammentreffen, ergibt sich der maximale Winkelfehler für ϑ zu:

$$F_{\vartheta,max} = \arctan\left(\frac{(2 \cdot F_{max})}{(r - 2 \cdot F_{max})}\right) + F_{Vertikalausrichtung} = 2,16^\circ \quad (15)$$

φ : Für φ ergibt sich eine ähnliche Vorgehensweise wie bei Theta, der maximale Winkelfehler hängt allerdings auch noch von der Ungenauigkeit des Drehtellers ab.

$$F_{\varphi,max} = \arctan\left(\frac{(2 \cdot F_{max})}{(r - 2 \cdot F_{max})}\right) + F_{Horizontalausrichtung} + F_{Drehteller} = 3,21^\circ \quad (16)$$

2.4.5. Bias

Später wurde dann über einen erweiterten Frequenzbereich (von 140 bis 3096 Hz) bei allen Positionen eine Durchschnittsabweichung ermittelt und diese korrigiert (siehe 4.1 - Entzerrung). Der Fehler belief sich auf $\varphi_B = -1,7^\circ$ und $\vartheta_B = +1,5^\circ$. Das bedeutet es musste jeweils das Gegenteilige dazu addiert werden.

Dadurch sind die unter 2.4.4 angeführten Punkte natürlich auch betroffen, weshalb nach dem Berechnen des Durchschnittsfehlers mit Sicherheit ein geringerer Streubereich anzunehmen ist, welcher allerdings nicht mehr exakt in Zahlen gefasst werden kann. Hierbei sei primär auf die Ergebnisse (Datenanalyse) verwiesen, deren Aussagekraft offensichtlich nicht zu sehr durch derlei Ungenauigkeiten gestört wurde.

3. Datenverarbeitung

In Abbildung 14 sind die einzelnen Stationen in Form eines Flussdiagramms aufgezeichnet, welche die Daten jedes einzelnen Messpunktes durchlaufen müssen. Die dafür notwendigen Funktionen in MATLAB¹² sind im Anhang A näher beschrieben, an dieser Stelle soll nur auf die dick umrandeten Stationen genauer eingegangen werden.

¹²MATLAB, v. 7.9.0.529 (R2009b), The MathWorks Inc., Natick, Massachusetts, USA

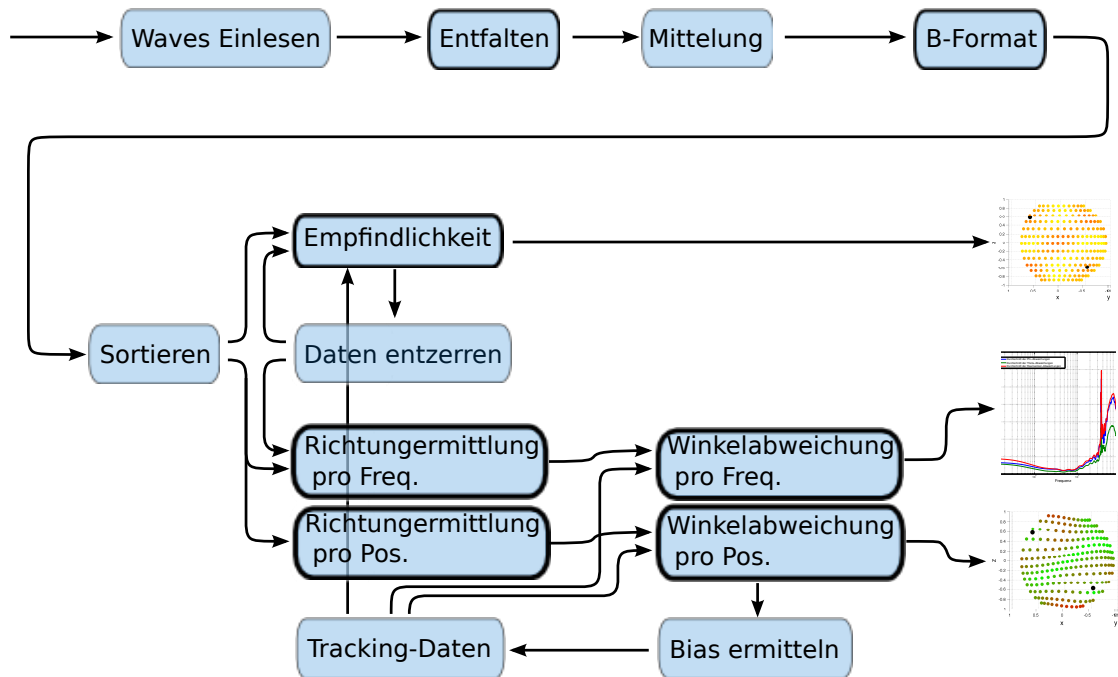


Abbildung 14: Flussdiagramm der Datenverarbeitung

3.1. Entfaltung

Im Folgenden sollen Beispiele zur Impulsantwort betrachtet werden, wobei Bezug auf die theoretischen Hintergründe in 1.2.3 genommen wird. In Abbildung 15 (links) ist die bereits gemittelte Impulsantwort bei $\varphi = 42^\circ$ und $\vartheta = -50^\circ$ über die gesamte Messdauer angeführt.

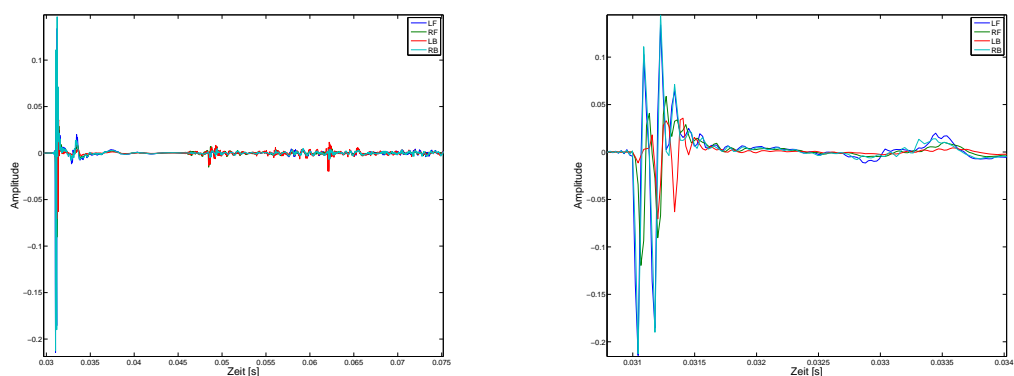


Abbildung 15: Impulsantwort (links) und Ausschnitt (rechts) der Impulsantwort vom ersten Peak bis zu der ersten Reflexion bei $\varphi = 42^\circ$ und $\vartheta = -50^\circ$. Die Kanäle sind, farblich markiert, übereinandergelegt. In der linken Abbildung sind der Direktschall, die ersten Reflexionen und nach dem sehr flachen Verlauf (geringes Rauschen) weitere Reflexionen, sowie die harmonischen Verzerrungen zu sehen.

Deutlich zu sehen ist der erste Impuls des Direktschalls und wenig später ein weiterer Peak durch die erste Reflexion (Boden), sowie rechts weitere Reflexionen und der Beginn eines diffusen Nachhalls. Das Hintergrundrauschen ist bei dieser Darstellung zu gering, als dass es erkennbare Einflüsse hätte. Die unterschiedlichen Farben stehen für je eine der Mikrofonkapseln (Bezeichnung siehe 1.2.2).

Abbildung 15 (rechts) zeigt den Ausschnitt dieser Impulsantwort bis zu der ersten Reflexion. Aus dem Zeitunterschied zwischen den Peaks des Direktschalls und den ersten Reflexionen kann bei angenommener Schallgeschwindigkeit von $c = 343 \text{ m/s}$ (Raumtemperatur $\vartheta_{Temp} = 20^\circ \text{ C}$) ein Abstand von ca. 0,78 m (2,27 ms, 100 Samples) ermittelt werden, welcher der Messanordnung entspricht.

Die harmonischen Verzerrungen, primär des Lautsprechers, wurden an das Ende gefaltet (Abbildung 16). Entsprechend der Theorie (siehe 1.2.4) sind die harmonischen Verzerrungen zum Teil trennbar, sie beginnen zeitlich versetzt. In Abbildung 16 sind neben dem ersten Spitzenwert zwei weitere Peaks zu erkennen, welche den zusätzlichen Scheinanregungen entsprechen (Abbildung 3).

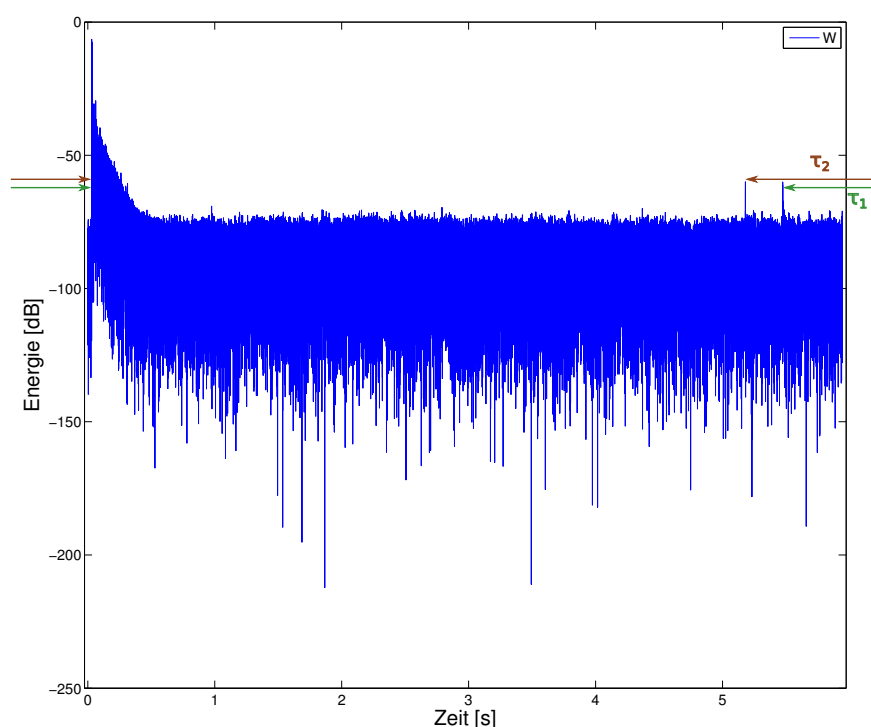


Abbildung 16: Energie pro Zeit: τ_1 und τ_2 kennzeichnen den Abstand zwischen den zusätzlichen Peaks der harmonischen Verzerrungen und dem des Direktschalls (siehe auch Abbildung 3)

Um die Peaks besser erkennen zu können, wurde die Energieskala in dB verwendet. Da die Harmonischen nur durch die Faltung im rechten Bereich erscheinen muss der Abstand vom Ende bis zum Direktschall ausgerechnet werden (s. Abbildung 16). In diesem Fall ergeben sich aus der Messung ein $\tau_1 = 0,5017$ s und ein $\tau_2 = 0,7952$ s. Theoretisch müssten sich $\tau_1 = 0,5014$ s und $\tau_2 = 0,7954$ s ergeben. Die geringeren Differenzen ergeben sich aufgrund des Einschwingverhaltens. Eine 3. Harmonische (k4) ist nicht mehr erkennbar.

3.2. B-Format

Die Theorie zum B-Format wird in 1.2.1 ausführlich behandelt. Im Folgenden soll wie im vorhergehenden Punkt die Impulsantwort betrachtet und exemplarisch die Richtungsinformation kontrolliert werden. Abbildung 17 (links) zeigt die gleiche Impulsantwort wie Abbildung 15, nur dass nun Schalldruck (w) und Schallschnelle (x, y, z) dargestellt sind. Abbildung 17 (rechts) zeigt den vergrößerten Bereich bis zur ersten Impulsantwort. Da

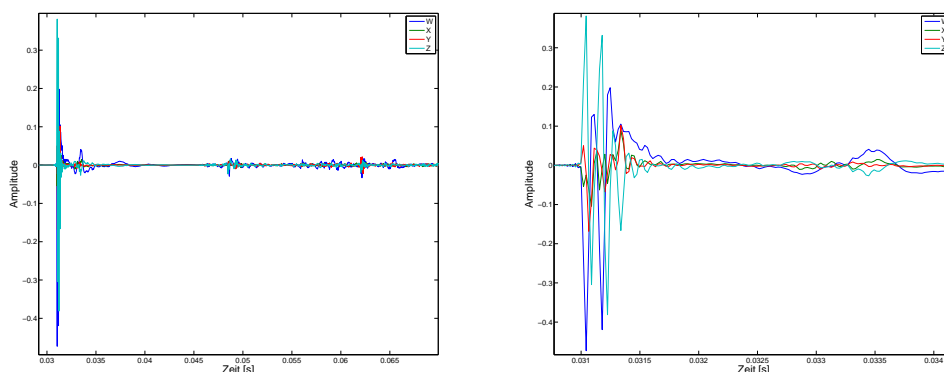


Abbildung 17: Impulsantwort bei $\varphi = 42^\circ$ und $\vartheta = -50^\circ$ im B-Format über den gesamten Frequenzbereich, inklusive der ersten paar Reflexionen (links) und nur bis zu der ersten Reflexion (rechts)

aufgrund des räumlichen Aliasings 1.2.2 die hohen Frequenzen für einen hohen Rauschanteil sorgen, lässt sich in dieser Darstellung keine sichere Richtungsinformation ablesen. Mit Hilfe einer Tiefpassfilterung (Abbildung 18) lässt sich der störende Anteil eliminieren, wodurch ein geglätteter Kurvenverlauf erreicht wird. $\varphi = 42^\circ$ und $\vartheta = -50^\circ$ bedeutet, dass der Schall die beiden unteren Mikrofonkapseln (LF und RB) früher als die oberen erreicht. In kartesischen Koordinaten gedacht bedeutet das, dass sich die Schallquelle im positiven x - und y -Bereich, jedoch im negativen z -Bereich befindet. Folglich müssen der x - und y -Kanal dem Ausschlag des w -Kanals folgen, während der z -Kanal genau

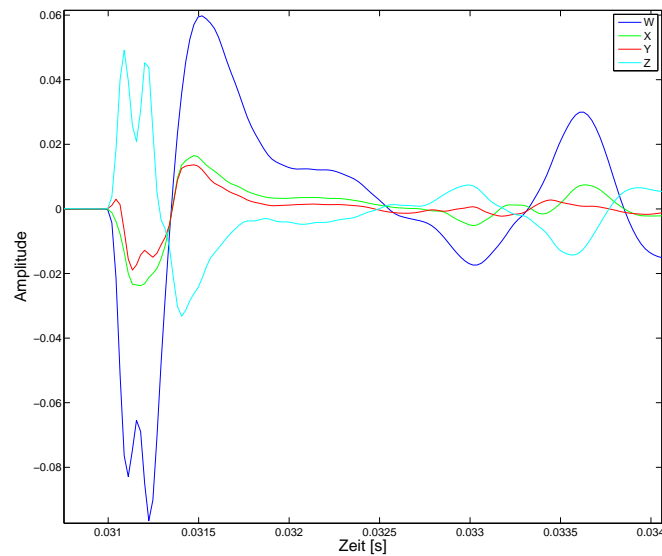


Abbildung 18: Tiefpassgefilterter Ausschnitt der Impulsantwort bei $\varphi = 42^\circ$ und $\vartheta = -50^\circ$ vom ersten Peak bis zu den ersten Reflexionen im B-Format. Tiefpass: $f_c = 2000$ Hz

gegengleich ausschlägt. Besonders im Bereich des Direktschalls ist dies in Abbildung 18 sehr gut zu sehen.

3.3. Empfindlichkeit

Die Betrachtung der Empfindlichkeit ist rein eine relative zwischen den Unterschiedlichen Richtungen. Die Daten wurden gleich mittels Scatterplots zur Darstellung gebracht. Sinnvollerweise wurden dafür wieder die mit dem Trackingsystem gemessenen Positionen verwendet. Dabei ergaben sich eindeutig erkennbare Bereiche, in denen Signale signifikant lauter oder leiser dargestellt wurden.

3.4. Richtungsermittlung

Bei der Ermittlung der Richtungsinformation (Theorie s. 1.2.5) sind die Phasenmehrdeutigkeiten zu beachten, da sonst bei den Ergebnissen nebeneinander liegender Punkte (Umrechnung von kartesischen in Kugelkoordinaten) Sprünge über 180° entstehen können. Diese Sprünge würden in diesem Fall zu unverhältnismäßig großen Unterschieden im eigentlich kontinuierlichen Richtungsänderungsverlauf führen.

Die Berechnung erfolgte zum einen pro Frequenz (summiert über alle Positionen) und zum anderen pro Position (summiert über spezielle Frequenzbereiche).

3.5. Winkelabweichung

Für die Winkelabweichung wurde die Differenz der gemessenen Daten und den protokollierten Trackingpositionen gebildet.

Bei der Berechnung des Raumwinkels, bzw. der Raumwinkelabweichung ist zu beachten, dass die Richtungsinformation bezüglich eines Koordinatensystems verloren geht, da die beiden Vektoren mit konstantem Raumwinkel γ beliebig im Raum angeordnet werden können.

Wie unter 3.4 wurde auch hier in Ergebnisse pro Frequenz und Ergebnisse pro Position unterschieden.

4. Datenanalyse

Die Ziele der Datenanalyse waren, zum einen die richtungsabhängige Empfindlichkeit der Schallaufnahme zu beschreiben, zum anderen einen Zusammenhang zwischen der Darstellungsgenauigkeit und der Frequenz zu finden. Weiters sollte die örtliche Abhängigkeit der Genauigkeit möglichst anschaulich erläutert werden. Um die Einflüsse des Lautsprechers auf die Messung möglichst gering zu halten, wurde eine Entzerrung vorgenommen.

Wie bereits unter 2.4.1 und 3.1 erwähnt, wurde von den Messergebnissen zur Minimierung von Störeinflüssen ein Analysefenster von 100 Samples verwendet.

4.1. Entzerrung

Zu Beginn ist es sinnvoll den Übertragungsfrequenzgang (Lautsprecher - Raum - Mikrofon) zu kennen. Dieser ergibt sich (bestmöglich) aus dem Durchschnitt aller gemessenen Messpunkte (siehe Abbildung 19).

Auch hier ist deutlich zu erkennen bei welcher Wellenlänge es zu Auslöschungen zwischen den einzelnen Kapselsignalen kommt. Der Frequenzbereich unterhalb vom Peak (1308 Hz) ist dagegen direkt auf den Einfluss des Lautsprechers zurückzuführen (siehe auch Abbildung 6), weshalb eine Entzerrung in Betracht zu ziehen ist.

Oberhalb des Peaks beginnt die Zerfurchung des Frequenzganges aufgrund bereits beschriebener Auslöschungen. Diesen Frequenzbereich eins zu eins so zu entzerren ist unsinnig, da es sich dabei um eine mikrofontypische Charakteristik handelt, welche bei jeglicher anderer Anwendung genauso zum Tragen kommt. Eine Entzerrfunktion in die-

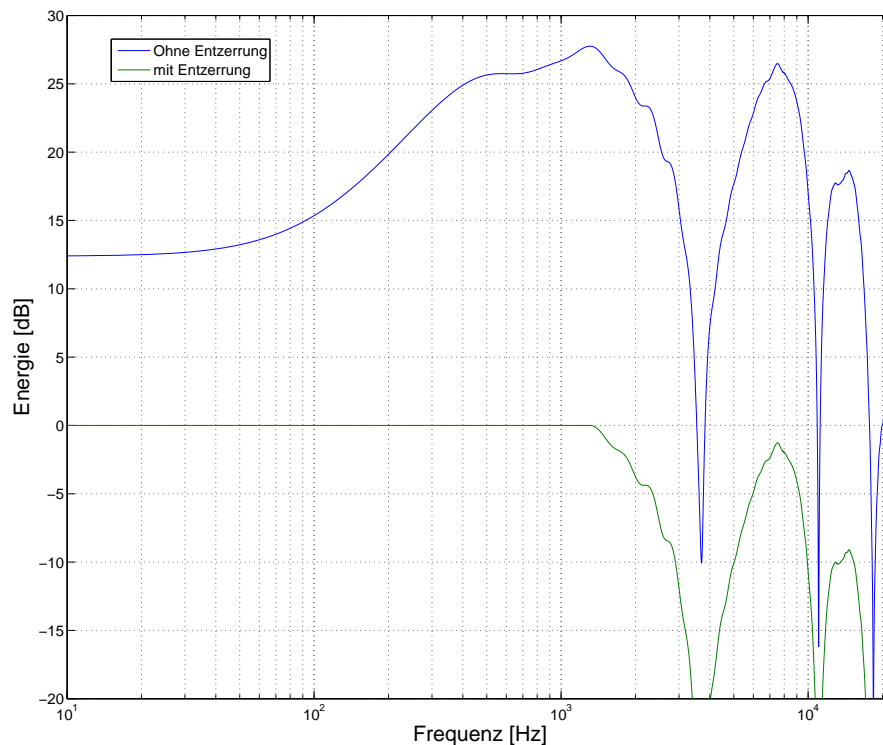


Abbildung 19: Frequenzgang des W-Kanals: Um den Einfluss des Frequenzgangs des Lautsprechers zu minimieren wird der Frequenzgang bis 1308 Hz entzerrt.

sem Bereich müsste also die Einschnitte ignorieren und in etwa nur über die Maxima interpolieren. Allerdings ist dabei schwer feststellbar, inwiefern die einzelnen Einflüsse wirklich vom Lautsprecher kommen und inwieweit diese durch das Mikrofon entstehen.

Des Weiteren kann man bei jedem realistischen (nicht elektronisch erzeugten) Signal von einem natürlichen Höhenabfall ausgehen, welcher sich in ähnlicher Form bemerkbar macht. Daher ist eine Entzerrfunktion sinnvoll, welche nur den tiefen Frequenzbereich verändert (1308 Hz). Die Entzerrfunktion bildet sich durch Invertierung des Frequenzgangs bis 1308 Hz. Um die Daten nun zu entzerren, wurde der jeweilige Frequenzgang mit der Entzerrfunktion multipliziert.

Da die Entzerrung im Vergleich, wie in Abbildung 20 zu sehen, keinen nennenswerten Unterschied erzeugt, dafür jedoch die Energie im oberhalb von 200 Hz drastisch reduziert, wird im weiteren Verlauf keine Entzerrung der Signale verwendet.

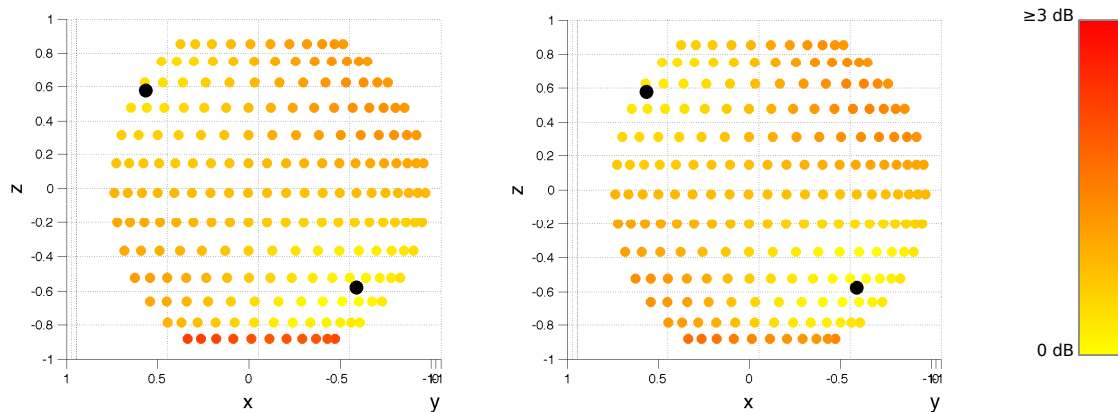


Abbildung 20: Vergleich zwischen aufgenommener Energie entzerrt (links) und nicht entzerrt (rechts): TP: $f_c = 3$ kHz, 2048 FFT, die Farbskala geht kontinuierlich über 3 dB, wobei der kleinste Wert gleich 0 dB entspricht. Deutlich sichtbar ist der geringe Unterschied, weshalb auf die Entzerrung verzichtet wurde.

4.2. Empfindlichkeit

Zusätzlich zur Betrachtung der örtlichen Darstellungsfehler (siehe 4.4), ist es auch interessant die generelle Empfindlichkeit des Mikrofons über den Aufnahmebereich zu betrachten. Dabei geht es nicht um absolute, sondern rein um relative Werte, das heißt ob Schall aus einer bestimmten Richtung lauter als aus einer anderen aufgenommen wird. Hierbei geht es nicht um die Richtungsinformation, sondern rein um den Gesamtpegel, die Gesamtenergie. Daher wird bei den folgenden Betrachtungen nur der w -Kanal verwendet, welcher bekanntlich keinerlei Richtungsinformation enthält. Bei den Diagrammen kann man anhand der Farbskala die Pegelverhältnisse ablesen. *Gelb* bedeutet hierbei immer wenig und *Rot* viel Pegel. Die Skala erstreckt sich bei allen weiteren Grafiken jeweils kontinuierlich über 10 dB. Das untere Ende der Skala wird vom jeweils geringsten im Diagramm vorkommenden Wert festgelegt. Falls es größere Unterschiede gibt (als 10 dB über dem Minimum), werden die lauterempfindlicheren Stellen weiterhin im maximalen Rot dargestellt.

4.2.1. Gesamtsignalbetrachtung

Um einen Überblick zu schaffen, was man hier zu erwarten hat, ist es sinnvoll zuerst den gesamten Frequenzbereich summiert zu betrachten (Abbildung 21).

Man kann sehr gut sehen, dass es ausgeprägte Tendenzen gibt, denen eine Systematik zu Grunde liegt. In Bereichen direkt über, genau zwischen zwei und genau zwischen drei Kapseln ist das Soundfield-Mikrofon empfindlicher als zwischen diesen Bereichen. Bei der

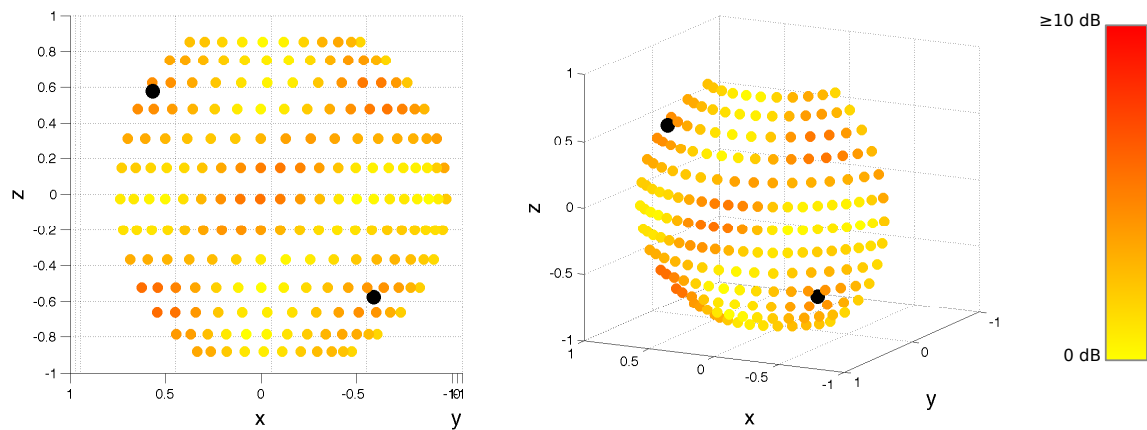


Abbildung 21: Energieverlauf W-Kanal über alle Frequenzen (2048 FFT). Die Skala beschreibt den Unterschied der Energieanteile kontinuierlich bis 10 dB, wobei 0 dB für den geringsten berechneten Wert gilt. In den folgenden Abbildungen gelten die gleichen Bedingungen

entzerrten Summe handelt es sich immerhin um 5,8 dB Unterschied zwischen minimaler und maximaler Empfindlichkeit.

4.2.2. Einzelne Frequenzbereiche

Die Einteilung der Frequenzbereiche erfolgt nach dem gleichen Schema wie unter 4.4.1:

Mitten (Idealbereich) Wie erwartet sieht diese Grafik (Abbildung 22) sehr homogen aus. Im Bereich zwischen 152 und 1034 Hz kommt es maximal zu 0,8 dB Unterschied zwischen den unterschiedlichen Richtungen. Dies bestätigt auch, dass in diesem Frequenzband die Richtcharakteristiken sehr sauber eingehalten werden und sich beinahe perfekt auf eine einheitliche Kugel addieren.

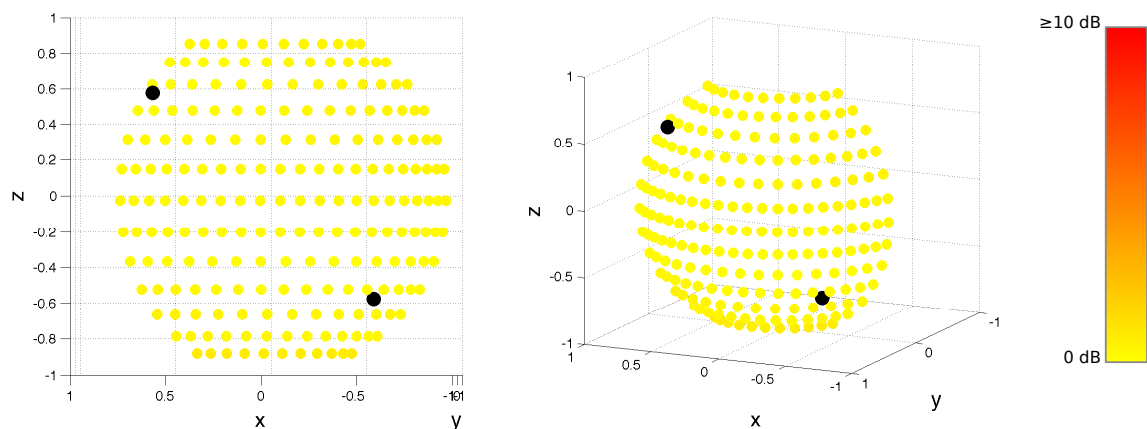


Abbildung 22: Energieverlauf W-Kanal zwischen 152 Hz und 1034 Hz (2048 FFT)

Hochmitten Auch in Abbildung 23 erkennt man aufs erste Hinsehen eine einheitliche Farbe, was für eine konstante Empfindlichkeit spricht. Allerdings ergibt sich im Bereich zwischen 1034 und 3618 Hz eine Differenz von 2,3 dB zwischen dem Maximum und dem Minimum, welche optisch kaum auffällt. Bei genauerer Betrachtung erkennt man jedoch eine Tendenz, welche die Richtungen direkt über der Kapsel leicht benachteiligt (leiser darstellt).

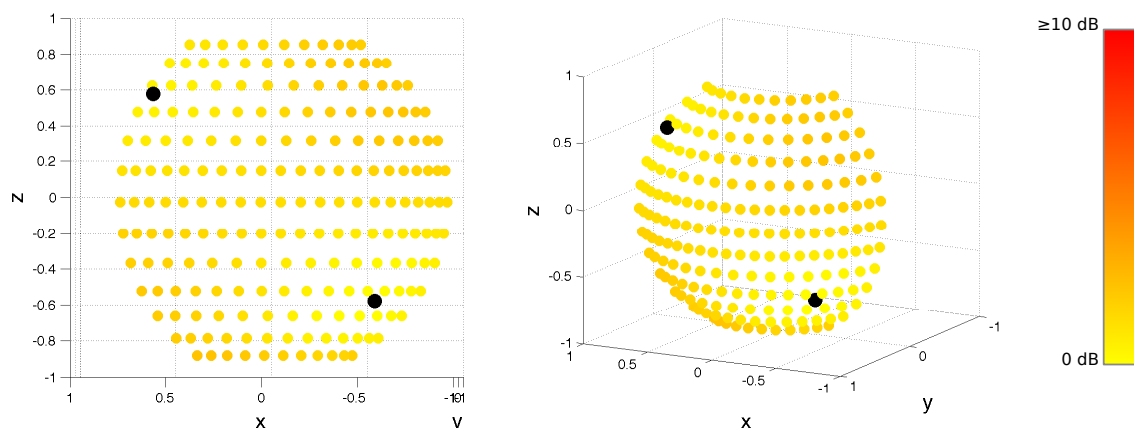


Abbildung 23: Energieverlauf W-Kanal zwischen 1034 Hz und 3618 Hz (2048 FFT)

Tiefen Theoretisch müssten die Richtcharakteristiken (Nieren) der Mikrofonkapseln bei tiefen Frequenzen besonders gut gehalten werden, praktisch kann man allerdings oft ein leichtes Kugelverhalten feststellen. Beim Soundfield-Mikrofon würde das bedeuten, dass die Empfindlichkeit in beiden Fällen sehr konstant sein sollte. Selbst wenn es sich im Extremfall um vier Kugel- statt Nierenmikrofone handeln würde, müsste die Empfindlichkeit in alle Richtungen gleich groß sein. In diesem Diagramm (Abbildung 24) allerdings ist das überhaupt nicht der Fall, der maximale Unterschied beträgt für Frequenzen zwischen 43 Hz und 151 Hz 10 dB.

Interessant ist, dass es offensichtlich nichts mit den Mikrofonpositionen zu tun hat (wie bisher), sondern rein vom Azimut bestimmt wird. Daher liegt die Vermutung nahe, dass die Bodenreflexion für dieses Phänomen zuständig ist.

Höhen Es ist zwar leicht möglich hier noch Tendenzen zu erkennen, jedoch springen die Werte frequenzabhängig sehr stark, da die Wellenlänge in den Bereich der Kapselabmessungen kommt. Aus diesem Grund ist eine genauere Betrachtung für den gesamten Höhenbereich ab ca 3,6 kHz nicht sinnvoll.

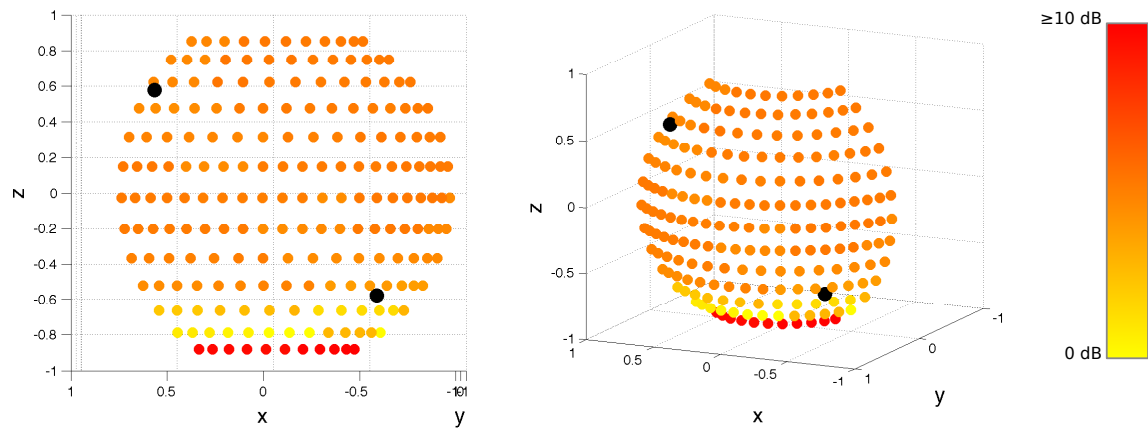


Abbildung 24: Energieverlauf W-Kanal zwischen 43 Hz und 151 Hz (2048 FFT)

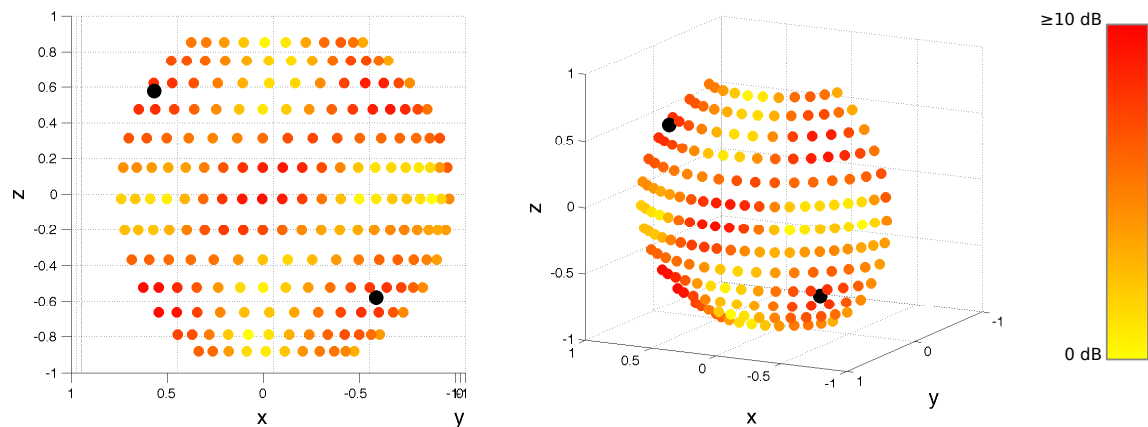


Abbildung 25: Energieverlauf W-Kanal zwischen 3618 Hz und 22028 Hz (2048 FFT)

4.3. Durchschnittliche Winkelabweichung

In diesem Punkt wird die Abbildungsgenauigkeit pro Frequenz gemittelt über alle Positionen betrachtet. Abbildung 26 (links) zeigt die durchschnittliche Winkelabweichung von φ , ϑ und γ über den gemessenen Frequenzbereich in Grad an. 'Durchschnittlich' bezieht sich dabei auf die Mittelung über alle Positionen in Abhängigkeit zur Frequenz. Beginnend im unteren Bereich sind die Abweichungen noch verhältnismäßig gering, bei sehr tiefen Frequenzen abfallend und ab 500 Hz wieder ansteigend, bis es bei ca. 3,6 kHz eine steilflankige Überhöhung gibt. Diese lässt sich dadurch erklären, dass die Wellenlänge in diesem Bereich zur Hälfte zwischen die Mikrofonkapseln passt, wodurch es zu Auslöschungen kommt und ab diesem Bereich eine Winkelmehrdeutigkeit entsteht. Das führt dazu, dass die Phase der eigentlichen Frequenz (für alle Frequenzen oberhalb dieses Bereichs) nicht mehr eindeutig bestimmbar ist. Dieser Effekt tritt auch auf, wenn der Kapselabstand der 1,5- und 2,5-fachen Wellenlänge entspricht (bei größerer Samplerate auch bei 3,5-facher Wellenlänge, etc.).

Oberhalb der ersten Spitze liegen aufgrund der Winkelmehrdeutigkeiten alle Winkelabweichungen teils weit über 15° , weshalb in diesem Frequenzbereich keine sinnvolle Bestimmung der Schallquellenposition mehr möglich ist. Auffällig ist hier, dass sich die Winkelabweichungen vor allem in horizontaler Richtung ausprägen, während die vertikalen Abweichungen im Schnitt mindestens 15° darunter liegen.

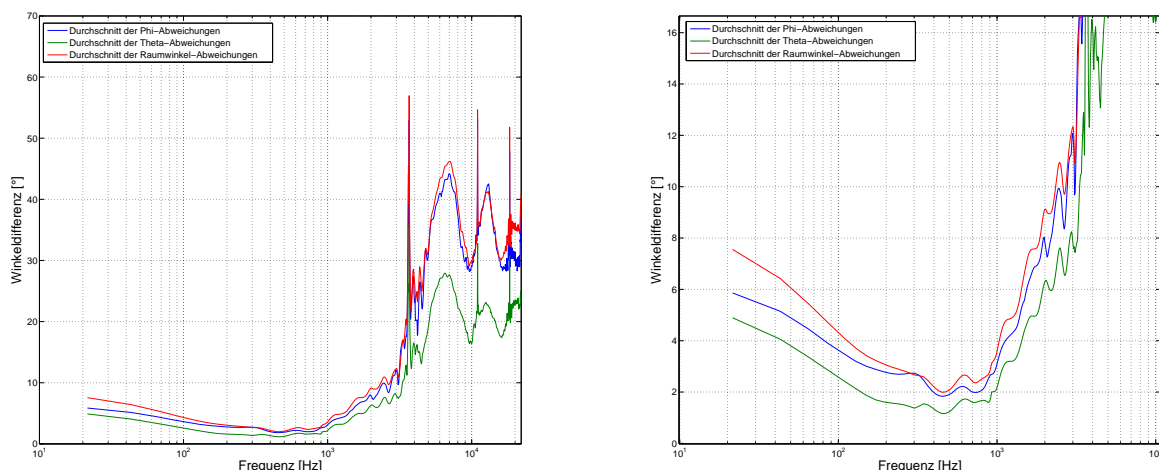


Abbildung 26: Gemittelte Winkelabweichung über alle Positionen, in Abhängigkeit zur Frequenz (2048 FFT). Der Abstand der Peaks zueinander wird von der Anordnung der Kapseln bestimmt. Links: gesamter Frequenzbereich ,rechts: vergrößerte Darstellung

Der untere Frequenzbereich lässt sich noch feiner untergliedern. Abbildung 26 (rechts) greift diesen Frequenzbereich heraus. Es ist deutlich sichtbar, dass alle Kurven im Bereich um 500 Hz ein Minimum besitzen, wobei aus der genaueren Analyse der Daten ein Bereich von 140 Hz bis etwa 1 kHz herausgegriffen werden kann, in dem alle Abweichungen unter 4° liegen. Dies kann in Anbetracht eventueller Ungenauigkeiten bei der Bestimmung der Messpositionen durch das Trackingsystem als hohe Genauigkeit eingestuft werden.

Unterhalb von 140 Hz steigt die Abweichung noch einmal bis 8° , bzw. 10° an, was durch die schlechte Abstrahlung des Lautsprechers in diesem Bereich und den damit schlechteren Rahmenbedingungen erklärbar ist. Oberhalb von 1 kHz nimmt die Abweichung stetig bis 10° , bzw. 20° zu und geht ab 3,6 kHz in die starke Überhöhung über. Der Bereich bis knapp unter 3,6 kHz kann noch als relativ genau eingestuft werden.

4.4. Winkelabweichung pro Position

Um örtliche Abweichungstendenzen formulieren zu können, muss ein Bezug der berechneten Positionen zu den Protokollierten hergestellt werden. Für die Darstellung wurde

die Raumwinkelabweichung verwendet. Wichtig ist allerdings, dass über den Raumwinkel γ keine Aussage getroffen werden kann, in welche Richtung der Vektor tatsächlich abweicht. Daher ist bei einer dreidimensionalen Darstellung die Kodierung der Raumwinkelabweichung mittels einer Farbskala an jeder einzelnen Position sinnvoll. Für die Farbskala wurde ein Übergang von *Grün* (Abweichung = 0°) bis *Rot* (Abweichung $\geq 12^\circ$) gewählt, um auch im Bereich zwischen 140 Hz und 1 kHz noch Unterschiede erkennen zu können. Zusätzlich kann die Toleranz oberhalb dieses Bereiches nur noch als genügend bis ungenügend betrachtet werden, weshalb eine stärker differenzierte Unterscheidung nicht mehr notwendig ist.

4.4.1. Einzelne Frequenzbereiche

In den folgenden Abbildungen sind die berechneten Positionen mit der entsprechenden Farbskala, nach dem am Anfang dieses Kapitels beschriebenen Prinzip, für verschiedene Frequenzbereiche dargestellt. Das Zeitfenster, welches um den Direktschall (erster Impuls) der Impulsantwort gelegt wurde, beträgt wieder 100 Samples, um den Einfluss von Reflexionen zu minimieren. Dabei erfolgt die Unterteilung nach folgendem Schema:

Mitten (Idealbereich): In diesem Bereich sind die Winkelabweichungen im Schnitt kleiner als 4° , weshalb dieser Bereich unter den gegebenen Bedingungen als ideal für die Richtungsbestimmung angenommen werden kann. In Abbildung 27 sind daher die meisten Punkte grün gefärbt, lediglich unterhalb der *rightfront*-Mikrofonkapsel ist eine höhere Abweichung erkennbar.

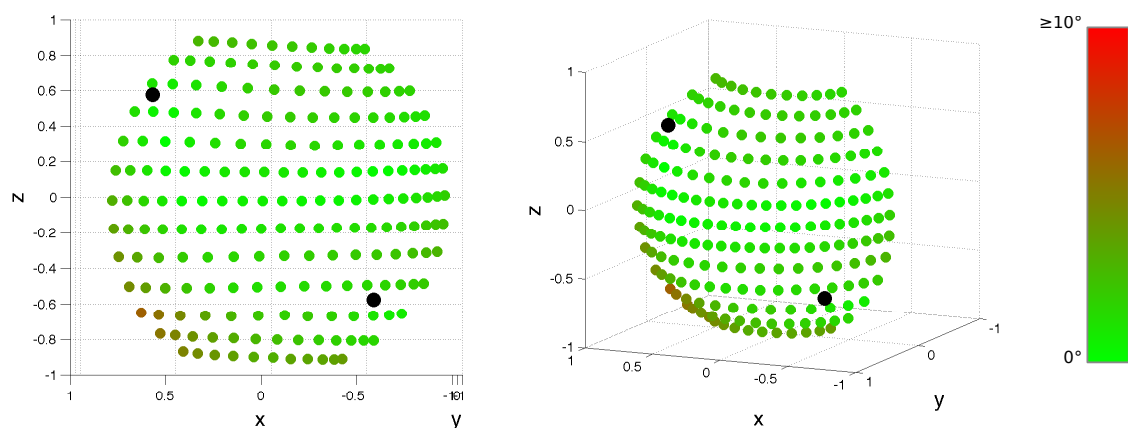


Abbildung 27: Raumwinkelabweichung (2048 FFT) Idealbereich: 152 Hz bis 1034 Hz. Die Skala beschreibt die Winkelabweichung kontinuierlich bis 12° . In den folgenden Abbildungen gelten die gleichen Bedingungen

Hochmitten (Driftbereich): Für diesen Frequenzbereich muss bezüglich der Raumwinkelabweichung eine örtliche Unterteilung in drei Bereiche gemacht werden (Abbildung 28):

- Der erste Bereich liegt jeweils über der Mikrofonkapsel und zeichnet sich durch eine geringe (im 4°-Bereich) Abweichung aus.
- Der zweite Bereich verläuft auf einer Linie, welche zu beiden Mikrofonkapseln den gleichen Abstand hat und hat ebenfalls eine geringe bis sehr geringe Abweichung an den einzelnen Punkten.
- Der dritte Bereich verläuft jeweils radialsymmetrisch um die Mikrofone und besitzt Punkte mit Abweichungen bis außerhalb der Skala.

Die Systematik, welcher der Verschiebung der Punkte zugrunde liegt, ist in 4.4.2 beschrieben.

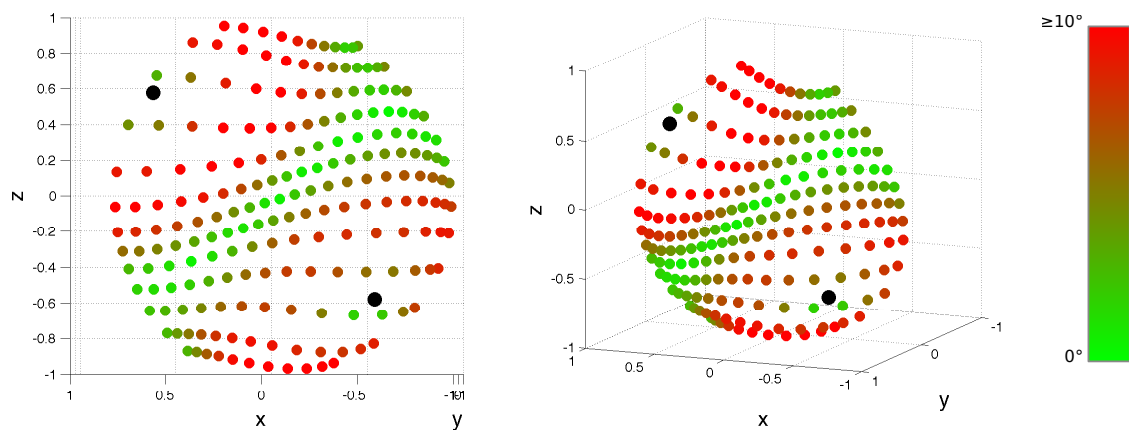


Abbildung 28: Raumwinkelabweichung (2048 FFT) Hochmittenbereich (Driftbereich): 1034 Hz bis 3618 Hz

Tiefen: In Abbildung 29 sieht die Verteilung der Raumwinkelabweichungen anders aus. Hier liegen die Punkte mit der geringsten Abweichung direkt über den Mikrofonkapseln und in direkter Verbindung dazu. Die Bereiche, welche von allen vier Mikrofonkapseln am weitesten entfernt sind, besitzen hier die größten Abweichungen, wobei auffällig ist, dass sie im oberen Mikrofonbereich maximal im 8°-Bereich liegen, im unteren dagegen teilweise weit über den 12°-Bereich. In diesem Fall lassen sich bei einem Zeitfenster von 100 Samples Einflüsse der ersten Reflexionen im unteren Bereich nicht vermeiden, weshalb die Richtungsinformation teilweise verloren geht.

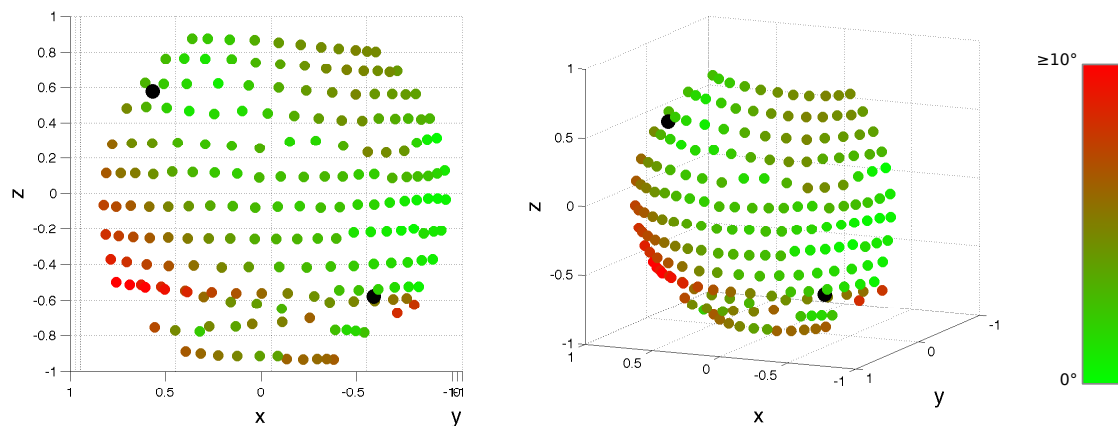


Abbildung 29: Raumwinkelabweichung (2048 FFT) Tiefen: 43 Hz bis 151 Hz über 100 Samples

Höhen: Abbildung 30 schließlich zeigt den für die Richtungsbestimmung unbrauchbaren Bereich. Wie in 4.3 bereits festgestellt, liegt hier die Winkelabweichung teilweise bei über 50° , was sich durch die eher stochastische Verteilung der berechneten Punkte bemerkbar macht. Einzig direkt über den Kapseln, an den Punkten mit der weitesten Entfernung und genau zwischen diesen vier genannten Punkten, liegen Bereiche, in denen die Raumwinkelabweichung innerhalb der Skala und vereinzelt im sehr guten Bereich liegt.

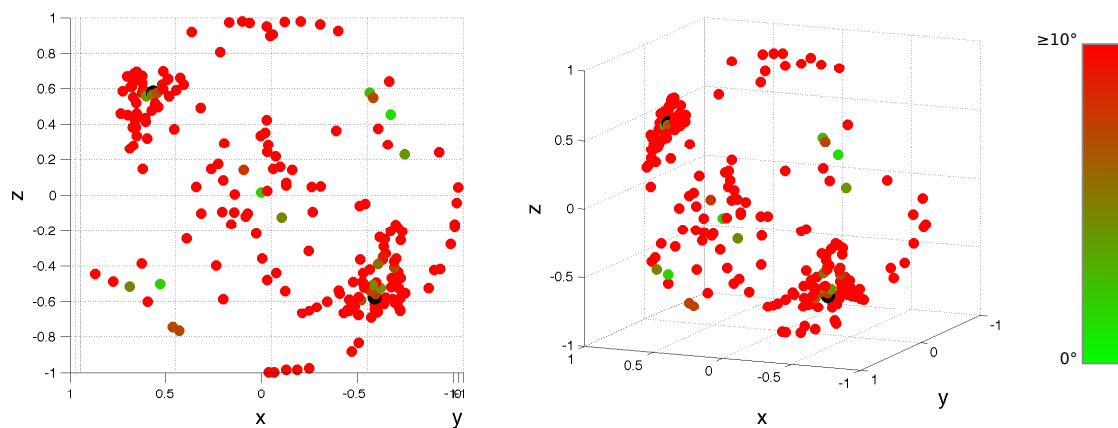


Abbildung 30: Raumwinkelabweichung (2048 FFT) Höhen: 3618 Hz bis 22028 Hz. Die stochastische Verteilung und die hohe Winkelabweichung zeigen, dass in diesem Bereich keine sinnvolle Interpretation mehr möglich ist.

4.4.2. Drift

In 4.4.1 zeigt sich bei Betrachtung der Frequenzbereiche, dass die Richtung der Winkelabweichung nicht unbedingt zufällig ist, sondern in einem gewissen Bereich symmetrisch

abdriftet. Es lässt sich feststellen, dass die Punkte dabei radialsymmetrisch vom Kapselmittelpunkt weg rücken. Daher scheint es, dass weniger Punkte im Bereich der Kapseln liegen als dazwischen. Dieser Drift entsteht oberhalb des Idealbereiches und lässt sich unter gewissen Einschränkungen auch im Höhenbereich feststellen. Im Bereich der tiefen Frequenzen gibt es keinen Drift.

Literatur

- [1] Zmölnig J. M. (2002): *Entwurf und Implementierung einer Mehrkanal-Beschallungsanlage*, unveröffentlichte Diplomarbeit, Kunstuniversität Graz
- [2] Craven P. G., Gerzon M. A. (1977): *Coincident microphone simulation covering three dimensional space and yielding various directional outputs*, United States Patent, Appl. No.: 593,244
- [3] Strauß M. (2001): *Mehrkanal-Wiedergabetechniken*, unveröffentlichte Seminararbeit, IEM, Kunstuniversität Graz
- [4] Pulkki V. (2007): *Spatial Sound Reproduction with Directional Audio Coding*, J. Audio Eng. Soc., Vol. 55, No. 6, S. 503-516
- [5] Pulkki V., Laitinen M. V., Vilkamo J., Ahonen J., Lokki T., Pihlajamäki T. (2009): *Directional audio coding - perception-based reproduction of spatial sound*, International Workshop on the Principles and Applications of Spatial Hearing, Zao, Miyagi, Japan
- [6] Farina A. (2000): *Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique*, J. Audio Eng. Soc. (Abstract's), Vol. 48, p. 350
- [7] Majdak P., Balazs P., Laback B. (2007): *Multiple Exponential Sweep Method for Fast Measurement of Head-Related Transfer Functions*, J. Audio Eng. Soc., Vol. 55, No. 7/8, S. 623-637
- [8] Warum S. (2003): *Messung von Außenohrübertragungsfunktionen und direktionalen Raumimpulsantworten als Basis der Auralisation realer Räume*, unveröffentlichte Diplomarbeit, Kunstuniversität Graz
- [9] Batke J. M. (2009): *The B-Format Microphone Revised*, Ambisonics Symposium 2009, Graz.

A. Anhang

A.1. MATLAB-Funktionen

Um die Fülle der Messdaten übersichtlich zu halten, sind die Messdaten nach jeder Weiterverarbeitung je Lautsprecherposition (vertikaler Winkel) in Structs zusammengefasst, in welche die einzelnen horizontalen Messpunkte gegliedert werden können. Die Benennung erfolgt automatisch, innerhalb der Structs nach Winkel (Azimuth) und außerhalb (Datei) nach Lautsprecherposition. Im Folgenden soll nun eine Übersicht über die verwendeten Funktionen und Skripte erstellt werden, wobei die in Schreibmaschinenschrift dargestellten Namen die entsprechenden Variablen kennzeichnen ([x] bezeichnet einen numerischen Platzhalter).

A.1.1. Datenerzeugung

Für die Datenerzeugung wird ein Sweep als Audiodatei benötigt, welcher in 5 Sekunden einmal über den gesamten Hörbereich (definiert von 20 Hz bis 20 kHz) läuft und nach jeweils 2 Sekunden Pause noch drei Mal wiederholt wird. Über die in MATLAB enthaltene Funktion *wavread* und die selbst geschriebene Funktion *messung* wurde eine entsprechende Datei erzeugt (Samplerate 44,1 kHz, Bittiefe 16 bit), wobei die Funktion *messung* *expsweep*¹³ zurückgreift um den eigentlichen Sweep zu generieren. Um die während der Messung im Waveformat aufgenommenen Dateien in MATLAB verwenden zu können, müssen die Daten in ein anderes Format umgewandelt werden, welches das Skript *WavesEinlesen* mit allen zugehörigen Automatisierungen für jeden Messpunkt übernimmt.

expsweep (function) Erzeugt einen Sweep innerhalb einer bestimmten Dauer, wobei die untere und obere Grenzfrequenz, sowie die Samplerate als weitere Parameter übergeben werden müssen.

messung (function) Generiert vier Sweeps hintereinander, wobei jeder Zweite mit einem Phasenoffset von 90° beginnt. Die Phasenänderung rührt von der ursprünglichen Annahme her, die Impulsantwort über das TDS (Time Delay Spectrometry) Verfahren erhalten zu können. Aufgrund der Einfachheit entschieden wir uns bei der Auswertung

¹³Majdak P (2010).: <http://piotr.majdak.com/download/mlib/expsweep.m>, 2011

des Materials dafür, die Impulsantwort rein über die Entfaltung zu berechnen. Die Phasensprünge müssen zwar berücksichtigt werden, ändern am Endergebnis jedoch nichts.

Zwischen den Sweeps sitzt eine Pause von 2 Sekunden. Sweepdauer und -bereich sind in der Funktion fixiert.

WavesEinlesen Lädt alle aufgenommenen Daten nacheinander und speichert sie in der in A.1 beschriebenen Weise ab. Es gilt:

`LSPos.name` Benennung jeder einzelnen Position

`LSPos.data` Speicherung der gemessenen Daten

`Pos[x]` Name des gespeicherten Structs

A.1.2. Impulsantwort

Um die Impulsantwort zu erhalten, ist eine Entfaltung nötig (Theorie s. 1.2.3). Dabei ist darauf zu achten, dass bei jedem Messpunkt vier Einzelmessungen durchgeführt wurden, welche für sich zu behandeln sind. Die vier entstandenen Impulsantworten können anschließend gemittelt werden, wodurch sich transiente Effekte während der Messung stark verringern. Die in Rohform vorliegenden Daten können nun einfach in das B-Format (s. 1.2.1) umgewandelt werden. Die Funktion *Sortieren* ist an dieser Stelle verwendet worden, um eine bessere Übersicht zu gewährleisten, da die Reihung der Daten innerhalb der Structs mit der ersten Ziffer als wichtigster zu einer Unordnung wie z.B. folgender führt: 0, 10, 100, 20, 30...

Entfaltung Lädt alle `Pos[x]` Daten nacheinander, separiert einen evtl. vorhandenen 5. Kanal (wurde bei der letzten Messung als Soundkarten-Ein- und Ausgangsverbindung aufgezeichnet) und speichert die vier relevanten Kanäle nach der Entfaltung mittels *Ir4Ch* nach folgendem Schema ab:

(Ch5) `Ir[x].name` Benennung jeder Position und Impulsantwort (Ch5 optional)

(Ch5) `Ir[x].data` Enthält die Daten der Impulsantworten

(Ch5) `Ir[y]Pos[x]` Name des gespeicherten Structs (Kanalnummer: $y = 1, 2, 3, 4$)

Ir4Ch_(function) Trennt die vier aufeinander folgenden Sweeps, transformiert sie in den Frequenzbereich und selektiert die Übertragungsfunktion durch Division mit dem

zugehörigen transformierten Anregungssignal. Die Impulsantwort ergibt sich durch Rücktransformation (Realteil). Optional können die Impulsantworten noch geplottet und als SVG-Datei abgespeichert werden.

Ir_summe Bildet den Mittelwert über die einzelnen Impulsantworten und speichert die Gesamtimpulsantwort nach gleichem Schema in ein Struct ab:

`Ir_G.name` Benennung jeder Position

`Ir_G.data` Enthält die Daten der Impulsantwort

`Ir_GPos [x]` Name des gespeicherten Structs

B-Format Umwandlung der Rohdaten aus der jeweiligen Gesamtimpulsantwort (A-Format) in das sog. B-Format. Die vier Vektoren w, x, y, z werden als Spaltenvektoren der einfacheren Handhabung halber in einer gemeinsamen Matrix vereint.

`Bfrmt.name` Benennung jeder Position

`Bfrmt.data` Enthält die Daten der Impulsantwort

`Bfrmt_Pos [x]` Name des gespeicherten Structs

Sortieren Sortieralgorithmus, welcher die einzelnen Daten über ihre Benennung (`[x].name`) sortiert.

`B_R_Pos [x]` Name des gespeicherten Structs

Entzerrung Liest den in *Empfindlichkeit_W* ermittelten Frequenzgang ein und erstellt einen Korrekturfrequenzgang (siehe 4.1). Da der Absolutpegel irrelevant ist, kommt es dabei auch zu einer generellen Pegelreduktion.

A.1.3. Richtung und Richtungsfehler

Für die Auswertung der Daten ist eine Berechnung der Richtungen unumgänglich. Da die gesamte Messung über den horizontalen und vertikalen Winkel aufgebaut und definiert ist, ist es naheliegend, die gemessene Winkelinformation zu extrahieren. Die Vorgangsweise führt hierbei über die Energie, indem das Produkt von Schalldruck p (B-Format: w) und Schallschnelle $v = (x|y|z)$ gebildet wird (siehe 1.2.5). Über die MATLAB-Funktion *cart2sph* sind die xyz -Koordinaten leicht in Winkel umzuwandeln (der berechnete Radius

enthält keine relevante Information). Aus den protokollierten Trackingdaten kann nun noch eine Winkelabweichung und eine Standardabweichung (pro Frequenzband und über die Frequenzbänder pro Position) berechnet werden, wobei die *Bias* Funktion einen evtl. vorhandenen systematischen Messfehler korrigiert. Um vom Koordinatensystem unabhängig zu sein, kann die Winkelabweichung auch in Form eines Raumvektors zwischen zwei beliebigen Vektoren betrachtet werden. Dieser Winkel ist allerdings immer positiv.

Richtungsvektor_pro_Band Generiert die Richtungsinformation aus dem B-Format. Die Daten werden durch pinweises Multiplizieren der x - y - z -Spektren mit dem w -Spektrum ermittelt, wobei keine Summation über die Frequenzen erfolgt, sondern die einzelnen Richtungen pro Frequenz in Kugelkoordinaten abgespeichert werden.

AbWink_B.name Benennung der einzelnen Positionen

AbWink_B.pos Speicherung der Daten $[\Delta\varphi \ \Delta\vartheta \ d\gamma]$

AbWink_B_[Frequenzauflösung]_Pos[x] Name des gespeicherten Structs der Einzelwerte

Bias Berechnet den Offsetfehler aus den berechneten Winkeln in Form einer arithmetischen Mittelung.

Trackingsystem_Positionen Speichert die protokollierten Daten des Trackingsystems sowohl in Rohform, als auch in einer in Bezug auf Bias und Koordinatenursprung korrigierten Form ab.

T_LSPos.name Benennung jeder einzelnen Position

T_LSPos.data Speicherung der gemessenen Daten

T_LSPos Name des gespeicherten Structs der Rohdaten

TK_LSPos.name Benennung jeder einzelnen Position

TK_LSPos.data Speicherung der gemessenen und korrigierten Daten

TK_LSPos Name des gespeicherten Structs der korrigierten Daten

Raumwink (*function*) Berechnung des Raumwinkels zwischen zwei (beliebigen) Vektoren, sowie die Winkelabweichung $\Delta\varphi$ und $\Delta\vartheta$. Die Berechnung der Winkel erfolgt dabei jeweils in rad.

Messfehler_pro_Band_Winkelabweichung Berechnet die Winkelabweichung zwischen den aus *Richtungsvektor_pro_Band* gespeicherten Daten zur jeweiligen mit dem Trackingsystem gemessenen Position. Schließlich wird über alle Positionen ein Durchschnitt pro Frequenz berechnet.

AbWink_B.name Benennung der einzelnen Positionen

AbWink_B.pos Speicherung der Daten [$\Delta\varphi$ $\Delta\vartheta$ $d\gamma$]

AbWink_B_Pos[x] Name der gespeicherten Structs der Einzelwerte

Durch_AbWink_B_[Frequenzauflösung] Name des gespeicherten Structs des Durchschnittswertes

Fkt_RiVek (function) Ermittelt den Richtungsvektor pro Position über den übergebenen Frequenzbereich bei übergebener Frequenzauflösung. Als Output wird ein Struct übergeben sowie gespeichert, welches alle Positionen mit zugehörigen Abweichungen enthält.

p.name Benennung der einzelnen Positionen

p.data Daten der Richtungsinformationen

p.wink Matrix mit folgenden Spalten: [φ , ϑ , R , $\Delta\varphi$, $\Delta\vartheta$]

P_[Frequenzauflösung]_[erster Verwendeter Frequenzpin]bis[letzter

verwendeter Frequenzpin]_Pos[x] Name der im Ordner *Richtungsvektor* abgespeicherten Structs.

Fkt_MessWinkAb (function) Bestimmt aus den Richtungsvektoren von *Fkt_RiVek* (welche als Input übergeben werden) nun die jeweiligen Abweichungen. Ansonsten gelten die selben Eingabeparameter wie bei der vorherigen Funktion. Als Output wird ein Struct erstellt, welches für alle Positionen die Abweichungen $\Delta\varphi$ $\Delta\vartheta$, Raumwinkel (dafür wurde gleichnamige Funktion verwendet), $\sigma\varphi$ und $\sigma\vartheta$ beinhaltet.

AbWink.name Benennung der einzelnen Positionen

AbWink.pos nach folgendem Schema gespeicherte Daten: [$\Delta\varphi$, $\Delta\vartheta$, $\Delta\gamma$, $\sigma\varphi$, $\sigma\vartheta$] ($\Delta\gamma$... Raumwinkelabweichung)

AbWink[Frequenzauflösung]_[erster Verwendeter Frequenzpin]bis[letzter verwendeter Frequenzpin] Name des *Richtungsvektoren* gespeicherten Structs.

A.1.4. Darstellung

Um die Positionen der einzelnen Messungen sowohl in berechneten Absolutwerten als auch in ihren Abweichungen zu berücksichtigen, wurde die MATLAB-Funktion *scatter3* und eine Farbskala verwendet. Damit lassen sich die Daten auf anschauliche Weise darstellen und analysieren. In den folgenden Skripten werden neben den eigentlichen Plots auch die Farbskalen sowie andere Darstellungen wie z.B. die durchschnittliche Winkelabweichung über die Frequenz etc. erzeugt.

Micplot (*function*) Funktion welche *scatterplots* erstellt, wobei die dafür nötigen Daten (Positionen, Fehlergröße, Bezeichnung usw.) als Inputs übergeben werden müssen. Diese Funktion wird von *Plots* aufgerufen.

Plots Das Skript ruft die Funktionen *Fkt_MessWinkAb*, *Fkt_RiVek*, *Bez* und *Micplot* bei unterschiedlichen Datenparametern (primär unterschiedlicher Frequenzbereich) auf, und erstellt damit alle nötigen scatterplots. Ferner wird das Ergebnis von *Messfehler_pro_Band_Winkelabweichung* über die Frequenz geplottet.

Empfindlichkeit_W Errechnet die Empfindlichkeit pro Position und über einzelne Frequenzbereiche und plottet diese. Zusätzlich wird ein genereller Frequenzgang (gemittelt über alle Positionen) des w -Kanals berechnet, gespeichert (siehe *Entzerrung*) und geplottet.

Durchschnitt_Zeitsignal Plottet die aufgenommenen Sweeps, sowohl mit Amplitude, als auch energetisch in dB, sowie auch eine Darstellung der Energie über die Frequenz.

IR_plots Plottet die Impulsantwort der Mikrofone und des B-Formats bei $\varphi = 42^\circ$ und $\vartheta = -50^\circ$ um diese als Beispiel verwenden zu können. Einmal als Amplitudenverlauf der einzelnen Mikrofone, sowie des B-Formates (linear und tiefpassgefiltert) und als Energieverlauf in dB (allerdings nur den w -Kanal).

Rauschspektrum Liest die Rauschaufnahmen ein und plottet diese. Dabei erfolgt die Darstellung sowohl in Amplitudenspektrum als auch Energiespektrum in dB. Um die

Ansicht übersichtlicher zu gestalten wurden auch Diagramme geplottet, in welchen die Daten über mehrere Frequenzpins gemittelt wurden.

A.1.5. Sonstige

Um die verschiedenen Ordnerstrukturen auf verschiedenen Betriebssystemen dynamischer zu gestalten (damit das Skript nicht jedes Mal erneut angepasst werden musste), wurde die Funktion *SelDir* geschrieben, welche im Wesentlichen eine Parametereingabe in MATLAB veranlasst (angenehm ist hier die MATLAB-Funktion *pwd*, welche das gegenwärtige Verzeichnis angibt).

SelDir_ (*function*) Abfrage des Zielverzeichnisses, um die verschiedenen Ordnerstrukturen auf mehreren Rechnern dynamischer handhaben zu können.

Bez (*function*) Bezeichnungsfunktion, erstellt aufgrund von Eingangsdaten (Frequenzauflösung, Frequenzbereich usw.) einen String, der als Bezeichnung für die Diagramme verwendet wird (teilweise auch bei der Datenbenennung).