

Messung und Analyse von zeit- und ortsabhängigen Raumimpulsantworten

Diplomarbeit

von

Gerhard Thallinger

Matr.Nr.: 9431086
Geburtsdatum: 26.03.1975

durchgeführt am
Institut für Nachrichtentechnik und Wellenausbreitung
der Technischen Universität Graz
in Zusammenarbeit mit dem Institut
für Elektronische Musik
der

Universität für Musik und darstellende Kunst Graz

im Rahmen des interuniversitären Diplomstudiums
Elektrotechnik-Toningenieur

Graz, März 2002

Begutachter: Univ.-Prof. Dr. techn. DI Gernot Kubin
Univ.-Prof. Dr. techn. DI Mag. Robert Höldrich

Betreuer: Univ.-Ass. DI Christian Feldbauer
Univ.-Ass. DI Alois Sontacchi

Kurzfassung

Diese Arbeit beschäftigt sich mit der Messung und Analyse von Raumimpulsantworten ("room impulse responses, RIR"). Es wird derjenige Ausschnitt im Zeit- und Frequenzbereich ermittelt, in dem man für die akustische Übertragungsfunktion reproduzierbare Messergebnisse erhält. Als Bewertungsmaß wird hierfür das SDR (signal-to-deviation ratio) herangezogen. Es wird gezeigt, dass die Temperatur eine zentrale Einflussgröße im Bezug auf die RIR darstellt, da schon eine Temperaturänderung im Zehntelgradbereich eine massive Änderung der RIR bewirkt.

Weiters wird die RIR für langsam veränderliche Räume (beispielsweise ein sich langsam öffnendes Fenster) gemessen und analysiert.

Die Messung der zeitvarianten Übertragungsfunktion wird auch mit einem Mikrofonarray vorgenommen, um einen Datensatz zu generieren, der für zukünftige Anwendungen genutzt werden kann. Dabei ist auf die Reproduzierbarkeit der Messung zu achten, um später eine genau Aussage bezüglich des Zusammenhangs der Impulsantwort mit der Veränderung im Raum erstellen zu können.

Anhand dieser Daten wird untersucht, ob der Verlauf der zeitvarianten RIR durch lineare Prädiktion vorhergesagt werden kann. Das Ergebnis der Prädiktion liefert die Aussage, dass zeitvariante RIRs in einem räumlichen Horizont von ca. 0,05 Wellenlängen prädiziert werden können.

Abstract

The task of this work was the measurement and analysis of room impulse responses (RIR). The range in the time and frequency domain was determined over which reproducible measurement results are obtained for the acoustic transfer function.

The SDR (signal-to-deviation ratio) is used as the yardstick to evaluate the measurements.

The temperature represents a central influence factor with respect to the RIR, because a temperature change within a few tenths of a degree causes a substantial modification of the RIR.

The RIR for slowly variable spaces (for example slowly opening of a window) is measured and analysed.

Furthermore, a time-varying transfer function was measured with a microphone array to generate a data record which can be used for future applications. It was necessary to pay attention to the reproducibility of the array measurement, in order to be able to establish the correspondence between the impulse response and the modification of the room.

On the basis of these data, it was analysed whether the process of the time-varying RIR can be linearly predicted. The result of the prediction experiment allows to say that time-varying RIRs can be predicted over a spatial horizon of approximately 0,05 wavelengths.

Danksagung

An dieser Stelle möchte ich mich bei all jenen bedanken, die mir bei der Erstellung dieser Arbeit behilflich waren.

Dies gilt an erster Stelle für die Herren Univ.-Prof. Dr. techn. DI Gernot Kubin und Univ.-Prof. Dr. techn. DI Mag. Robert Höldrich, die mit Geschick den Verlauf dieser Arbeit zu lenken wussten und durch ihre große Kompetenz für mich immer ein zusätzlicher Ansporn waren.

Weiters möchte ich mich bei den Herren Univ.-Ass. DI Christian Feldbauer und Univ.-Ass. DI Alois Sontacchi bedanken, welche mir als Betreuer zur Verfügung standen.

Auch der weiteren Belegschaft des INWs und IEMs, die mich durch unzählige Hilfestellungen unterstützte, möchte ich meinen Dank aussprechen.

Besonders bedanken möchte ich mich noch bei meiner Familie, die mir dieses Studium ermöglichte, und meiner Freundin Silvia, die mich während dieser Zeit begleitete.

Herzlichen Dank!

Einleitung

Diese Arbeit beschäftigt sich mit der Messung, Analyse und Prädiktion von akustischen Raumimpulsantworten.

Unter Raumimpulsantwort (auch einfach nur Impulsantwort oder als Abkürzung „RIR“ (room impulse response)) versteht man das Anregungssignal eines Systems, nämlich der Kaskadierung von Lautsprecher – Raum – Mikrofon wenn dieses mit einem Impuls angeregt wird.

Es wird untersucht, wie genau sich Impulsantworten messen lassen, sprich, wo der reproduzierbare Bereich im Zeit- und Frequenzbereich liegt und welche Fehlerquellen identifiziert bzw. beseitigt werden können.

Die Messungen beziehen sich sowohl auf stationäre aber auch auf nicht stationäre Verhältnisse.

Im stationären Fall wird einfach in einem Raum gemessen, in dem keine geometrischen Veränderungen passieren, in dem keine Personen sind und sich nichts bewegt.

Im nicht stationären Fall wird ein Gegenstand mit konstanter Geschwindigkeit linear im Raum bewegt. Anhand der Messergebnisse des nicht stationären Falles soll untersucht werden, in wie weit es möglich ist, Impulsantworten vorherzusagen.

In Kapitel 1 wird ein kurzer Überblick über die Grundlagen der Raumakustik gegeben. Es wird die klassische Einteilung von wellentheoretischer, geometrischer und statistischer Raumakustik vorgenommen und erklärt.

Kapitel 2 beschäftigt sich mit der Darstellung von zeitkontinuierlichen und zeitdiskreten Signalen. Es werden Übertragungssysteme und deren Eigenschaften behandelt, um die akustische Übertragungsfunktion einer Klasse zuordnen zu können.

Weiters werden Möglichkeiten zur Bestimmung der Übertragungsfunktion gezeigt.

Auf die bei der Messung verwendete MLS-Methode wird dabei näher eingegangen. Es wird geprüft, ob und in welcher Form sich RIRs in die Klasse der LTI-Systeme einordnen lassen.

Kapitel 3 zeigt einige Anwendungsbeispiele, die sich durch die exakte Kenntnis der Raumimpulsantwort ergeben.

Kapitel 4 umfasst den messtechnischen Teil dieser Diplomarbeit, wobei zwischen stationären und nicht stationären Messungen unterschieden werden kann. Auf die Messung des Arraydatensatzes wird dabei detaillierter eingegangen.

In Kapitel 5 werden die Messergebnisse der stationären Messungen ausgewertet. Hier wird besonderes Augenmerk auf die Reproduzierbarkeit von Raumimpulsantworten gelegt.

Im letzten Teil dieser Arbeit wird die Kompaktheit des Arraydatensatzes untersucht. Weiters sollen die Veränderungen der Raumimpulsantwort aufgrund eines nicht stationären Vorganges im Raum (bewegtes Objekt) analysiert werden. Hierzu dient unter anderem eine 4 Quadranten-Darstellung der zeitvarianten Impulsantwortdaten.

Abschließend soll noch untersucht werden, ob und inwieweit sich Raumimpulsantworten mittels linearer Prädiktion vorhersagen lassen.

Inhaltsverzeichnis

Einleitung	1
1 Grundlagen der Raumakustik	3
1.1 Einleitung	3
1.2 Teilgebiete der Raumakustik	4
1.2.1 Wellentheoretische Raumakustik	5
1.2.2 Geometrische Raumakustik	8
1.2.3 Statistische Raumakustik	10
1.2.3.1 Schallabsorptionsgrad	11
1.2.3.2 Nachhallzeit	12
1.2.3.3 Messung der Nachhallzeit	13
1.2.3.4 Hallradius	14

2	Systemtheoret. Betrachtung der Impulsantwort	16
2.1	Beschreibung deterministischer Signale	16
2.1.1	Darstellung im Zeitbereich	16
2.1.2	Fouriertransformation	17
2.1.2.1	Kontinuierliche Darstellung	17
2.1.2.2	Zeitdiskrete Darstellung	17
2.1.3	Faltung	17
2.1.3.1	Kommutativität	17
2.1.4	Korrelation	18
2.1.4.1	Berechnung der Korrelationsfunktion über Fourier-Transformation	18
2.2	Impulsantwort und Übertragungsfunktion	19
2.2.1	LTI-Systeme	19
2.2.1.1	Linearität	19
2.2.1.2	Zeitinvarianz	20
2.2.2	Übertragungsfunktion eines LTI-Systems	20
2.2.3	Akustische Übertragungsfunktion des Raumes	21
2.2.4	Bestimmung einer Übertragungsfunktion	23
2.2.4.1	Impuls als Anregungssignal	23
2.2.4.2	Harmonisches Anregungssignal	23
2.2.4.3	Breitbandiges Anregungssignal	24
2.2.4.4	Breitbandiges, deterministisches Anregungssignal	24
2.3	MLS-Verfahren zur Bestimmung der Raumimpulsantwort	25
2.3.1	Herleitung für Berechnung der Impulsantwort	27
2.3.2	Impulsantwortberechnung im Frequenzbereich	29
2.3.2.1	Unterschied linearer zu zirkulärer Faltung	29
2.3.2.2	Realisierung der zeitdiskreten Korrelationsberechnung	30
3	Anwendungsbereiche der Raumimpulsantwort	33
3.1	Echokompensation	33
3.1.1	Echokompensation durch Modellieren der Raumübertragungsfunktion	37
3.1.2	Übertragungsbereiche	38
3.1.3	SDR (Signal to Deviation Ratio)	38
3.1.3.1	Beschreibung der Signalqualität	39
3.1.3.2	Definition des SDR bzw. SDRF	39
3.2	Anwendung in der Aufnahmetechnik	41
3.2.1	Entzerren des Aufnahmebereiches	41
3.2.2	Entzerren des Wiedergabebereiches	42
3.2.3	Entzerren des Aufnahme- und Wiedergabebereiches	43
3.3	Verwendung als Hallprogramm	43

4	Messtechnik	44
4.1	Einteilung der Raumverhältnisse	44
4.1.1	Stationäre Verhältnisse	44
4.1.2	Nichtstationäre Verhältnisse	45
4.2	Allgemeiner Überblick der Messungen	46
4.3	Messgeräte	48
4.3.1	Studiogeräte	48
4.3.2	Messschlitten	48
4.4	Durchgeführte Messungen	50
4.4.1	Stationäre Messungen	50
4.4.1.1	Langzeitmessung	50
4.4.1.2	Unterschiedliche Abstände	50
4.4.1.3	Verschiedene Pegel	50
4.4.1.4	Sprachsignal als Anregung	51
4.4.2	Nichtstationäre Messungen	51
4.4.2.1	Bewegte Schachtel mit fixpositioniertem Mikrofon	51
4.4.2.2	Bewegte Schachtel im unveränderlichen Raum	52
4.4.2.3	Bewegte Schachtel mit Array aufgenommen	52
4.5	Beschreibung des Mikrofonarrays	53
4.5.1	Kontinuierliche Apertur	53
4.5.2	Diskrete Apertur	56
4.5.3	Beamformer	58
4.6	Planung des Mikrofonarrays	59
4.6.1	Stationäre Raumbedingung	61
4.6.2	Verstärkungskompensation	62
4.6.3	Kompensation der verschiedenen Mikrofonfrequenzgänge	62
4.6.4	Kompensation der Geschwindigkeitsfehlers	62
4.6.5	Protokollierung der Temperaturveränderung	63
4.7	Ablauf der Arraymessungen	64
4.7.1	Zeitplan der Teilmessung	64
4.7.2	Ablauf zwischen den Teilmessungen	66
5	Verfahren zur Auswertung der Messdaten	67
5.1	Darstellung der Raumimpulsantwort	67
5.1.1	Zeitbereich	67
5.1.2	Frequenzbereich	67
5.2	Definition der Raumimpulsantwort-Matrix M	69
5.3	Einige Betrachtungen zum SDR	69
5.3.1	SDR angewandt auf verschiedene Bereiche der Impulsantwort	69
5.3.2	SDR bei Abschneiden beider Raumimpulsantworten	70

5.3.3	SDR bei Abschneiden einer Raumimpulsantworten	71
5.3.4	Mittelung von Raumimpulsantworten	72
5.3.5	SDR in Abhängigkeit des Abstandes	73
5.3.6	SDR in Abhängigkeit des Beschallungspegels	73
5.4	Impulsantwort im Frequenzbereich	74
5.5	Impulsantwort in Zeit- und Frequenzbereich	74
5.6	Kenngrößen von Raumimpulsantworten	75
5.6.1	Nachhallzeit	75
5.6.2	Nulldurchgangsrate	77
5.6.3	Untersuchung zum Hallradius	77
5.7	Entstehen von Kammfiltereffekten	79
5.8	Temperatureinfluss	83
5.8.1	Simulierter Temperaturerhöhung	84
5.8.2	Reale Temperaturerhöhung	85
5.8.3	Erwärmungseffekt des Lautsprechers	87
5.9	Impulsantwortschätzung aus Sprachdaten	88
6	Messergebnisse bei nichtstationären Raumverhältnissen	91
6.1	Untersuchung zur Homogenität des Arraydatensatzes	91
6.1.1	Kompensation der Mikrofonverstärkung	91
6.1.2	Kompensation der Mikrofonfrequenzgänge	92
6.1.3	Temperaturkompensation	94
6.1.4	Geschwindigkeitskompensation	95
6.2	Vier-Quadranten-Darstellung	96
6.3	Darstellung einzelner Vektoren der Matrix M	99
6.3.1	Orts-Zeit-Darstellung	99
6.3.2	Orts-, Frequenzdarstellung	101
6.3.3	Ortskurven	102
6.4	Grafische Vier-Quadranten-Darstellung der Matrix M	103
6.5	Synthetischer Betragsfrequenzzeiger	108
6.6	Lineare Prädiktion (LP)	113
6.7	Prädiktion der RIR im nichtstationären Fall	117
6.7.1	Örtliche Prädiktion von Frequenzzeigern	117
6.7.2	Prädiktion im Zeitbereich	123
Anhang A		125
A1	Grundriss CUBE	125
A2	Aufriss CUBE	126

A3	Detailaufstellung der Arraymessung	126
A4	Höhe der Mikrofone	127
A5	Abmessung der Lautsprecheraufstellung	127
A6	Abmessungen des bewegten Hindernisses	128
A7	Fotos der Messungen im CUBE	128
Anhang B		131
B1	x-Positionen des Arrays	131
B2	Weg-, Zeitdaten Messschlitten 1	132
B3	Koordinaten sämtlicher Arraysensorpositionen	133
Anhang C		134
C1	Frequenzgang Studiolautsprecher Genelec 1032 A, No. 780	134
C2	Datenblatt Genelec	135
C3	Mikrofonvorverstärker	136
C4	Verwendete Mikrofone	136
C5	Technische Daten der Arraymikrofone	137
C6	Technische Daten des AD/DA Konverters	137
Quellenverzeichnis		138

Grundlagen der Raumakustik

1.3 Einleitung

Unter Schall versteht man mechanische Schwingungen eines elastischen Mediums. Damit Schallausbreitung möglich wird, ist ein Übertragungsmedium nötig. Schall breitet sich in gasförmigen Medien vorzugsweise als Longitudinalwelle, mit der Schallgeschwindigkeit c aus. Wir betrachten im Folgenden nur die Schallausbreitung in Luft. Dabei werden nur abgeschlossene Bereiche (Räume) betrachtet.

Schallwellen werden in 3 Bereiche eingeteilt:

- Infraschall (unter 16 Hz)
- Hörbereich (16 Hz - 20 kHz)
- Ultraschall (über 20 kHz)

Dabei interessiert uns in dieser Arbeit nur der Hörbereich. Infraschall und Ultraschall werden beispielsweise in der Bauakustik und in medizinischen Anwendungen behandelt.

Die Schallgeschwindigkeit c in Luft ist temperaturabhängig und wird wie folgt angegeben:

$$c(\vartheta) \approx 331,3 + 0,6 \cdot \vartheta \quad [m/s] \quad (1.1)$$

ϑ ... Lufttemperatur in °C

Die Schallgeschwindigkeit ist jedoch nicht zu verwechseln mit der Schallschnelle v , welche die Bewegung der Luftmoleküle angibt, uns jedoch nicht weiter beschäftigen soll.

Die Schallgeschwindigkeit c beträgt also bei 20°C Raumtemperatur 343,3 m/s. Die Wellenlänge λ ist umgekehrt proportional der Frequenz.

$$\lambda = \frac{c}{f} \quad (1.2)$$

Das Gehör ist somit in der Lage, Wellenlängen λ von ca. 1,7 cm bei 20 kHz bis 21 m bei 16 Hz wahrzunehmen.

Da die Schallwellen im Hörbereich einen so großen Wellenlängenbereich ausfüllen, ist auch keine pauschale Beschreibung bei der Schallausbreitung möglich. Effekte der Wellenausbreitung wie Absorption, Beugung, Abschattung und Stehwellen müssen in den jeweiligen Frequenzbereichen verschieden betrachtet werden.

Ein Raum umfasst beispielsweise Objekte mit sämtlichen Abmessungen (z.B. Kugelschreiber, Flasche, Kasten, Raumlänge), die der Wellenlängenbereich abdeckt. Die verschiedenen Einrichtungsgegenstände wirken also im jeweiligen Frequenzbereich unterschiedlich auf die Schallwellen.

Die Schallwellenausbreitung in Luft erfolgt über die Luftdruckänderung des Medium. Dabei wird als Schalldruck nur der Wechseldruck verstanden, welcher im Vergleich zum atmosphärischem Luftdruck sehr gering ist. Der Schalldruck p ist ein Skalar, also eine ungerichtete Größe. Sie dient zur Beschreibung des Schallfeldes als zentrale Größe. Dabei entsprechen $p_0 = 20\mu Pa$ Wechseldruck der Ruhehörschwelle (wird in der logarithmischen Skala mit 0 dB definiert) und $p_s = 100 Pa$ der Schmerzgrenze des Gehörs. Es ergibt sich also ein Dynamikumfang von 134 dB des Gehörs. Diesem großen Umfang ist man jedoch mit Ausnahme von extremem Lärm nicht ausgesetzt. Das Gehör wäre auch gar nicht in der Lage, einen so großen Dynamikbereich gleichzeitig zu verarbeiten. Erfolgt beispielsweise ein Knall mit 100 dB, empfindet das Gehör einen darauffolgenden Geräuschpegel mit 40 dB als völlige Ruhe, welches durch die zeitliche Verdeckung erklärt wird [Lit. 1]. Nach einer kurzen Erholungszeit (ca. 100 ms) wird daraufhin das Grundgeräusch wieder wahrgenommen. Aufgrund dieses beschränkten simultanen Dynamikumfangs von ca. 60 dB werden Störgeräusche, die sich gerade unterhalb dieses Dynamikfensters befinden, auch nicht wahrgenommen. Dieser Dynamikumfang von 60 dB dient auch zum Definieren der Nachhallzeit, was später noch gezeigt wird.

1.4 Teilgebiete der Raumakustik

Die Raumakustik kann in drei Teilbereiche gegliedert werden, die sich in der Praxis bereits bewährt haben:

- Wellentheoretische Raumakustik
- Geometrische Raumakustik
- Statistische Raumakustik

Es ist zu erwähnen, dass keiner diese drei Bereiche ausreicht, das Schallfeld eines Raumes komplett zu beschreiben. Es können lediglich Teile der zugehörigen Raumimpulsantwort damit erklärt werden, oder für bestimmte Anwendungen kompakt mit einer Zahl ausgedrückt werden.

In Abb. 1.1 ist nun die prinzipielle Struktur einer Raumimpulsantwort dargestellt. Dabei können 3 Bereiche unterschieden werden:

- Direktsignal
- 1. Reflexionen
- diffuser Nachhall

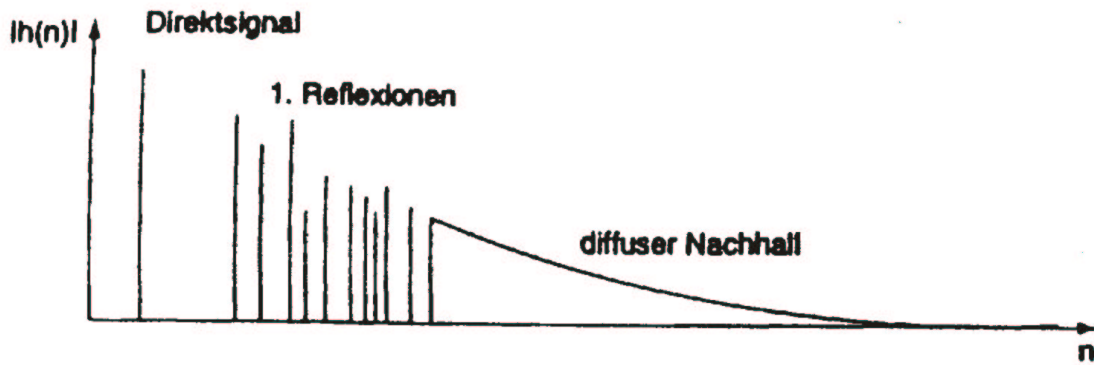


Abb. 1.1: Einteilung der Raumimpulsantwort $h(n)$ in Bereiche [Lit. 2]

Zur besseren grafischen Darstellung ist dabei in Abb. 1.1 der Betrag der RIR zu sehen, n bezeichnet den zeitdiskreten Zeitindex. Findet eine Schallausbreitung in einem Raum statt, treten zusätzlich zum Direktschall Reflexionen an den Wänden auf. Im vorderen Teil der RIR, können die einzelnen Reflexionen noch zeitlich voneinander unterschieden werden. Im hinteren Teil der Impulsantwort, kann die Unterscheidung der Impulse jedoch aufgrund der Vielzahl der Einzelreflexionen, welche sich überlagern, nicht mehr erfolgen.

Dabei kann der Teil erster Reflexionen und der Direktschall in Abb. 1.1 der geometrischen Raumakustik zugeordnet werden.

Der diffuse Nachhallanteil wird der statistischen Raumakustik zugeordnet.

Die wellentheoretische Raumakustik liefert Aussagen über Frequenzgang eines Raumes. Anhand dieser Betrachtung können starke Resonanzerscheinungen von Räumen erklärt werden.

1.4.1 Wellentheoretische Raumakustik

Als Grundlage der Wellentheoretischen Raumakustik werden die Wellengleichungen herangezogen. Sie sind jedoch nur für einfache Raumformen lösbar. Diese Form der Raumakustik wird vor allem im tiefen Frequenzbereich angewandt, wo die Wellenlänge ungefähr im Bereich der Raumabmessungen liegt.

Ausgehend von der Wellengleichung für das Geschwindigkeitspotential [Lit. 3], lassen sich Schalldruck und Schallschnelle wie folgt berechnen:

$$\Delta\phi = \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} \tag{1.3}$$

ϕ Geschwindigkeitspotential

$$v = -\text{grad}\phi \quad (1.4)$$

v Schallschnelle

$$p = \frac{d\phi}{dt} \quad (1.5)$$

p Schalldruck

Es werden folgende vereinfachende Annahmen getroffen:

- Schallharte Wände (Schnelle $v = 0$)
- Harmonische Anregung des Raumes (zeitliche Ableitungen ergeben: $j\omega, -\omega^2$)

$$\Delta\phi = -\frac{\omega^2}{c^2}\phi(\omega) = -k^2\phi \quad (1.6)$$

mit $k = \frac{\omega}{c} = \frac{2\pi f}{c}$ Wellenzahl

Die Differenzialgleichung (1.6) lässt sich ausschließlich für einfach geformte Räume geschlossen lösen. Bei Annahme eines Quaders mit $l_x > l_y > l_z$ lässt sich folgender Ansatz anschreiben:

$$\phi = \hat{\phi} \cdot \cos\left(\frac{n_x \cdot \pi \cdot x}{l_x}\right) \cdot \cos\left(\frac{n_y \cdot \pi \cdot y}{l_y}\right) \cdot \cos\left(\frac{n_z \cdot \pi \cdot z}{l_z}\right) \cdot e^{j\omega t} \quad (1.7)$$

Mit dem Ansatz aus (1.7) kann nun die Wellengleichung für diskrete Frequenzen f_R gelöst werden.

Die diskreten Frequenzen f_R beschreiben die Eigenschwingungen bzw. Resonanzen des Raumes.

$$f_R^2 = \left(\frac{c}{2}\right)^2 \left[\left(\frac{n_x}{l_x}\right)^2 + \left(\frac{n_y}{l_y}\right)^2 + \left(\frac{n_z}{l_z}\right)^2 \right] \quad (1.8)$$

n_x, n_y, n_z natürliche Zahlen

Mit jedem Wertetripel n_x, n_y, n_z wird die Eigenschwingung eines Raumes charakterisiert, deren Frequenz durch (1.8) gegeben ist. Die Ordnungszahlen n_x, n_y, n_z werden für die Angabe der Zahl der Knotenebenen des Schalldrucks in der jeweiligen Ebene verwendet.

Um sich einen genauen Überblick über die Menge der Eigenschwingungen und die dazugehörigen Eigenfrequenzen zu verschaffen, werden die Gitterpunkte mit den Koordinaten in einem kartesischen Koordinatensystem mit den Achsen f_x, f_y, f_z eingezeichnet. Dabei stellt in diesem Frequenzraum jeder Gitterpunkt eine Eigenschwingung dar. Die Frequenz wird durch den Abstand des Gitterpunktes vom Koordinatenursprung angegeben.

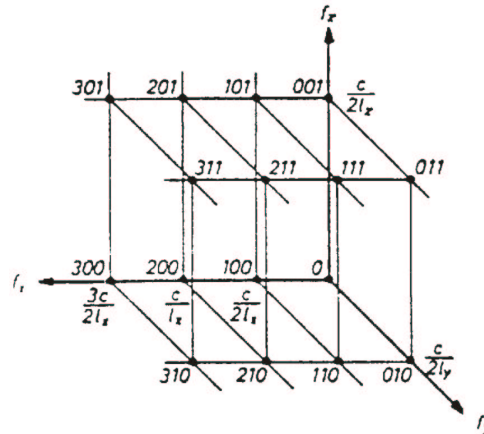


Abb. 1.2: Darstellung der Eigenschwingungen eines Quaders [Lit. 3]

Werden nun sämtliche Eigenfrequenzvektoren bis zu einer bestimmten Grenzfrequenz f_{GR} aufgezählt, so erhält man die Anzahl der verschiedenen Resonanzfrequenzen eines Raumes.

$$N = \frac{4\pi f_{GR}^3 V}{3c^3} + \frac{\pi f_{GR}^2 A}{4c^2} + \frac{f_{GR} L}{8c} \quad (1.9)$$

- f_{GR} Grenzfrequenz
- V Raumvolumen
- A gesamte Wandfläche
- L Summe der Kantenlängen

Prinzipiell ist zu sagen, dass Resonanzfrequenzen in kleineren Räumen besser voneinander getrennt werden können, da sich diese isolierter voneinander ausbilden. Bei großen Räumen ist bereits bei tieferen Frequenzen eine so große Resonanzfrequenzdichte vorhanden, dass man keine einzelnen Resonanzfrequenzen hören kann.

Je höher der betrachtete Frequenzbereich eines Raumes liegt, desto größer wird die Resonanzfrequenzdichte.

Beispiel: Aula der TU-Graz: zwischen 1000 und 1001 Hz: 510 Eigenfrequenzen

1.4.2 Geometrische Raumakustik

Dieser Teilbereich (auch „optische Raumakustik“ genannt) geht von der strahlenförmigen Schallausbreitung aus (Schallstrahl), d.h. die Schallausbreitung wird als rein geometrischer Vorgang betrachtet.

Im ersten Teil der Schallausbreitung kann zweckmäßig von der statistischen Betrachtung des Schalls abgesehen werden. Es wird das Modell der geometrischen Raumakustik verwendet.

Sie findet hauptsächlich im Bereich Direktschall und erste Reflexionen Anwendung.

Die Schallabstrahlung einer akustischen Quelle erfolgt durch Überlagerung der Abstrahlcharakteristik und der Art der Welle (Ebene-, Kugelwelle).

Im folgenden gehen wir von einer Kugelwelle einer akustischen Quelle mit gleichmäßiger Abstrahlcharakteristik aus, welche für tiefe Frequenzen gilt. Bei höheren Frequenzen kommt es i.a. zur Keulenausbildung, die aber hier vernachlässigt werden soll. Die Schallausbreitung erfolgt klarerweise dreidimensional, soll aber in Abb. 1.3 wegen der einfacheren Darstellung auf den zweidimensionalen Fall reduziert werden.

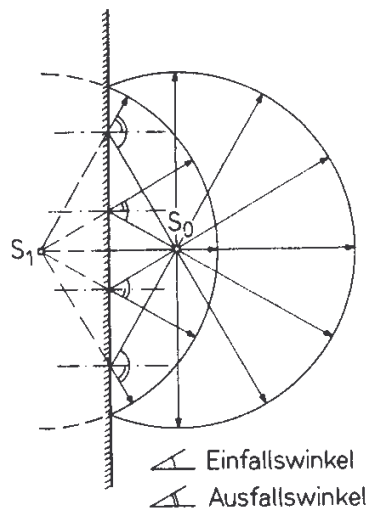


Abb. 1.3: Spiegelquelle 1. Ordnung und Wellenfront [Lit. 4]

Abb. 1.3 zeigt also für die 2 dimensionale Darstellung die kompletten Wellenfronten. Dabei bezeichnet S_0 die akustische Quelle und S_1 die dazugehörige Spiegelquelle. Anhand der Spiegelquelle S_1 lässt sich eine reflektierte Schallfront auf einfache Weise darstellen.

Will man die ersten Reflexionen einer RIR interpretieren, ist es sinnvoll, nur diejenigen Schallstrahlen zu zeichnen, die auch wirklich den Punkt treffen, für den die RIR betrachtet wird. Es wird somit die Wellenfront auf einen Schallstrahl reduziert.

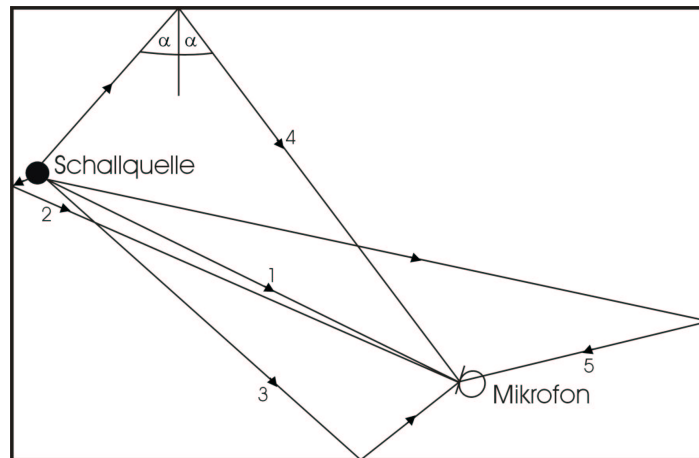


Abb. 1.4: Schallstrahlenverlauf bis zu den Reflexionen 1. Ordnung

In Abb. 1.4 können nun einzelne Schallstrahl verfolgt werden, welche nach der ersten Reflexion das Mikrofon treffen. Dabei gilt das Gesetz: Einfallswinkel = Ausfallswinkel. Die Nummerierung der Schallstrahlen erfolgt dabei nach der Reihe des Eintreffens der Schallstrahlen beim Mikrofon. Die einzelnen Wellenfronten einer Impulsantwort können später anhand der Laufzeit speziellen Reflexionen (Wände, Decke, Fußboden) zugeordnet werden.

Kommt die Rauigkeit einer Oberfläche jedoch in die Größenordnung der Wellenlänge, wirkt eine solche Wand eher als Diffusor, denn als idealer Reflektor. Bei jeder Reflexion findet außerdem eine Filterung und Absorption der Schallfronten statt. Da die Absorptionsgrade und die Filterfunktionen der verschiedenen Oberflächen (z.B. Holzdecke, Tapetenwand, Teppichboden, Fenster) frequenzabhängig und im allgemeinen unterschiedlich sind, ist die mathematische Beschreibung des Strahlverlaufs über viele Reflexionen nicht wirklich möglich. Weitere erschwerende Effekte, die eine exakte mathematische Beschreibung verhindern, sind Abschattung und Beugung von Schallwellen. Unter letzterem versteht man, dass sich Schallwellen um ein Hindernis beugen, d.h. auch jenen Bereich bis zu einem gewissen Grad akustisch ausfüllen, der von einem Hindernis abgeschattet wird. Beugung findet für Frequenzen statt, deren Wellenlänge ca. gleich groß bzw. größer als das Hindernis ist (Abb. 1.5 – oberstes Hindernis).

Ist die Wellenlänge kleiner als die Hindernisgröße, so kommt es zur Abschattung. Diese Effekte könnte man beispielsweise in einem Konzertsaal hinter einer Säule beobachten. Der Tieftonbereich ist unverändert zu hören, der Mitteltonbereich wird leicht bedämpft und der Hochtonbereich ist kaum mehr wahrzunehmen (siehe Abb. 1.5).

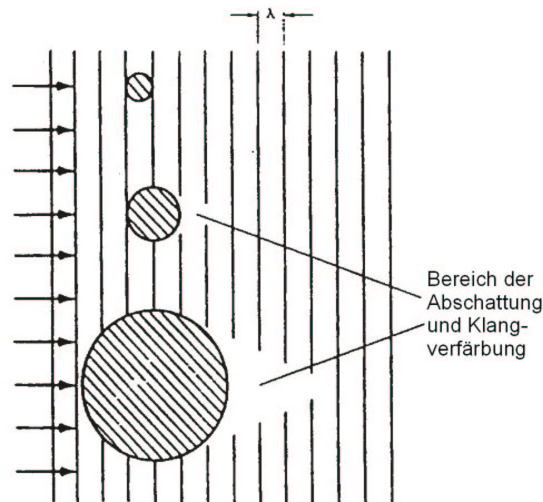


Abb. 1.5: Beugung an verschiedenen Hindernisgrößen [Lit. 5]

1.4.3 Statistische Raumakustik

Aufgrund der hohen Dichte der Eigenfrequenzen bei mittleren und großen Räumen sinkt die Aussagekraft der wellentheoretischen Raumakustik. Eine weitere Komplikation liegt darin, dass Räume in der Praxis keine einfachen geometrischen Formen besitzen. Die einfache Annahme eines quaderförmigen Raumes wird durch seine Einrichtungsgegenstände für höhere Frequenzen erheblich gestört.

Die geometrische Raumakustik ist nur für den ersten Teil der RIR geeignet. Mit steigender Anzahl der Reflexionen wird das mathematische Modell einfach zu komplex. Es erfolgt somit der Übergang zur statistischen Raumakustik, da über das Schallfeld nur noch pauschale Aussagen gemacht werden können.

Dieses Teilgebiet der Raumakustik hat vor allem die Aufgabe, den Zeitverlauf von Schallfeldparametern zu untersuchen.

Folgende Größen werden dabei genauer betrachtet:

- Leistungsbilanz
- An-/Abklingen im Raum
- Stationärer Teil der Energiedichte

Dabei tritt die geometrische Form des Raumes in den Hintergrund. Relevante Größen, die zur Berechnung herangezogen werden, sind das Raumvolumen V , die Oberfläche A und der Schallabsorptionsgrad α .

Schallabsorptionsgrad

$$\text{Schallabsorptionsgrad } \alpha = \frac{\text{absorbierte Schalleistung}}{\text{auftretende Schalleistung}} \quad (1.10)$$

Der Reflexionsgrad sei der Vollständigkeit halber erwähnt

$$\text{Reflexionsgrad } \rho = \frac{\text{reflektierte Schalleistung}}{\text{auftreffende Schalleistung}} \quad (1.11)$$

Wobei folgende Beziehung gilt:

$$\alpha + \rho = 1$$

Gesamtschallabsorptionsgrad:

$$\alpha_{ges} = \frac{1}{A} \sum_n \alpha_n \cdot A_n \quad (1.12)$$

A_n ... Teilflächen

Aus der Summe der Teilflächen erhält man somit die Gesamtfläche:

$$A = \sum_n A_n \quad (1.13)$$

Durch die Absorption der Wände, sinkt nach dem Ausschalten der Schallquelle die Gesamtenergiedichte im Raum mit zunehmender Zeit. Diese Energieabnahme kann anhand einer abklingenden Exponentialfunktion modelliert werden.

Zusätzlich zur Absorption an Wänden, dämpft auch das Medium Luft die Schallwellen. Die Schallabsorption in Luft ist aber im Vergleich zur Dämpfung an Wänden vor allem im tieffrequenten Bereich sehr gering.

Die Schallabsorption in Luft beträgt:

Frequenz	(-1dB Dämpfung)
1 kHz	5000 m
10 kHz	50 m
100 kHz	0,5 m

Tab. 1.1: Schallabsorption in Luft [Elektroakustik AK]

Nachhallzeit

Die, im vorigen Punkt angegebene Schallabsorption in Luft, sei für die Herleitung der Nachhallzeit vernachlässigt. Es wird also angenommen, dass nur durch Reflexion an Wänden und sonstigen Einrichtungsgegenständen, Schallenergie abgebaut wird.

Ein Wellenpaket mit der Anfangsenergie ΔW , besitzt also nach der ersten Reflexion nur noch die Energie $\Delta W(1-\alpha)$. Nach n Reflexionen beträgt die Energie des Wellenpakets demnach $\Delta W \cdot (1-\alpha)^n$. Um nun eine Aussage über die durchschnittliche Anzahl der Reflexionen im Raum zu bekommen, wird eine mittlere freie Weglänge l_m [Lit. 6] berechnet.

$$l_m = \frac{4V}{A} \quad (1.14)$$

Es ergibt sich die mittlere Stoßzahl n :

$$n = \frac{A \cdot c}{4V} \quad \left[\frac{1}{s} \right] \quad (1.15)$$

Daraus folgt für die Energie nach $n \cdot t$ Reflexionen:

$$\Delta W \cdot (1-\alpha)^{\frac{A \cdot c \cdot t}{4V}} \quad (1.16)$$

A Gesamtoberfläche des Raumes

V Raumvolumen

Wird nun die zur Zeit $t = 0$ vorhandene Energiedichte mit E_0 angenommen, so ergibt sich für die Schallenergiedichte folgende Zeitabhängigkeit:

$$E(t) = E_0 \cdot (1-\alpha)^{\frac{A \cdot c \cdot t}{4V}} = E_0 e^{\frac{A \cdot c \cdot \ln(1-\alpha) \cdot t}{4V}} \quad (1.17)$$

Aufgrund des logarithmischen Lautstärkeempfindens des Gehörs, wird die exponentielle Abnahme der Energiedichte als lineare Abnahme der Lautstärke empfunden.

Die Nachhallzeit entspricht nun jener Zeitdauer, in der die Energie auf den millionsten bzw. der Schalldruck auf den tausendsten Teil abklingt, was einer Abnahme um 60 dB entspricht.

$$\frac{E}{E_0} = 10^{-6} = e^{\frac{A \cdot c \cdot \ln(1-\alpha) \cdot t}{4V}} \quad (1.18)$$

Durch Umformung erhält man:

$$t = T_N = \frac{\ln(10^{-6}) \cdot 4V}{A \cdot c \cdot \ln(1-\alpha)} = 13,82 \cdot \frac{4V}{-A \cdot c \cdot \ln(1-\alpha)} \quad (1.19)$$

Wird die Schallgeschwindigkeit mit $c = 344 \text{ m/s}$ in (1.19) eingesetzt, erhält man die Nachhallzeit nach Eyring:

$$T_N = 0,161 \cdot \left[\frac{s}{m} \right] \frac{V}{-A \cdot \ln(1-\alpha)} \quad (1.20)$$

Für kleine und mittlere Schallabsorptionsgrade kann $\ln(1-\alpha)$ durch $-\alpha$ angenähert werden. Dadurch ergibt sich die Nachhallzeit nach Sabine:

$$T_N = 0,161 \left[\frac{s}{m} \right] \cdot \frac{V}{\alpha A} \quad (1.21)$$

Messung der Nachhallzeit

Die Nachhallzeit wird nun an dem in einem Diagramm dargestellten Abklingverhalten der Schallenergie abgelesen. Nachdem die Abklingkurve wie gesagt im logarithmischen Maßstab linear ist, kann dadurch die Nachhallzeit durch die zweifache Zeitdauer eines 30 dB Abfalls ermittelt werden. Dies hat den Vorteil, dass die, zur Ermittlung der Nachhallzeit benötigte Schallenergie, nicht so hoch sein muss (hat der Grundgeräuschpegel eines Raumes bezogen auf die Ruheschwelle 50 dB, müsste der Raum mit mindestens 110 dB beschallt werden).

Ist die Impulsantwort eines Raumes bekannt, kann die Nachhallzeit auch durch die Schröder-Rückwärtsintegration ermittelt werden. Diese Methode ist durch die Integration frei von statistischen Nebeneffekten im Abklingverhalten des Raumes und wird wie folgt angegeben:

$$r^2(t) = \int_t^{\infty} h^2(\tau) d\tau \quad (1.22)$$

$r(t)$ Abklingkurve

Die Nachhallzeit entspricht nun der Zeitdauer, in der $r^2(t)$ um 60 dB abfällt.

Als Anregungssignale eines Raumes kommen schmalbandige Signale, kurze breitbandige Impulse oder auch breitbandige MLS-Signale in Frage.

Wird die Impulsantwort in Frequenzteilbänder zerlegt, kann die Nachhallzeit auch für verschiedene Frequenzen bestimmt werden.

Hallradius

Das Schallfeld kann prinzipiell in 2 Teile zerlegt werden:

Beim Direktschallfeld breitet sich der Schall näherungsweise als Kugelwelle aus. Die Direktschallenergiedichte nimmt dabei proportional $\frac{1}{r^2}$ ab.

$$E_{Dir} = \frac{P}{4\pi \cdot r^2 \cdot c} \left[\frac{J}{m^3} \right] \quad (1.23)$$

P Akustische Leistung der Schallquelle
 r Abstand zur Quelle

In Räumen bildet sich zusätzlich auch das diffuse Schallfeld aus. Dabei wird die Schallenergie einerseits zum Teil reflektiert und andererseits absorbiert. Bei stationärer Schallabstrahlung ist dabei die abgestrahlte Energie der Schallquelle gleich der absorbierten Energie der Wände, sofern die Luftabsorption vernachlässigt wird.

Durch Reflexionen wird die Energie im Raum sozusagen zusammengehalten. Die diffuse Energiedichte, welche den Nachhall charakterisiert, wird über die Raumeigenschaften berechnet.

$$E_{Diff} = \frac{4P}{A \cdot c \cdot \alpha} \quad (1.24)$$

Als Hallradius ist nun jener Abstand definiert, bei dem die Direktschallenergie und die Diffusschallenergie gleich groß sind .

$$E_{Dir} \stackrel{!}{=} E_{Diff} \quad \Leftrightarrow \quad \frac{4P}{A \cdot c \cdot \alpha} = \frac{P}{4\pi \cdot r^2 \cdot c}$$

Daraus erhält man nun:

$$r_H = \sqrt{\frac{\alpha A}{16\pi}} \quad (1.25)$$

Wird nun (1.21) eingefügt, ergibt sich schließlich:

$$r_H = 0,057 \sqrt{\frac{V}{T_N}} \quad (1.26)$$

Abb. 1.6 zeigt nun die Überlagerung von Direkt- und Diffusschallfeld in einem Raum.

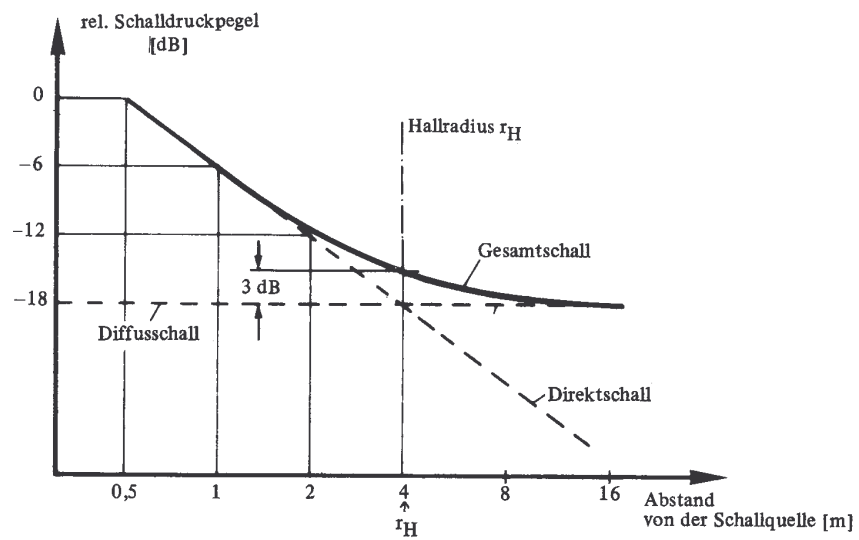


Abb. 1.6: Zusammensetzung von Direkt- und Diffusschall [Lit. 7]

Systemtheoretische Betrachtung der Impuls-antwort

1.5 Beschreibung deterministischer Signale

1.5.1 Darstellung im Zeitbereich

Darstellung einer Kosinus-Schwingung:

$$x(t) = A \cdot \cos(\omega t + \varphi) \quad (2.27)$$

$x(t)$ Zeitsignal
 A Amplitude
 φ Anfangswinkel

In Exponentialform angeschrieben:

$$x(t) = \frac{A}{2} (e^{j\omega t + \varphi} + e^{-j\omega t - \varphi}) \quad (2.28)$$

Durch Abtastung des kontinuierlichen Zeitsignals kann $x(t)$ nun auch in zeitdiskreter Form dargestellt werden.

$$x(n) = A \cdot (e^{j\theta n + \varphi} + e^{-j\theta n - \varphi}) \quad (2.29)$$

$t = nT$
 T Abtastintervall
 $\frac{1}{T} = f_s$ Abtastfrequenz
 $\omega = 2\pi f$ Kreisfrequenz im analogen Bereich
 $\theta = \omega T = \frac{2\pi f}{f_s}$ Kreisfrequenz im zeitdiskreten Bereich

Um das zeitdiskrete Signal wieder im zeitkontinuierlichen Bereich darstellen zu können, muss gewährleistet sein, dass $0 \leq \theta < \pi$ ist.

Es kann also im Zeitdiskreten nur folgender Frequenzbereich dargestellt werden:

$$0 \leq f < \frac{f_s}{2} \quad \text{..... Abtasttheorem}$$

1.5.2 Fouriertransformation

Die Darstellung von Signalen im Frequenzbereich erfolgt durch Fourier-Transformation der Zeitbereichsdarstellung.

Kontinuierliche Darstellung

$$\underline{X}(f) = F\{x(t)\} = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} df \quad (2.30)$$

Durch Rücktransformation gelangt man wieder in den Zeitbereich:

$$x(t) = F^{-1}\{\underline{X}(f)\} = \int_{-\infty}^{\infty} \underline{X}(f) e^{j2\pi ft} df \quad (2.31)$$

Zeitdiskrete Darstellung

$$\underline{X}(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{\frac{-j2\pi \cdot n \cdot k}{N}} \quad (2.32)$$

$$x(n) = \frac{1}{N} \sum_{K=0}^{N-1} \underline{X}(k) \cdot e^{j \frac{2\pi \cdot n \cdot k}{N}} \quad (2.33)$$

1.5.3 Faltung

Faltung im Zeitbereich:

$$y(t) = \int_{-\infty}^{\infty} v(\tau) w(t - \tau) d\tau = v(t) * w(t) \quad (2.34)$$

Die Faltung im Zeitbereich entspricht dabei einer Multiplikation im Frequenzbereich.

$$\underline{Y}(f) = \underline{V}(f) \underline{W}(f) \quad (2.35)$$

Kommutativität

$$\int_{-\infty}^{\infty} v(\tau) \cdot w(t - \tau) d\tau = \int_{-\infty}^{\infty} w(\tau) \cdot v(t - \tau) d\tau \quad (2.36)$$

$$v * w = w * v \quad (2.37)$$

1.5.4 Korrelation

Korrelationsfunktion für reelle Signale:

$$R_{v,w}(t) = \int_{-\infty}^{\infty} v(\tau) \cdot w(\tau - t) d\tau \quad (2.38)$$

Es gilt: $R_{w,v}(t) = R_{v,w}(-t)$ (2.39)

Berechnung der Korrelationsfunktion über Fourier-Transformation

$$R_{w,v}(t) = R_{v,w}(-t) = \int_{-\infty}^{\infty} v(\tau) \cdot w(t + \tau) d\tau \quad (2.40)$$

$$\begin{aligned} F\{R_{w,v}(t)\} &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} v(\tau) \cdot w(t + \tau) d\tau \right] \cdot e^{-j\omega t} dt \\ &= \int_{-\infty}^{\infty} v(\tau) \cdot \left[\int_{-\infty}^{\infty} w(t + \tau) e^{-j\omega t} dt \right] d\lambda \\ &= \int_{-\infty}^{\infty} v(\tau) \left[\underline{W}(f) \cdot e^{j\omega\tau} \right] d\tau \\ &= \left[\int_{-\infty}^{\infty} v(\tau) \cdot e^{+j\omega\tau} d\tau \right] \underline{W}(f) \end{aligned} \quad (2.41)$$

Dualität:

Der, in eckiger Klammer befindliche Teil aus (2.15), wird nochmals in gewohnter Form angeschrieben:

$$\left[\int_{-\infty}^{\infty} \underline{V}(f) \cdot e^{j\omega f} df \right] = v(t) \quad (2.42)$$

Ist $v(t)$ dual mit $\underline{V}(f)$,

$$v(t) \Leftrightarrow \underline{V}(f)$$

dann folgt

$$\underline{V}(t) \Leftrightarrow v(-f) \quad (2.43)$$

Aus (2.16) in (2.15) eingesetzt, folgt somit:

$$\underline{R}_{w,v}(f) = \underline{V}(-f) \cdot \underline{W}(f) \quad (2.44)$$

Da $\underline{v}(t)$ reell ist, gilt $\underline{v}(-f) = \underline{v}^*(f)$

$$\underline{R}_{w,v}(f) = \underline{V}^*(f) \cdot \underline{W}(f) \quad (2.45)$$

Somit zeigt sich, dass die Korrelation sehr ähnlich der Faltung zweier Signale berechnet wird. Es muss lediglich $\underline{V}(f)$ komplex konjugiert werden, womit die Korrelationsberechnung sehr effizient im Frequenzbereich erfolgen kann.

1.6 Impulsantwort und Übertragungsfunktion

Es sei folgendes System gegeben, das zu einem Eingangssignal $x(t)$ über einen Systemoperator T , das Ausgangssignal $y(t)$ liefert:



Abb. 2.1: Übertragungssystem

Dabei sollen nun zwei Voraussetzungen für das System getroffen werden, die es zu einem „LTI“-System machen:

1.6.1 LTI-Systeme

Als LTI-Systeme (linear time invariant) werden Übertragungssysteme bezeichnet, welche die Eigenschaft der Linearität und Zeitinvarianz aufweisen.

Linearität

Gilt: $y_1(t) = T[x_1(t)]$

$$y_2(t) = T[x_2(t)]$$

$$x_3 = a \cdot x_1(t) + b \cdot x_2(t)$$

so muss für lineare Systeme gelten: $y_3 = a \cdot y_1(t) + b \cdot y_2(t)$ (2.46)

Zeitinvarianz

Gilt: $x_2(t) = x_1(t - t_0)$

muss für zeitinvariante Systeme gelten: $y_2(t) = y_1(t - t_0)$ (2.47)

Weist also ein Übertragungssystem die Eigenschaften der Linearität und Zeitinvarianz auf, kann das System mit einer Impulsantwort $h(n)$ charakterisiert werden.

Für zeitvariante Systeme ist die Impulsantwort zeitabhängig.

Jedes LTI-System ist dabei durch die Kenntnis der Impulsantwort eindeutig beschrieben. Die Eingangs-, Ausgangsbeziehung des Systems kann durch die, in Punkt 2.1.3 angegebene Faltung im Zeitbereich, oder durch Multiplikation im Frequenzbereich angegeben werden.

1.6.2 Übertragungsfunktion eines LTI-Systems

Im Allgemeinen hat eine Übertragungsfunktion im Laplace-Bereich folgende Form:

$$\underline{H}(s) = \frac{a_0 + a_1s + \dots + a_ms^m}{b_0 + b_1s + \dots + b_ns^n} = \frac{\underline{Y}(s)}{\underline{X}(s)} \quad (2.48)$$

wobei $s = \delta + j\omega$

δ Wachstmaß

Wir beschränken uns jedoch auf den Fourier-Bereich, welcher als Teilbereich des Laplace-Bereichs (Fourier-Bereich entspricht imaginärer Achse des Laplace-Bereichs) verstanden werden kann.

$$\delta = 0 \Rightarrow s = j\omega = j2\pi f$$

$$\underline{H}(f) = \frac{\underline{Y}(f)}{\underline{X}(f)} \quad (2.49)$$

1.6.3 Akustische Übertragungsfunktion des Raumes

Aus systemtheoretischer Sicht würde man gerne die Akustik eines ganzen Raumes in ein LTI-System verpacken. Der Raum wäre damit akustisch eindeutig beschrieben. Prinzipiell ist festzustellen, dass eine RIR nur das akustische Übertragungsverhalten eines Raumes zwischen 2 Punkten beschreibt.

Jedoch nur mit Einschränkungen lässt sich eine RIR als LTI-System einordnen. Betrachtet man einen Raum, worin sich niemand befindet und keine mechanische Veränderungen passieren, würde man eigentlich das Übertragungsverhalten von einem Lautsprecher zu einem Mikrofon als konstant annehmen.

Über längere Zeit stellt sich jedoch eine Drift ein, die wir später (Kapitel 5) auf die Temperatur zurückführen werden. Im Normalfall wird die Temperaturveränderung im Vergleich zur Messdauer einer RIR langsam verlaufen und zumindest den Messvorgang nicht stören. Jedoch das Öffnen eines Fensters im Winter, oder starke Sonneneinwirkung bei schlecht wärmegeprägten Gebäuden kann auch in wenigen Minuten Temperaturänderungen im Bereich von 1 °C bringen. Diese Veränderung bewirkt jedoch schon eine Änderung der Raumimpulsantwort, dass diese für Entzerrungszwecke oder Echounterdrückung (siehe Kapitel 3) unbrauchbar wird.

Befinden sich Personen im Raum, die sich bewegen, oder verändert sich die Position von Einrichtungsgegenständen, so ist die Raumimpulsantwort sowieso zeitvariant. Hier wird die Zeitvarianz sogar so groß, dass selbst die Zeitdauer für die Messung einer Impulsantwort (ca. 0,5 - 2 Sekunden) nicht mehr als stationär angesehen werden kann.

Das akustische Übertragungsverhalten eines Raumes muss also als zeitvariantes System modelliert werden, kann aber unter bestimmten Bedingungen als quasistationäres LTI-System betrachtet werden. Unter quasistationär verstehen wir in diesem Zusammenhang, dass die Raumimpulsantwort in kurzen Zeitabschnitten als zeitinvariant betrachtet werden kann (Sekundenbereich bis wenige Minuten), diese jedoch über längeren Betrachtungszeitraum (Stundenbereich) zeitvariant ist. Die quasistationäre Impulsantwort hat also nur eine beschränkte Gültigkeit, die von der Veränderung des Raumes abhängt.

Es wird im folgenden ein Raum mit einem Lautsprecher an der Position 1 und ein Mikrofon an der Stelle 2 angenommen. Dieses System kann nun als abgeschlossen betrachtet werden, sofern sich keine akustisch relevanten Parameter innerhalb des Raumes ändern. $\underline{H}_{L,M}(f)$ wird dabei als die Übertragungsfunktion vom Lautsprecher zum Mikrofon bezeichnet. Nun werden der Lautsprecher und das Mikrofon vertauscht $\underline{H}_{L,M}(f) \Rightarrow \underline{H}_{M,L}(f)$. Mit der geometrischen Betrachtungsweise der Raumakustik anhand des Strahlenmodells (siehe Abb. 1.4), können einfach alle Schallstrahlen richtungsmäßig invertiert werden, und es gilt:

$$\underline{H}_{L,M}(f) = \underline{H}_{M,L}(f) \tag{2.50}$$

Im folgenden wird die Gesamtübertragungskette zur Messung einer Impulsantwort gezeigt:



Abb. 2.2: Gesamtübertragungskette zur Impulsantwortmessung

Generell ist hier zu sagen, dass die Impulsantwort nur in dem Bereich bestimmt werden kann, welcher im Messbereich der gesamten Messkette von Abb. 2.2 zu sehen ist. Dabei kann als untere Grenzfrequenz 40Hz (Lautsprecher) und als obere Grenzfrequenz 20kHz (Antialiasingfilter des A/D-Wandlers) angegeben werden (siehe Anhang C).

Innerhalb der Bandbegrenzung können nun näherungsweise sowohl die Verstärker als auch die D/A- bzw. die A/D-Wandler mit $\underline{H}(f)=1$ angenommen werden. Somit reduziert sich das Gesamtsystem auf die unten stehenden drei relevanten Übertragungsfunktionen.

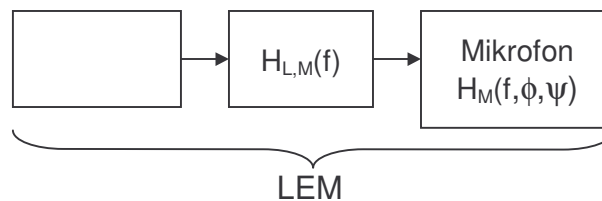


Abb. 2.3: Blockschaltbild der verbleibenden relevanten Übertragungsfunktionen

Da eine Raumimpulsantwort nur mit einem Lautsprecher und einem Mikrofon ermittelt werden kann, beinhaltet das Ergebnis der Impulsantwortmessung immer die, in Abb. 2.3 angegebene Messkette. Um also die akustische Übertragungsfunktion vom Lautsprecher zum Mikrofon zu erhalten, müssten jedoch noch die Frequenzgänge von Lautsprecher und Mikrofon herausgerechnet werden.

Genaugenommen dürften die Übertragungsfunktionen des Lautsprechers und des Mikrofons nicht als Funktionen der Frequenz alleine angegeben werden, da noch die Einfallsrichtung des Schalls berücksichtigt werden müsste, welche in Abb. 2.3 mit ϕ = Azimut und ψ = Elevationswinkel im Raum, angegeben werden. Das würde aber die Komplexität des Systems erheblich erhöhen, und wird daher nicht extra berücksichtigt.

Aus diesem Grund wird ab jetzt die Raumimpulsantwort als die Impulsantwort des Systems „Lautsprecher – Raum – Mikrofon“, auch LEM (loudspeaker enclosure microphone), betrachtet. Jede Veränderung in einem der 3 Blöcke der Abb. 2.3, wird also als Änderung der RIR betrachtet.

1.6.4 Bestimmung einer Übertragungsfunktion

Um nun die Übertragungsfunktion eines Systems zu bestimmen, werden verschiedene mögliche Eingangssignale kurz angesprochen. Man betrachte dabei Abb. 2.1, jedoch mit der Voraussetzung, dass es sich um ein LTI-System handelt.

Impuls als Anregungssignal

Laut Definition erhält man die Impulsantwort eines Systems, wenn dieses mit einem Impuls erregt wird. Diese Methode ist jedoch nur in digitalen Systemen sinnvoll.

In realen Systemen kann nämlich aufgrund endlicher Flankensteilheit kein idealer Impuls erzeugt werden.

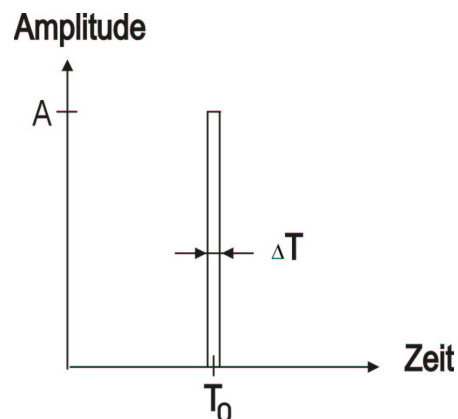


Abb. 2.4: Näherung des Dirac-Impulses

Die Möglichkeit, den Dirac-Impuls anzunähern ($\Delta T \rightarrow 0$, A aufgrund der analogen Aussteuergrenze in seiner Größe jedoch beschränkt) hat jedoch zur Folge, dass die Anregungsenergie sehr klein wird und das Messergebnis der Impulsantwort aufgrund des kleinen SNR (Signal-zu-Rauschleistungsverhältnis) darunter leidet.

Harmonisches Anregungssignal

Für eine Frequenz ω kann die Übertragungsfunktion auch anhand der Eingangs-, Ausgangsbeziehung eines sinusförmig erregten Systems bestimmt werden.

$$\underline{H}(j\omega) = \frac{\underline{Y}(j\omega)}{\underline{X}(j\omega)} \quad (2.51)$$

Der gesamte Frequenzgang könnte beispielsweise durch die TDS-Methode [Lit. 8, 9] ermittelt werden. Dabei wird über einen Frequenz-Wobbel (Sweep) die Übertragungsfunktion des kompletten Frequenzbereichs bestimmt.

Breitbandiges Anregungssignal

Eine Übertragungsfunktion kann immer nur innerhalb des Frequenzbereichs gemessen werden, welcher vom Anregungssignal abgedeckt ist. Es drängt sich somit weißes Rauschen auf, welches alle Frequenzbereiche abdeckt.

Aufgrund seiner konstanten Spektraldichtefunktion wird in jedem Frequenzbereich mit gleichem SNR gemessen.

In der Praxis gibt es jedoch nur "bandbegrenzt weißes" Rauschen, welches konstante Spektraldichte in einem endlichen Frequenzbereich aufweist, da sonst die Leistung des Rauschsignals unendlich werden würde.

$$P_y = \int_{-\infty}^{\infty} G_y(f) df = \int_{-\infty}^{\infty} |H(f)|^2 G_x(f) df \quad (2.52)$$

$G_x(f)$, $G_y(f)$ Spektraldichtefunktion des Eingangs-, bzw. Ausgangssignals

Ist also $G_x(f) = \text{konstant}$, folgt $P_y \rightarrow \infty$ sowie $P_x \rightarrow \infty$, falls $\int_{-\infty}^{\infty} |H(f)|^2 df$ konvergiert.

Breitbandiges, deterministisches Anregungssignal

Ein Signal, welches die benötigten Erfordernisse eines "bandbegrenzt weißen" Spektrums bietet, ist beispielsweise eine MLS (Maximum Length Sequence).

Aus messtechnischer Sicht eignet sich dieses Signal besonders zur Impulsantwortbestimmung, da es sich um ein deterministisch-periodisches (jedoch breitbandiges) Signal handelt, welches die Messtechnik vereinfacht.

Die genauere Beschreibung des MLS-Verfahrens erfolgt im nächsten Punkt.

1.7 MLS-Verfahren zur Bestimmung der Raumimpulsantwort

Zur Bestimmung der RIR wurde das MLS-Verfahren (Maximum Length Sequence) verwendet. Es findet in der Praxis häufig Anwendung.

Man benötigt für dieses Verfahren lediglich einen Lautsprecher, ein Mikrofon, das MLS-Signal und einen PC, welcher als Signalgenerator, Aufnahmegerät und zur Auswertung der Signalverarbeitung dient.

Folgende Vorteile sprechen für dieses Verfahren:

- es wird kein teures Spezialmessgerät benötigt
- gesamte Impulsantwort messbar
- gesamter Frequenzbereich fein aufgelöst messbar
- unempfindlich gegen Störgeräusche
- universell einsetzbar
- kurze Messzeit der RIR aufgrund des breitbandigen Messvorganges

Nachteil:

- Messsignal belastet Hochtonlautsprecher stark

Da das Frequenzspektrum des MLS-Signals breitbandig (konstant über den Messbereich) ist, wird der Hochtonlautsprecher um ein Vielfaches mehr belastet als bei normaler Musikwiedergabe, wo die Leistungsverteilung von Bass, Mittel, Höhen ca. 60 zu 30 zu 10 Prozent beträgt. Bei weißem Spektrum ist die Leistungsverteilung proportional dem Frequenzbereich der jeweiligen Lautsprecher. Bei einem Zwei-Kanal-Lautsprecher mit einer Trennfrequenz von 2 kHz würde somit am Tieftonlautsprecher nur ein Zehntel der Leistung anliegen.

Um also den Lautsprecher vor thermischer Überlastung zu schützen, gibt es mehrere Möglichkeiten:

- Spektrumanpassung (MLS-Signal in den Höhen bedämpfen, um den Lautsprecher thermisch zu entlasten). Später wird diese Filterung durch inverses Filtern entzerrt.
- Überlastschutz der Box (Overload Led)
- Begrenzer vor Verstärker

Der Überlastschutz der Box bzw. der Limiter dürfen nur als Indikator verwendet werden. Da die MLS-Methode Linearität voraussetzt, müssen der Lautsprecher bzw. der Begrenzer demnach auch im linearen Bereich betrieben werden, also unterhalb der Indikatorschwelle, da es sonst zu irreversiblen Verzerrungen bei der Messung kommen würde!

Wird keine dieser Möglichkeiten verwendet, muss ein größerer Sicherheitsabstand zum maximalen Aussteuerpegel berücksichtigt werden. Dieser verringert zwar das Signal zu Rauschverhältnis, schützt jedoch die Hochtöner vor thermischer Überlastung.

Ein Signalausschnitt einer MLS $s[n]$ ist nun in Abb. 2.5 dargestellt.

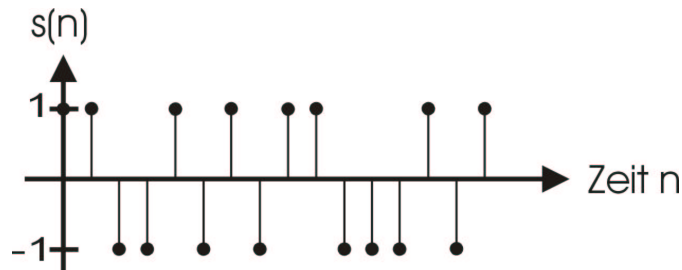


Abb. 2.5: Signalausschnitt einer MLS

Eine MLS mit Periode L besitzt die Länge:

$$L = 2^N - 1 \quad N \varepsilon \text{ natürliche Zahlen} \quad (2.53)$$

Weiters muss N so gewählt werden, dass die Länge einer MLS Periode größer als die Länge der zu messenden Raumimpulsantwort ist. Beträgt die Länge einer RIR beispielsweise 35000 Abtastwerte, so wird idealerweise eine MLS mit $N=16$ ("16 bit MLS") verwendet, welche eine Länge von 65535 Abtastwerten besitzt.

Zwar könnte eine MLS mit $N > 16$ verwendet werden, dies hätte aber eine längere Messzeit der RIR zur Folge.

Die Messzeit T_M für eine RIR erhält man wie folgt:

$$T_M = \frac{L}{f_s} \quad (2.54)$$

Eine Periode der MLS s hat folgende Eigenschaften:

- Sie besitzt nur Amplitudenwerte von +1 und -1.
Die Amplitudenverteilung dieser zwei Werte ist gleich, bis auf ein Sample, das zusätzlich den Wert -1 hat (exakte Gleichverteilung der Amplitudenwerte ist nicht möglich, da die MLS eine ungerade Sampleanzahl aufweist)
- Ihre Autokorrelationsfunktion ist praktisch ein Dirac-Impuls (Gleichung (2.33))

1.7.1 Herleitung für Berechnung der Impulsantwort

Es wird nun die Berechnung der Impulsantwort im zeitdiskreten Bereich hergeleitet [Lit. 10].

' Zeichen der Periodizität innerhalb einer Folge

Φ ... Zeichen für Kreuz- bzw. Autokorrelation, periodisch

\otimes ... zyklische Faltung

$$s'(n): r_{ss}(n) = s'(n) \Phi s'(n) = \sum_{k=0}^{L-1} s'(k) \cdot s'(k+n) \quad (2.55)$$

Autokorrelationsfunktion $r_{ss}(n)$ ohne Normierung $\frac{1}{L}$

$$r_{ss}(0) = L \quad \text{für } n = 0$$

$$r_{ss}(n) = -1 \quad \text{sonst: } 0 < n < L$$

Die Autokorrelationsfunktionen werden nun mit dem Faktor $\frac{1}{L+1}$ gewichtet und als Ω_{ss} bezeichnet.

$$\Omega_{ss}(n) = \frac{1}{L+1} s'(n) \Phi s'(n) = \frac{1}{L+1} \sum_{k=0}^{L-1} s'(k) \cdot s'(k+n) \quad (2.56)$$

$$\Omega_{ss}(0) = \frac{L}{L+1} \quad \text{für } n = 0 \quad (2.57)$$

$$\Omega_{ss}(n) = \frac{-1}{L+1} \quad \text{für } 0 < n < L \quad (2.58)$$

Der maximale Spitze–Spitze Wert der Funktion $\Omega_{ss} = 1$. Damit kann die nächste Gleichung wie folgt angeschrieben werden:

$$\Omega_{ss}(n) = \delta(n) - \frac{1}{L+1} \quad 0 \leq n < L \quad (2.59)$$

Das Ergebnis der Autokorrelationsfunktion bringt also eine Überlagerung des Dirac-Impulses mit einem kleinen negativen Gleichanteil.

Wenn L sehr groß ist, fällt dieser Term nicht ins Gewicht und kann daher vernachlässigt werden.

Gleichung (2.33) ist sehr wichtig. Vernachlässigt man den zweiten Term in (2.33), so liegt die Korrelationsfunktion von perfekt weißem Rauschen vor, wenn man auch von der L -Periodizität in $\Omega_{ss}(n)$ absieht.

$$y'(n) = s'(n) \otimes h'(n) = \sum_{k=0}^{L-1} s'(k) \cdot h'(n-k) \quad (2.60)$$

Zyklische Faltung der MLS mit der Impulsantwort

$$\Omega_{xy}(n) = s'(n) \Phi y'(n) = \frac{1}{L+1} \sum_{k=0}^{L-1} s'(k) y'(n+k) \quad (2.61)$$

Kreuzkorrelation der MLS mit dem Ausgangssignal y'

Wird (2.33) in (2.34) eingesetzt erhält man:

$$\begin{aligned} \Omega_{xy}(n) &= s'(n) \Phi [s'(n) \otimes h'(n)] = [s'(n) \Phi s'(n)] \otimes h'(n) = \\ &= \Omega_{ss}(n) \otimes h'(n) = \sum_{k=0}^{L-1} \Omega'_{ss}(n) \cdot h'(n-k) \end{aligned}$$

$$\Omega'_{xy}(n) = \Omega'_{ss}(n) \otimes h'(n) \quad (2.62)$$

$$\Omega_{ss}(n) = \delta(n) - \frac{1}{L+1} \quad 0 \leq n < L \quad (2.63)$$

$$\Omega'_{xy}(n) = \left[\delta(n) - \frac{1}{L+1} \right] \otimes h'(n) = h'(n) - \frac{1}{L+1} \sum_{k=0}^{L-1} h'(k) \quad (2.64)$$

$$\Omega'_{xy}(n) = h'(n) - \underbrace{\frac{1}{L} \sum_{k=0}^{L-1} h'(k)}_{\text{Gleichanteil Komponente}} + \frac{1}{L(L+1)} \sum_{k=0}^{L-1} h'(k) \quad (2.65)$$

Der Term $\frac{1}{L+1} \cdot \sum_{k=0}^{L-1} h'(k)$ wird also durch die Umformung von (2.38) zu (2.39) in den Gleichanteil und einem weiteren Term zerlegt.

Durch Kreuzkorrelation der MLS mit dem Mikrofonsignal erhält man also die Raumimpulsantwort (LEM). Da eine Raumimpulsantwort frei von einem Gleichanteil ist (die Strecke LEM kann keinen Gleichanteil übertragen), stören die beiden letzten Terme in (2.35) nicht. Erstens ist der Summe der Impulsantwort $\sum_{k=0}^{L-1} h'(k)$ sehr klein (idealerweise ist sie null), und zweitens wird diese Summe mit dem ebenfalls sehr kleinen Faktor $\frac{1}{L}$ bzw. $\frac{1}{L(L+1)}$ multipliziert.

Bei einer 16 Bit MLS beträgt der Faktor $\frac{1}{L}$ -96 dB.

Der zweite Faktor, welcher durch die Umformung von (2.38) zu (2.39) entsteht, ist noch viel weniger relevant (-192 dB).

Um nun die in Abschnitt 2.1.4.1 gezeigte Berechnung der Korrelation über Multiplikation im Frequenzbereich auch im digitalen Bereich zu realisieren, werden noch einige Überlegungen angestellt.

Wird eine lineare Faltung zweier zeitdiskreten Folgen

$x(n)$ Länge N
 $h(n)$ Länge M

durchgeführt, so ergibt sich eine Folge y der Länge $N+M-1$.

Wird nun die lineare Faltung über den Frequenzbereich berechnet, müssen vor der FFT, die beiden Folgen mittels Zero-Padding mindestens auf die Länge $N+M-1$ gebracht werden, um nach der IFFT dasselbe Ergebnis wie das im Zeitbereich zu erhalten. Wird diese Forderung jedoch nicht erfüllt, muss man die Eigenschaften der zirkulären Faltung berücksichtigen.

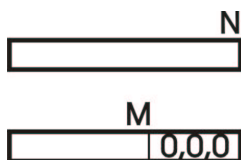


Abb. 2.6a: Blöcke ohne Zero-Padding

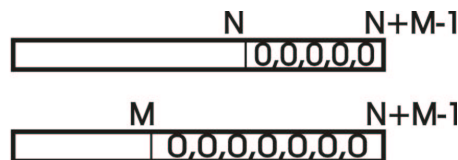


Abb. 2.6b: Blöcke mit Zero-Padding

Werden die beiden Folgen von Abb. 2.6a ohne Zero-Padding über den Frequenzbereich gefaltet, ist das Ergebnis der zirkulären Faltung nach anschließender Rücktransformation in den Zeitbereich jedenfalls nur N Punkte lang und somit nicht mit dem Ergebnis der linearen Faltung (Länge $N+M-1$) identisch.

Die Erklärung dafür liefert die nächste Abbildung.

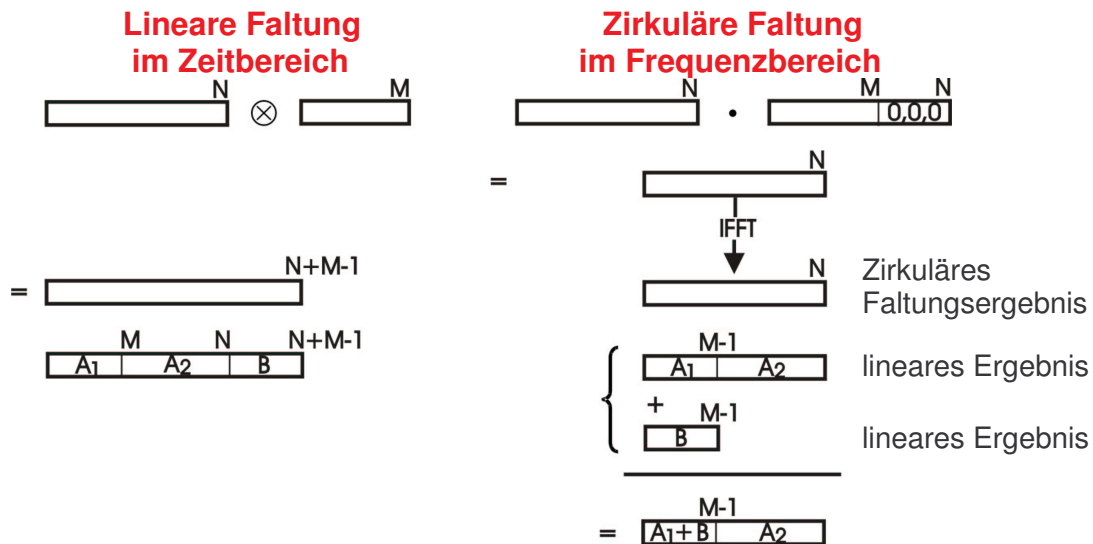


Abb. 2.7: Gegenüberstellung lineare – zirkuläre Faltung

Die linke Darstellung zeigt eine lineare Faltung im Zeitbereich, welche nun ein Ergebnis der Länge $N+M-1$ liefert. Dieses Ergebnis soll jetzt in 3 Teile mit Inhalt A_1 , A_2 , B gespaltet werden, wobei beiden Längen von A_1 und B gleich $M-1$ sind. Die rechte Seite beschreibt die Faltung im Frequenzbereich einer diskreten Fouriertransformation, welche durch Multiplikation durchgeführt wird. Dieses Faltungsergebnis ist nun im Zeitbereich nur N Punkte lang, womit es sich also vom linearen Faltungsergebnis unterscheidet.

Wegen der implizierten periodischen Fortsetzung bzw. Zirkularität bei der Fourieranalyse zeitdiskreter Folgen mit endlicher Länge, ist das zirkuläre Faltungsergebnis bis zur Länge $M-1$ gleich der Summe von A_1 und B des linearen Faltungsergebnisses, der Teil A_2 ist gleich dem linearen Faltungsergebnis.

Bei der zeitdiskreten Realisierung der Korrelation über den Frequenzbereich ist auf Folgendes zu achten. Damit keine zirkuläre Faltung entsteht, muss das Zero-Padding im vorderen Teil der Folgen passieren. Durch das nötige Konjugieren des MSL-Signals $s(n)$ des FFT-Produktes (siehe (2.19)), entsteht sonst eine zirkuläre Faltung.

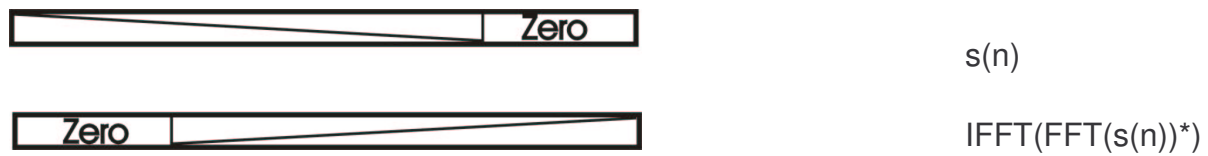


Abb. 2.8: Konjugieren einer reellen Zeitfunktion im Frequenzbereich resultiert in zeitlicher Spiegelung der Folge

Die Folge $s(n)$ wird also durch das Konjugieren in den hinteren Folgenblock gespiegelt, welcher jedoch nach Abb. 2.6b zum Verhindern der zirkulären Faltung null sein sollte.

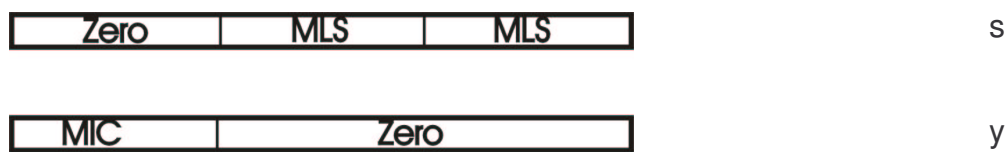


Abb. 2.9: Richtiges Zero-Padding für Korrelation im Frequenzbereich

Weiters ist noch wichtig, dass mindestens 2 MLS-Blöcke pro Impulsantwortberechnung verwendet werden, da an den Randbereichen der Kreuzkorrelation (Bereich, wo sich die Korrelationsblöcke nicht vollständig überlagern) nicht das richtige Ergebnis berechnet wird.

Da die MLS periodisch ist, muss nur eine Periode eingelesen werden. Jedoch ist es notwendig, diese zur Berechnung 2 Mal in den FFT-Transformationsblock zu schreiben.

Es ist nun möglich, die gesuchte Korrelation von Mikrofonsignal und MLS über die FFT zu berechnen. Die Blockfiltertechnik der FFT kann somit angewandt werden. Dies bietet z.B. die Möglichkeit, effektiv mehr als eine Impulsantwort pro MLS-Periode berechnen zu können, falls eine Überlappung der FFT Frames vorgenommen wird, welche in Abb. 2.10 gezeigt wird. Besitzt das Mikrofonsignal beispielsweise eine Länge von 655350 Abtastwerten, so können bei einem Offset von 65535 exakt 10 RIRs berechnet werden. Wird jedoch ein Offset von 20000 Abtastwerten verwendet, so erhält man bei der Berechnung 32 RIRs. Es ergibt sich somit eine Erhöhung der Zahl der RIRs pro MLS-Periode von 1 auf 3,2.

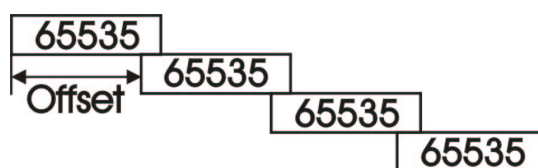


Abb. 2.10: Überlappung der zur Berechnung einer RIR verwendeten Aufnahmedatenblöcke

Generell wurde bei der Berechnung sämtlicher Raumimpulsantworten eine Offsetgröße der MLS-Periodendauer von 65535 Sample verwendet, die Geschwindigkeitskompensation bei der Arraymessung wurde jedoch durch Wahl eines entsprechend geänderten Offsets realisiert. Die Geschwindigkeitskompensation wird in Punkt 4.6.4 behandelt.

Eine stark zeitvariante RIR kann aber trotzdem nicht durch größere Überlappung der FFT-Frames kompensiert werden, da innerhalb einer MLS-Periode die Quasistationaritäts-Bedingung trotzdem verletzt wird.

3 Anwendungsbereiche der Raumimpulsantwort

In diesem Kapitel wird anhand einiger Beispiele gezeigt, welche Anwendungsmöglichkeiten sich durch die Kenntnis der Raumimpulsantwort ergeben.

3.1 Echokompensation

Abb. 3.1 zeigt die Prinzipsdarstellung einer Konferenzschaltung.

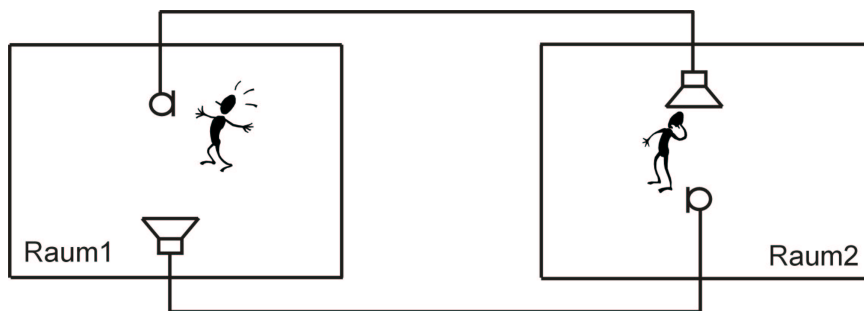


Abb. 3.1: Prinzipschaltbild des Echoproblems

In diesem Fall, kommt es bei der Übertragung zu dem Effekt der Rückkopplung. Der Sprecher in Raum 1 spricht hier mit dem Gesprächspartner in Raum 2. Dabei wird aber auch Mikrofon 2 beschallt, worauf über den Lautsprecher des Redners sein eigenes Echo zu hören ist. Dies wird einerseits als sehr störend empfunden, und kann andererseits bei mehrfachem Durchlaufen der Schleife zu Übersteuerung der Signalwege führen, was zu einem selbsterregt schwindenden, lauten Pfeifen oder Brummen führt, je nachdem welcher Frequenzbereich die Übersteuerung hervorgerufen hat.

Dieses Echoproblem kann man beispielsweise bei Rundfunksendungen feststellen, wenn ein Anrufer live auf Sendung geschaltet wird und dieser sein Empfangsgerät mit selbigem Live-Programm in der Nähe seines Telefons laut eingeschaltet hat.

Abb. 3.1 wird nun in ein Blockschaltbild übergeführt:

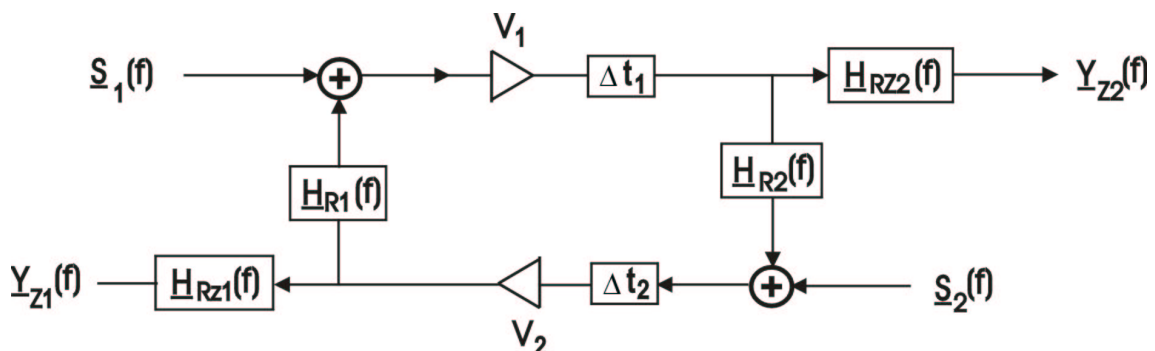


Abb. 3.2: Blockschaltbild des Echoproblems

$\underline{S}_1(f), \underline{S}_2(f)$	Eingangssignale der beiden Sprecher
$\underline{Y}_{z1}(f), \underline{Y}_{z2}(f)$	Ausgangssignale beim jeweiligen Zuhörer
$\underline{H}_{R1}(f), \underline{H}_{R2}(f)$	Übertragungsfunktionen vom Lautsprecher zum Mikrofon
$\underline{H}_{Rz1}(f), \underline{H}_{Rz2}(f)$	Übertragungsfunktionen vom Lautsprecher zum Zuhörer
V_1, V_2	Verstärkungsfaktoren
$\Delta t_1, \Delta t_2$	Gesamtverzögerungen des Signalpfades

Die bevorzugte Pfeif- oder Brummfrequenz f_0 im Rückkopplungsfall erhält man wie folgt:

$$f_0 = \frac{1}{\Delta t_1 + \Delta t_2}$$

Ein System neigt bei f_0 besonders zur Instabilität, da bei f_0 eine Phasenverschiebung des Rückkopplungsgrades um 360° erfolgt, also konstruktive Interferenz der Schwingung bei f_0 herrscht. Bei ganzzahligen Teilern von f_0 ergibt sich die konstruktive Interferenz ebenfalls.

Abb. 3.2 zeigt nun die Überlagerung zweier mitgekoppelter Systeme. Wenn $\underline{H}_{R1}(f)$ bzw. $\underline{H}_{R2}(f)$ ein FIR-System (System mit endlich langer Impulsantwort) ist, entsteht durch die Rückkopplung ein IIR-System (System mit unendlich langer Impulsantwort), das leicht instabil werden kann.

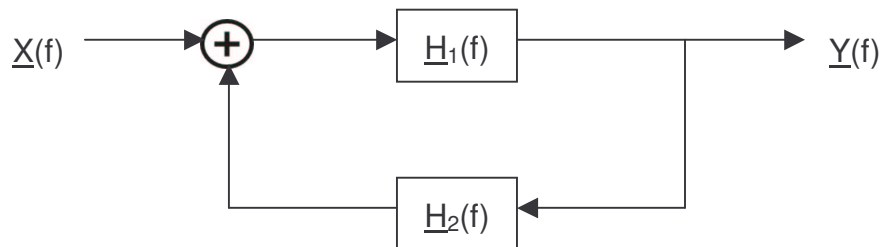


Abb. 3.3: Mitkopplung eines Übertragungssystems

Die Übertragungsfunktion aus Abb. 3.3 kann folgendermaßen angeschrieben werden:

$$\underline{Y}(f) = \left[\frac{\underline{H}_1(f)}{1 - \underline{H}_1(f)\underline{H}_2(f)} \right] \cdot \underline{X}(f) = \underline{H}(f) \cdot \underline{X}(f) \quad (3.66)$$

Das Übertragungsverhalten von Sprecher 1 und Sprecher 2 zum Zuhörer 2 von Abb. 3.2 wird wie folgt angegeben:

$$\frac{\underline{Y}_{z2}(f)}{\underline{S}_1(f)} = \frac{V_1}{1 - \underline{H}_{R1}(f)\underline{H}_{R2}(f)V_1V_2} \underline{H}_{Rz2}(f) \quad (3.67)$$

$$\frac{\underline{Y}_{Z2}(f)}{\underline{S}_2(f)} = \frac{\underline{H}_{R1}(f)V_1V_2}{1 - \underline{H}_{R1}(f)\underline{H}_{R2}(f)V_1V_2} \underline{H}_{RZ2}(f) \quad (3.68)$$

Es können 2 Effekte bei der Rückkopplung unterschieden werden:

1. Störendes Echo, welches die Konzentrationsfähigkeit des Sprechers reduziert
2. Instabilität des Systems aufgrund der Rückkopplung

Für die Beurteilung der Stabilität wird der Term $\frac{1}{1 - \underline{H}_{R1}(f)\underline{H}_{R2}(f)V_1V_2}$ aus (3.2) bzw. (3.3) herangezogen. Dabei wird bei der Stabilitätsbetrachtung jene Frequenz f_0 (z.B. Brummen oder Pfeifen) beurteilt, bei welcher das System am leichtesten instabil wird.

Instabilität kann nun folgendermaßen unterschieden werden:

Fall 1:

$$0 < 1 - |\underline{H}_{R1}(f_0)\underline{H}_{R2}(f_0)V_1V_2| \ll 1 \quad \text{für Frequenz } f_0$$

Hier handelt es sich zwar prinzipiell um ein stabiles System, welches jedoch aufgrund Überschreitens der eingeschränkten analogen Aussteuerungsgrenze (Mikrofon, Verstärker, Lautsprecher) instabil wird. Bei einer genügend großen Aussteuerungsreserve des Systems, muss das System jedoch nicht unbedingt instabil sein.

Fall 2:

$$0 \geq 1 - |\underline{H}_{R1}(f_0)\underline{H}_{R2}(f_0)V_1V_2|$$

Dieser Bereich beinhaltet den Fall, wo das System unabhängig von der analogen Aussteuerungsgrenze instabil ist. Sieht man vom Spezialfall des Oszillators ab ($\underline{H}_{R1}(f_0)\underline{H}_{R2}(f_0)V_1V_2 = -1$), so ist jedes rückgekoppelte System des Falls 2 instabil.

Um das Echo zu reduzieren bzw. Übersteuerung der Signalwege zu vermeiden, gäbe es prinzipiell folgende Möglichkeiten, die in jeden der beiden Signalwege eingebaut werden könnten:

- Verstärkungsregelung $G(t)$
- Verzögerungsstufe $\Delta\tau$
- Frequenzverschiebung Δf

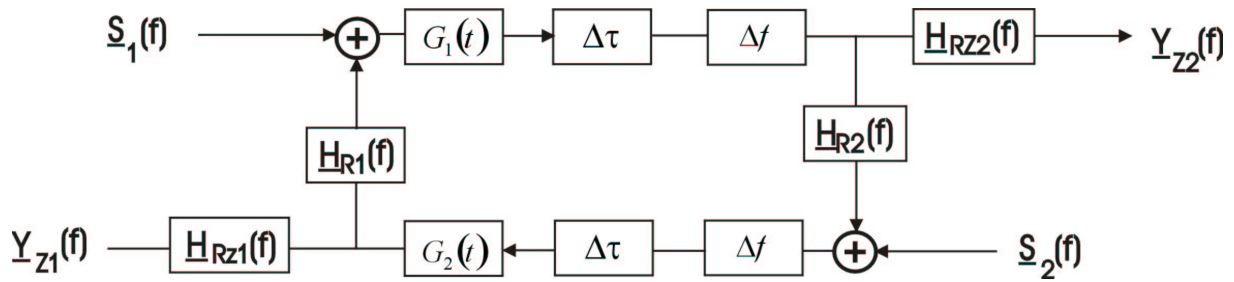


Abb. 3.4: Prinzipielle Möglichkeiten zur Reduktion des Echoproblems

Verstärkungsregelung

Da im Normalfall bei einer Konferenzschaltung nicht beide Personen gleichzeitig sprechen, kann der Mikrofonpfad des gerade Zuhörenden eigentlich stumm geschaltet werden.

$G_1(t)$ und $G_2(t)$ wirken dabei als Regelverstärker, die jeweils nicht beide gleichzeitig voll verstärken sollten. Dadurch wird die Rückkopplungsschleife stark gedämpft, womit die Echounterdrückung realisiert wird. Die Verstärkungsregelung wird dabei von der Signalenergie des aktiven Sprechers gesteuert.

Sprechen jedoch beide, hängt es dann von der konkreten Implementation ab, wie die beiden Regelverstärker reagieren.

Verzögerungsstufe

Eine weitere Möglichkeit besteht im Zeitversatz der Übertragung durch $\Delta\tau$. Wird $\Delta\tau$ größer als die Zeitspanne, über die ein Sprachsignal stark korreliert ist, kann die Stabilität des Systems für Sprachsignale erhöht werden. Für schmalbandige Störgeräusche im Hintergrund bringt die Methode des Zeitversatzes weniger, da diese auch über längere Zeit korreliert sind. Die Signalverzögerung darf außerdem nicht zu groß sein, damit die Reaktionszeit des Gesprächspartners nicht unnatürlich lange wird. Der Störeinfluss des Echos wird durch diese Methode jedoch nicht kompensiert, da das Echo nur verzögert wird, um die Systeminstabilität zu vermeiden.

Frequenzverschiebung

Hier wird die Neigung zur Instabilität zufolge der Rückkopplung dadurch verringert, dass die Signale mit einem Frequenzversatz Δf übertragen werden. Das Sprachsignal wird dabei beim Durchlaufen der gesamten Schleife um $2 \times \Delta f$ im Frequenzbereich verschoben. Damit kann die Stabilität des Systems auch bei stationären Störgeräuschen verbessert werden. Das störende Echo wird durch diese Methode jedoch ebenfalls nicht eliminiert.

1.7.3 Echokompensation durch Modellierung der Raumübertragungsfunktion

Ist die Raumimpulsantwort jedoch bekannt, kann das durch die Rückkopplung verursachte Echo sehr einfach vermieden werden.

Es wird nun ein Raum mit zugehöriger Beschallungseinrichtung im Blockschaltbild dargestellt. Hier wird noch einmal verdeutlicht, dass es keineswegs nachteilig ist, nur die Impulsantwort des Systems LEM (siehe Abb. 2.3) zu kennen, da genau diese Übertragungsfunktion für den Echokompensationspfad benötigt wird. Es sei nochmals darauf hingewiesen, dass $h_{LEM}(n)$ von uns als Raumimpulsantwort bezeichnet wird.

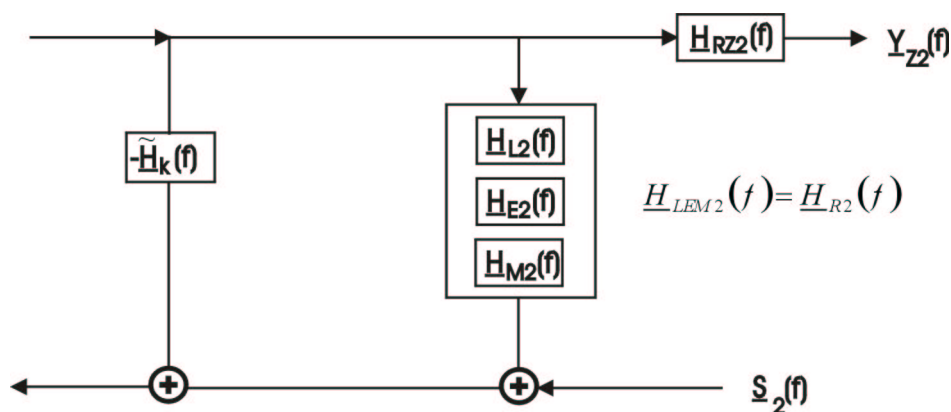


Abb. 3.5: Echokompensation auf Grundlage der bekannten Raumimpulsantwort

Da nun die zur Echokompensation verwendete Übertragungsfunktion nicht exakt gleich der tatsächlichen Raumübertragungsfunktion ist, bezeichnen wir nun die Übertragungsfunktion des Kompensationspfades als $-\tilde{H}_k(f)$. Je ähnlicher $\tilde{H}_k(f)$ der tatsächlichen Raumübertragungsfunktion $\underline{H}_{R2}(f)$ ist, desto besser funktioniert also die Echokompensation.

Wie in Kapitel 2 schon angesprochen wurde, haben wir mit der Zeitvarianz der RIR zu rechnen. Es wird sich noch zeigen, ob und in welchem Maße die Leistungsfähigkeit der Echokompensation darunter leidet.

1.7.4 Übertragungsbereiche

Für die Echokompensation ist nur der Frequenzbereich der Raumimpulsantwort wichtig, der auch übertragen wird.

Es lassen sich folgende Bereiche angeben:

- Telephonie 300 – 3400 Hz
- Breitbandsprachübertragung 50 – 7000 Hz
- Hifi – Bereich 16 – 20000 Hz

In dieser Arbeit soll vor allem die Echokompensation für den Bereich der Breitbandsprachübertragung untersucht werden. Somit wäre eine Abtastrate von 16 kHz ausreichend. Da jedoch die in dieser Arbeit durchgeführte Messung auch für andere Bereiche Aussagekraft haben sollen, wurde mit $f_s = 44,1$ kHz aufgenommen. Es werden also sämtliche Betrachtungen für diesen in der Audioverarbeitung üblichen Frequenzbereich ausgelegt.

1.7.5 SDR (Signal to Deviation Ratio)

Damit die in 3.1.1 vorgestellte Echounterdrückung gut funktioniert, ist folgende Bedingung wichtig:

$\underline{H}_{R2}(f)$ wird nun als wahrer Wert der Raumimpulsantwort $\underline{H}_w(f)$ zum gerade betrachteten Zeitpunkt bezeichnet. Durch Bilden der Differenz mit $\underline{H}_k = \tilde{\underline{H}}_{R2}(f)$, erhält man das Fehlerspektrum der Echounterdrückung.

$$\underline{H}_w(f) - \underline{H}_k(f) = \underline{E}(f) \quad \int_{-\infty}^{\infty} |\underline{E}(f)|^2 df \ll \int_{-\infty}^{\infty} |\underline{H}_w(f)|^2 df \quad (3.69)$$

Oder im Zeitbereich

$$h_w(n) - h_k(n) = e(n) \quad e(n) \ll h_w(n) \quad (3.70)$$

- $\underline{H}_w(f)$ Wahrer Wert der RIR im Frequenzbereich
- $\underline{H}_k(f)$ Kompensationswert oder zur Echokompensation verwendete Funktion
- $\underline{E}(f)$ Frequenzgang des Echokompensatorfehlers

Die Echounterdrückung funktioniert nun umso besser, je kleiner die Energie des Kompensationsfehlers im Bezug auf die Impulsantwort ist.

3.1.3.1 Beschreibung der Signalqualität

Meist wird zur Beschreibung der Signalqualität das SNR (Signal to Noise Ratio) angegeben. Diese Größe liefert jedoch keine Aussage über die Ähnlichkeit zweier Signale.

Eine Größe, die auch Aussage über die Ähnlichkeit von Signalen gibt, ist das SDR [Lit. 11]. Hier werden die einzelnen Zeitwerte bzw. Frequenzbins zweier Signale verglichen.

Für Raumimpulsantworten ist es nun besser, das SDR zur Beurteilung der Signalqualität und der Reproduzierbarkeit zu verwenden, da für unseren Zweck der Echokompensation das SNR allein zu wenig Aussagekraft bietet. Schließlich haben wir es nicht nur mit geräuschartigen Störungen zu tun, sondern auch mit systematischen Abweichungen zweier Raumimpulsantworten, die auf deterministische Einflüsse wie der Temperaturdrift oder der Raumänderung zurückgeführt werden.

Betrachtet man zwei Impulsantworten, wovon eine die um einen Abtastwert verzögerte Version der anderen ist, so wird das SNR beider Signale durch die Signalverzögerung nicht verändert. Jedoch zur Echokompensation wäre die verzögerte Version unbrauchbar, da $e(n)$ durch die Verzögerung sogar eine größere Gesamtenergie als $h_v(n)$ aufweisen kann (siehe Gl. (3.5)).

Ganz generell kann man also feststellen, dass das SDR auch für die Beschreibung der Signalqualität in zeitvarianten Systemen geeignet ist.

Das S/N-Verhältnis kann in der Aussagekraft als Spezialfall des SDR aufgefasst werden, bei dem Unterschiede zwischen zwei Raumimpulsantworten ausschließlich auf den Einfluss von Messrauschen zurückgeführt wird.

3.1.3.2 Definition des SDR bzw. SDRF

$$SDR = 10 \log_{10} \frac{\sum_{n=0}^{N-1} \{h_1[n]\}^2}{\sum_{n=0}^{N-1} \{h_1[n] - h_2[n]\}^2} \quad [dB] \quad (3.71)$$

h_1 Referenzimpulsantwort, wahrer Wert, aktueller Bezugswert

h_2 Driftende Impulsantwort, Kompensationswert,
aktuell gemessener Vergleichswert der Impulsantwort

Der Wert des SDR wird umso größer je ähnlicher bzw. umso kleiner je unterschiedlicher die Referenz- und die gemessene Impulsantwort sind.

Analog zu (3.6) kann noch das SDRF definiert werden. Hier werden statt den zeitdiskreten Abtastwerten Frequenzbins der Übertragungsfaktor verglichen, um in

einem bestimmten Frequenzbereich eine Aussage über die relative Veränderung einer Impulsantwort treffen zu können.

$$SDRF = 10 \log_{10} \frac{\sum_{k=k_1}^{k_2} \{ |H_1[k]| \}^2}{\sum_{k=k_1}^{k_2} \{ |H_1[k] - H_2[k]| \}^2} \quad [dB] \quad (3.72)$$

k_1 Untere Grenze des zu vergleichenden Frequenzbereichs

k_2 Obere Grenze des zu vergleichenden Frequenzbereichs

Um die tatsächlichen Frequenzen zu erhalten, muss mit der Frequenzauflösung der DFT ($\frac{1}{T_{RIR}}$) multipliziert werden.

$$f_{u,o} = k_{1,2} \cdot \frac{1}{T_{RIR}} \quad (3.73)$$

T_{RIR} Zeitfenster der Impulsantwort für die DFT in Sekunden

Das SDR bzw. SDRF beschreiben nun genau die Qualität der bestenfalls zu erreichenden Echokompensation. Um von (3.5) auf das SDR zu kommen, muss (3.5) lediglich quadriert, summiert, auf die Energie der tatsächlichen RIR bezogen und schließlich logarithmiert werden.

Ein SDR von beispielsweise 30 dB bedeutet für den Anwendungsfall der Echokompensation eine Unterdrückungsleistung von -30 dB des Echosignals (dies gilt jedoch nur als Gesamtwert über den gesamten Frequenzbereiche).

Will man bestimmte Frequenzbereiche bezüglich der Echokompensation genauer untersuchen, so verwendet man das SDRF. Das SDR bzw. SDRF wird nun als zentrale Größe zur Beurteilung von Raumimpulsantworten verwendet.

Ein SDR-Wert von 0 dB bedeutet für die Echokompensation keinerlei Verbesserung der Rückkopplung. Negative SDR-Werte verschlechtern sogar die Echosituation durch Einsatz des Kompensators.

Damit also die Echokompensation gut funktioniert, sollte das SDR der wahren Impulsantwort im Bezug auf die zur Echounterdrückung verwendete Kompensationsimpulsantwort so groß wie möglich sein.

Das messtechnisch beobachtete SDR von Raumimpulsantworten wird in Kapitel 5 und 6 genau behandelt.

3.2 Anwendung in der Aufnahmetechnik

Das Signal jeder Mikrofonaufnahme wird bei der Wiedergabe über Lautsprecher in einem anderen Raum mit den beiden Übertragungsfunktionen des Aufnahme- und Wiedergaberaumes multipliziert.

Kennt man beide Raumübertragungsfunktionen, wären folgende Modifikationen möglich, um eine möglichst originaltreuer Wiedergabe zu erzielen:

3.2.1 Entzerren des Aufnahme Raumes

Bei Kenntnis der Übertragungsfunktion von der Aufnahmeschallquelle zum Aufnahmefunktion, kann der Aufnahme Raum entzerrt werden. Es wird jedoch vorausgesetzt, dass die Ortskoordinaten der Schallquelle bzw. des Mikrofons zeitlich gleich bleiben.

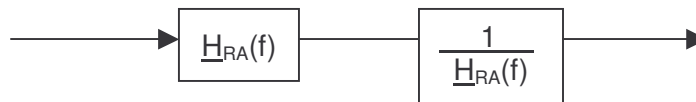


Abb. 3.6: Entzerrung des Aufnahme Raumes

$H_{RA}(f)$ Übertragungsfunktion des Aufnahme Raumes

Vor allem störende Echos aufgrund einer ungünstigen Positionierung des Mikrofons bzw. einer generell schlechten Raumakustik können somit kompensiert werden.

3.2.2 Entzerren des Wiedergaberaumes

Bei Kenntnis der Übertragungsfunktion von Lautsprecher zum Zuhörer kann auch der Wiedergaberaum entzerrt werden. Dabei wird jedoch vorausgesetzt, dass die Abhörposition immer dieselbe ist, sich der Zuhörer also kaum bewegt.

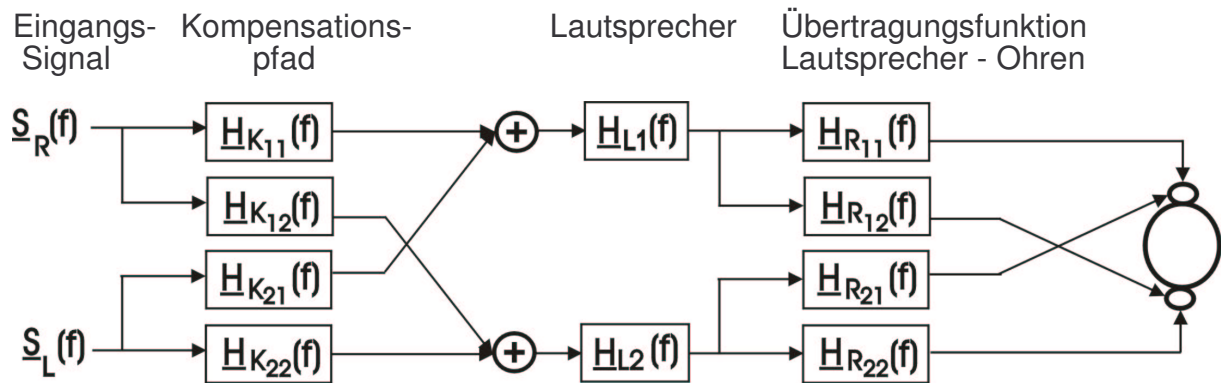


Abb. 3.7: Abbildung jedes Lautsprechers auf beide Ohren

- $H_{R12}(f)$ Übertragungsfunktion des Wiedergaberaumes vom rechten Lautsprecher zum linken Ohr
- $H_{K12}(f)$ Echokompensationspfad, welcher der Übertragungsstrecke $H_{R12}(f)$ entspricht

In Abb. 3.7 werden die Lautsprecher nun mit konstantem Frequenzgang ($H_{L1,2}(f)=1$) angenommen und können daher für die Berechnung vernachlässigt werden.

Wie in Abb. 3.7 ersichtlich, wirkt jeder Lautsprecher auf beiden Ohren des Zuhörers. Diese vier Raumimpulsantworten werden nun in der Matrix $H_R(f)$ zusammengefasst:

$$\underline{H}_R(f) = \begin{bmatrix} \underline{H}_{R11}(f) & \underline{H}_{R12}(f) \\ \underline{H}_{R21}(f) & \underline{H}_{R22}(f) \end{bmatrix}$$

Für eine Entzerrung des Wiedergaberaumes muss die Kompensationsmatrix $H_K(f)$ nun folgende Bedingung erfüllen (Lautsprecher werden vernachlässigt):

$$\underline{H}_K(f)\underline{H}_R(f) = E$$

E Einheitsmatrix

Für die gesuchte Kompensationsmatrix $H_K(f)$ erhält man somit:

$$\underline{H}_K(f) = \underline{H}_R^{-1}(f) = \begin{bmatrix} \underline{H}_{K11}(f) & \underline{H}_{K12}(f) \\ \underline{H}_{K21}(f) & \underline{H}_{K22}(f) \end{bmatrix}$$

3.2.3 Entzerren des Aufnahme- und Wiedergaberaumes

Bei Kenntnis der Übertragungsfunktionen des Aufnahme- und Wiedergaberaumes, können auch beide Räume entzerrt werden und durch ein beliebiges Hallprogramm bzw. eine spezielle Raumimpulsantwort ersetzt werden. Es gelten klarerweise die Voraussetzungen von 3.2.1 und 3.2.2. bzw. deren Einschränkungen.

3.3 Verwendung als Hallprogramm

Eine RIR kann auch verwendet werden, um ein Signal zu verhallen. Dabei wird nicht nur eine Raumart (Kirche, Oper, ...) durch ein Hallprogramm simuliert, sondern es kann ein Modell eines speziellen Raums durch die Kenntnis seiner Raumimpulsantwort sehr genau dem Signal aufgeprägt werden. Der nötige Berechnungsschritt wird über die Faltung aus (2.8) realisiert.

Messtechnik

Die Messungen der Raumimpulsantworten wurden im Cube des IEM, Inffeldgasse 10 der KUG durchgeführt. Dabei ist Cube die Bezeichnung für den Aufnahmeraum des IEM. Die Abmessungen des Cubes sind im Anhang A1 und A2 zu sehen. Die Nachhallzeit des Raumes beträgt ca. 0,7 Sekunden und war bereits vor der Messung bekannt.

1.8 Einteilung der Raumverhältnisse

Die Messungen gliederten sich prinzipiell in 2 Teile:

1. Stationäre Messung: einziger Systemparameter, der sich veränderte, war die Temperatur
2. Nichtstationäre Messung: Bewegung eines Einrichtungsgegenstandes bzw. Veränderung der Mikrofonposition

Die Gesamtheit aller einzelnen Messungen, die teilweise an verschiedenen Tagen durchgeführt wurden, ist nicht miteinander vergleichbar, da z.B. Einrichtungsgegenstände während dieser Zeit umgestellt wurden bzw. sich die Raumtemperatur verändert hat.

Der Änderung der Raumtemperatur würde man anfänglich eher weniger Bedeutung schenken, es hat sich jedoch herausgestellt, dass die Temperatur einen sehr großen Einfluss auf die RIR hat.

1.8.1 Stationäre Verhältnisse

Die Messung für stationäre Verhältnisse sollte abklären, wie reproduzierbar Impulsantworten überhaupt messbar sind.

Strenggenommen lässt sich jedoch keine der Messungen in diese Kategorie einordnen, da die RIR über längere Zeit aufgrund von Temperaturveränderungen (und anderen nicht genau identifizierbaren Drifterscheinungen) stets als zeitvariantes System betrachtet werden müsste.

Ist jedoch der Raum relativ temperaturstabil, bzw. die Messzeit von so kurzer Dauer, dass die Drifterscheinungen kaum Einfluss auf die Messung nehmen, kann der Raum als stationär beschrieben werden.

Als stationäre Messungen werden in der Folge daher jene bezeichnet, wo keinerlei Bewegung im Raum stattfindet.

1.8.2 Nichtstationäre Verhältnisse

Bei den Messungen für nichtstationäre Verhältnisse im Raum sollten die Auswirkungen eines bewegten Objektes auf die Impulsantwort gemessen werden. Diese könnten zum Beispiel die Bewegung einer Person, eine Handbewegung, das Verrücken eines Sessels oder das Öffnen eines Fensters sein.

Um diese zeitvarianten Veränderungen nachvollziehbar und reproduzierbar zu erfassen, wurden mögliche Änderungen im Raum durch eine sich linear mit konstanter Geschwindigkeit bewegend Schachtel simuliert. Aus den dabei gemessenen Daten lässt sich ca. alle 1,5 Sekunden (siehe Gleichung (2.28)) ein neuer RIR-Wert generieren, der dann einer bestimmten Position des bewegten Objekts zugeordnet werden kann. Die von der fahrenden Schachtel hervorgerufenen Veränderungen der RIR sollten so gering sein, dass eine spätere Analyse und im weitern die Prädiktion der zeitvarianten RIR möglich ist, jedoch deutlich die Auswirkungen einer Temperaturdrift auf die RIR übersteigen.

Es wird dabei versucht, die Raumimpulsantwort des zeitvarianten Systems "Raumimpulsantwort" unter quasistationären Bedingungen zu ermitteln. Damit innerhalb einer MLS-Periode von einem quasistationären Fall ausgegangen werden kann, wurde die Geschwindigkeit der Schachtel entsprechend klein gewählt. Als Anhaltspunkt für die Geschwindigkeit wurde dabei die Annahme getroffen, dass sich das Hindernis während einer Messperiode für eine Impulsantwort nur um 1/100 der kleinsten zu messenden Wellenlänge bewegen durfte. Dies würde bei 20 kHz einer Geschwindigkeit von 0,115 mm/sec entsprechen. Die Messdauer wäre bei dieser Geschwindigkeit für 2 m ca. 5 Stunden. Bei dieser Vorgangsweise würde man sich jedoch durch die Temperaturdrift einen viel größeren Fehler einhandeln, als man durch das strenge Einhalten der quasistationären Bedingungen gewinnt. Es ist somit besser, einen vernünftigen Kompromiss zwischen dem Einhalten der quasistationären Voraussetzung für einen möglichst großen Frequenzbereich und einer möglichst kurzen Messzeit zu finden. Für diesen Kompromiss wurden schließlich Geschwindigkeiten von 0,43 – 1,22 mm/sec gewählt.

Bei der höchsten verwendeten Geschwindigkeit bewegt sich das Hindernis also um 1/10 der kleinsten zu messenden Wellenlänge. Bei der höchsten für Breitbandsprache relevanten Frequenz von 7 kHz entspricht die Bewegung nur noch 1/30 der Wellenlänge.

1.9 Allgemeiner Überblick der Messungen

Bei den Messungen wurden MLS Signale je nach Anwendungsfall mit verschiedener Länge abgespielt (zw. 1 und 27 min) und mit einer Abtastrate von 44,1 kHz aufgenommen. Die Zahl der berechneten Raumimpulsantworten, erhält man anhand Division zwischen der Gesamtzeit des MLS-Signals und der MLS-Periodendauer.

Eine Messperiode der MLS beträgt für eine 16 bit MLS-Folge 65535 Abtastwerte, was einer Zeitdauer von 1,486 sec bei $f_s = 44100$ Hz entspricht.

Geht man bei der RIR-Berechnung von nicht überlappenden FFT Frames aus, wird pro MLS-Periode eine RIR berechnet.

Die Aufnahme der über den Raum gefilterten MLS erfolgte über Harddisc-Recording-Systeme, sowohl auf einem Macintosh Rechner mit "Pro Tools" [Lit. 12] als auch auf einem PC unter "Linux", wo die Aufnahme über "Pd" (Pure data) [Lit. 13] gesteuert wurde. Unter "Pro Tools" wurden die ersten Aufnahmen durchgeführt. Hier mussten jedoch für die weitere Bearbeitung oder Archivierung der Aufnahmedaten auf einem PC die Spuren zuerst gebounced¹ (nur in Echtzeit der Aufnahmespuren möglich) und dann ins PC-Wave-Format (higher and lower Byte vertauschen) konvertiert werden.

Für kleinere Datenmengen bei der Aufnahme war das auch kein großer Aufwand. Bei der Arraymessung fielen aber ca. 15 GB Daten an, wo dies sehr wohl ein großer Mehraufwand gewesen wäre.

Die Arraymessung wurde deshalb mit "Pd" durchgeführt, da hier die Aufnahmedaten in ein Mehrspur-Wavefile gespeichert werden konnten und kein weiterer Bearbeitungsaufwand für direkte Nutzung unter "Matlab" am PC bestand.

Während sämtlicher Aufnahmen befand sich natürlich niemand im Raum, da die Anwesenheit (und unwillkürliche Bewegung) einer Person auch Auswirkungen auf die RIR hat.

Für die Messungen wurden Mikrofone mit Kugelcharakteristik verwendet, um sämtliche Schalleinfallrichtungen möglichst gleichförmig bewertet aufzunehmen.

Die Aufstellung des Mikrofons erfolgte prinzipiell außerhalb des Hallradiuses im Raum. Der Grund liegt darin, dass genügend Reflexionsenergie gemessen werden soll, und nicht nur die direkte Übertragungstrecke vom Lautsprecher zum Mikrofon.

Abb. 4.1a und 4.1b zeigen zwei RIRs aus später durchgeführten Messungen, die in 1 bzw. 10 Meter Abstand zum Lautsprecher durchgeführt wurden. Abb. 4.1a zeigt eine RIR, welche innerhalb des Hallradius aufgenommen wurde und nur wenig Information über Raumreflexionen beinhaltet. Diese Impulsantwort wäre für die Untersuchung von zeitvarianten RIRs nicht sehr geeignet. In Abb. 4.1b wurde hingegen die RIR deutlich außerhalb des Hallradius aufgenommen. Hier sind die Reflexionen im Vergleich zum Direktschall sehr gut sichtbar.

¹ Bouncen: Berechnung des Stereofiles aus einer Mehrspuraufnahme bei Harddiscrecording-Systemen

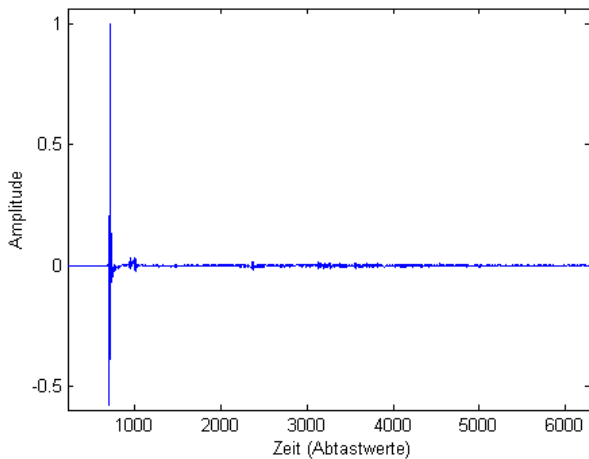


Abb. 4.1a: RIR innerhalb des Hallradiuses

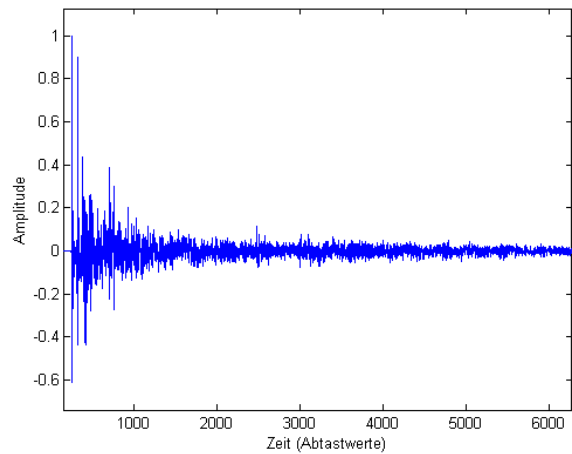


Abb. 4.1b: RIR außerhalb des Hallradiuses

Aus der Formel nach Sabine (1.21) ergibt sich für den Hallradius:

$$r_h = 1,59 \text{ m}$$

Das Raumvolumen V wurde abgeschätzt mit (siehe Anhang A1 und A2):

$$V \approx G \cdot h_{\text{mittel}} = 115 \text{ m}^2 \cdot 4,75 \text{ m} = 546,25 \text{ m}^3$$

$$G \approx 10 \text{ m} \cdot 11,5 \text{ m} = 115 \text{ m}^2$$

Der Wert des Hallradius ist stark frequenzabhängig und sollte nur als Richtwert verstanden werden. Das Mikrofon wurde deshalb für sämtliche Messungen nicht stationärer Raumimpulsantworten um mehr als das Doppelte des Hallradius vom Lautsprecher entfernt aufgestellt.

Betrachtet man Abb. 4.1, so ist zu erkennen, dass die Komplexität der Raumimpulsantwort abnimmt, je kleiner der Abstand Lautsprecher zu Mikrofon wird. Im Anwendungsfall der Konferenzschaltung, kann jedoch dieser Abstand nicht beliebig klein gewählt werden. Wird nämlich die Empfindlichkeit des Mikrofons so eingestellt, dass sich der Sprecher außerhalb des Hallradiuses befindet (es kann beispielsweise von sämtlichen Sitzplätzen in einem Lehrsaal übertragen werden), so würde es zu einer Übersteuerung des Mikrofonkreises kommen, falls ein Sprecher in geringem Abstand in das Mikrofon spricht (Abstand von Sprecher zu Mikrofon kleiner als der Hallradius). Genau dieses Problem tritt jedoch auf, falls der Lautsprecher zu nahe an das Mikrofon gestellt wird.

1.10 Messgeräte

Die für sämtliche Messungen verwendeten Geräte werden wie folgt angeführt:

4.3.1 Studiogeräte

- Lautsprecher Genelec 1032 A (Aktivbox)
- 6 Messmikrofone (Kugelcharakteristik), 1 Mikrofon für Temperaturkompensation
- Harddisc-Recording-System ("Pd"), Harddisc-Recording-System ("Pro Tools")
- RME A/D-D/A-Wandler
- Mikrofonvorverstärker "Midiman"
- Digitalmischpult "Makie D8B"
- Mikrofonkalibriergerät "Brüel & Kjaer"

Dazugehörige technische Daten siehe Anhang C.

4.3.2 Messschlitten

Um nun einerseits die Schachtel zu bewegen und andererseits das lineare Mikrofonarray zu transportieren, wurden folgende Geräte benötigt:

- 2 Messschlitten
- 2 Antriebsmotoren
- 2 fahrbare Untersätze (transportieren die Schachtel bzw. das Mikrofonarray)

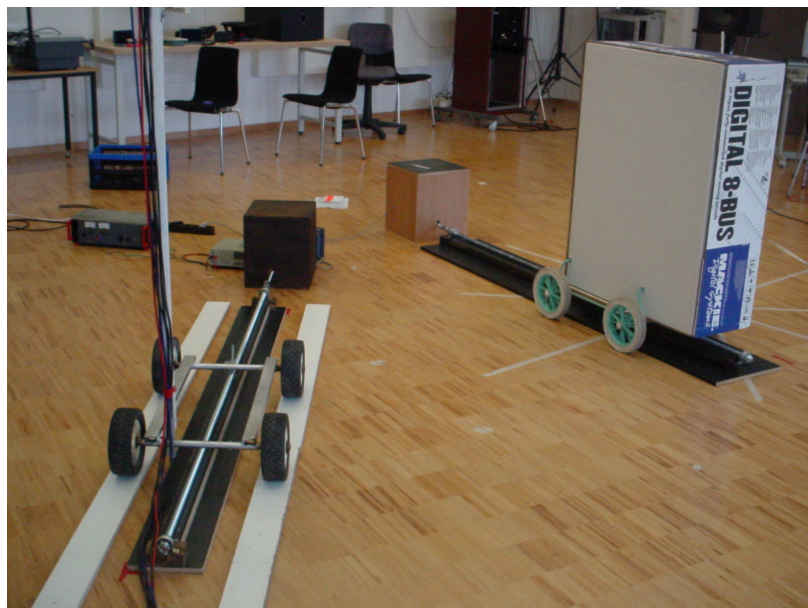


Abb. 4.2: Messschlitten 1 (rechts) und 2 (links)

In Abb. 4.2 sind beide Messschlitten zu sehen, wobei das linke Schlittensystem (Messschlitten 2) nur für die Arraymessung benötigt wurde. Bei den restlichen nicht

stationären Messungen wurde statt des linken Schlittens mit Mikrofonarray ein fixplaziertes Mikrofon aufgestellt. Nur bei der Arraymessung wurden die genauen Ortskoordinaten der Geräte und des Raumes aufgenommen, bei den restlichen nur einige Größen, wie der Abstand von Lautsprecher zu Mikrofon oder die Höhe des Lautsprechers bzw. des Mikrofons.

Aufgrund hoher Kosten eines Präzisions-Spindeltriebs für 2 m Länge wurden die für die Messungen nötigen Geräte selbst gebaut.

Für die Messschlitten wurden 2 Gewindestangen M20 als Spindelersatz verwendet. Eine darauf befindliche Mutter diente dabei als Vorschub, welche dann jeweils das fahrbare Untergestell linear in Richtung der Gewindestange zog.

Als Antriebsmotoren wurden Kfz-Scheibenwischermotoren verwendet. Diese wurden zwecks Schalldämmung in eine mit Schaumstoff gefüllte Holzkiste eingebaut. Die Motoren boten aufgrund des großen Drehmoments bei geringer Geräuschentwicklung gute Voraussetzungen für die Messungen. Die Drehzahl wurde dabei über die Versorgungsspannung eingestellt.

Die technischen Daten der Schlittensysteme können wie folgt zusammengefasst werden:

- 2 Drehrichtungen
- Drehzahlbereich von 0,2 – 1,5 Umdrehungen pro Sekunde
- Spannungsversorgung ca. 5 – 15 V (max. 2 A)
- Max. Messbereich 2 m
- Gewindesteigung 2,5 mm/U (M20)
- Spiel der Gewindemutter < 0,5 mm

Die Hysterese der Gewindemutter bei der Drehrichtungsänderung spielte keine Rolle, da die genaue Positionierung des Hindernisschlittens bzw. des Arrayschlittens immer aus einer Richtung erfolgte.

1.11 Durchgeführte Messungen

Im folgenden werden die durchgeführten Messungen beschrieben. Die Auswertung dieser Messdaten erfolgt für die stationären Messungen in Kapitel 5, für die nichtstationären Messungen in Kapitel 6.

1.1.1 Stationäre Messungen

Diese Messungen waren relativ einfach durchzuführen, da hier nur der Lautsprecher und ein Mikrofon benötigt wurden.

Folgende 4 Messungen wurden vorgenommen:

- Langzeitmessung
- Unterschiedliche Abstände
- Verschiedene Pegel
- Sprache als Anregungssignal

4.4.1.1 Langzeitmessung

Hier sollte die Reproduzierbarkeit einer RIR untersucht werden, wobei die Messung ca. 16 Minuten dauerte.

Es wurden insgesamt 3 Blöcke von je 5 Minuten abgespielt, die durch 2 Pausen von 2 bzw. 23 Sekunden getrennt waren. Die Signalpausen wurden eingefügt, um ihre eventuelle Auswirkungen auf RIRs kurz vor bzw. nach der Pause zu sehen.

4.4.1.2 Unterschiedliche Abstände

Bei diesen Messungen wurde das Mikrofon in verschiedenen Entfernungen im Bereich von 0,5 – 10 Meter zum Lautsprecher aufgestellt. Dabei betrug die Messdauer je Abstand ca. 1 Minute.

4.4.1.3 Verschiedene Pegel

Aufgabe dieser Messung war es, den Einfluss des Lautsprecherpegels auf das SDR zu untersuchen. Dabei wurde mit 5 verschiedenen Pegelstufen (-20, -12, -6, -3, 0 dB) des Lautsprechers gemessen.

Bei jeder Pegelstufe (5 Minuten Messdauer) wurden zusätzlich 4 Signalpausen in der Dauer von 2 – 20 Sekunden eingeschoben, um einen eventuellen Erwärmungseffekt des Lautsprechers nachzuweisen, welcher anhand von Ergebnissen der Messung gemäß Punkt 4.4.1.1 vermutet wurde.

4.4.1.4 Sprachsignal als Anregung

Bei dieser Messung wurde der Raum zusätzlich zur MLS auch mit einem Sprachsignal beschallt.

Block	Signal	Dauer [sec]
1	MLS	0 – 58
2	Sprache	61 – 200
3	MLS	204 – 263
4	Sprache	267 – 408
5	MLS	415 – 472

Tab. 4.1: Einteilung der Beschallungsdaten

Es sollte später versucht werden, aus den Sprachdaten die Impulsantwort zu bestimmen. Als Referenzgrößen sollten dabei die aus den MLS-Daten generierten RIRs verwendet werden.

1.1.2 Nichtstationäre Messungen

Dabei handelt es sich um die Messungen bei nichtstationären Raumbedingungen.

Es können 3 Messungen unterschieden werden:

- Bewegte Schachtel (Hindernis), fixpositioniertes Mikrofon
- Bewegtes Mikrofon
- Bewegte Schachtel mit Array aufgenommen

4.4.2.1 Bewegte Schachtel mit fixpositioniertem Mikrofon

Hier wurde während der Messung ein Hindernis im Raum bewegt.

Der Abstand zwischen Lautsprecher und Mikrofon betrug dabei 5,32 m. Der Messschlitten mit Hindernis wurde ca. in der Mitte zwischen Lautsprecher und Mikrofon aufgestellt. Die Aufstellung der Geräte war dabei ähnlich, wie in Abb. 4.11 zu sehen ist. Im Bereich des Mikrofonarrays befand sich jedoch stattdessen ein fixplaziertes Mikrofon.

Der Messschlitten wurde so zwischen Lautsprecher und Mikrofon plaziert, dass am Beginn der Messung die erste Bodenreflexion noch beim Mikrofon ankam, jedoch mit Weiterbewegung der Schachtel immer mehr Bodenreflexion abgeschattet wurde, bis sich die Schachtel schließlich komplett zwischen Lautsprecher und Mikrofon befand.

Das vollständige Abschatten der ersten Bodenreflexion gilt jedoch nur für Wellenlängen, die kleiner sind als das Hindernis (siehe Abb. 1.5). Tiefe Frequenzen wurden hingegen um das Hindernis gebeugt.

Der Direktschall wurde dabei nicht beeinträchtigt, da sich die Schachtel mit 122 cm Höhe ca. 20 cm unter der Direktverbindung Lautsprecher und Mikrofon befand.

Der Messablauf sah dabei folgendermaßen aus:

1. Start von Messschlitten 1 mit Hindernis
2. Verlassen des Raumes
3. Beginn der MLS-Beschallung des Raumes
4. Start der Aufnahme

Die Messdauer betrug ca. 20 Minuten. Es wurde dabei die Start- und Endposition des Messschlittens, sowie die dafür benötigte Zeit aufgenommen. Die daraus berechnete Durchschnittsgeschwindigkeit für den Messschlitten ergab 0,934 mm/sec.

Aktion	Position	Zeit [sec]
Start Messschlitten 1	0	0
Stop Messschlitten 1	116,5	1247,79

Tab. 4.2: Start- und Stopposition des Messschlittens

4.4.2.2 Bewegtes Mikrofon im unveränderlichen Raum

Bei der zweiten nicht stationären Messung wurde kein Hindernis verwendet. Statt dessen wurde das Mikrofon samt dem Ständer auf den fahrbaren Untersatz gestellt und bewegt.

Die Änderung der Impulsantwort ist bei dieser Messung am größten, da sich hier in Bezug auf das Mikrofon (falls man als Bezugssystem das Mikrofon betrachtet) nicht nur eine Schachtel sondern der ganze Raum bewegt. Die Geschwindigkeit des Messschlittens wurde deshalb mit 0,43 mm/sec sehr gering gewählt.

Die Stecke des Mikrofons von 62 cm wurde in einer Messzeit von 1444,4 Minuten zurückgelegt. Der Abstand des Lautsprechers zum Mikrofon betrug ca. 3,26 m (veränderte sich während der Bewegung leicht, da das Mikrofon seine Position änderte).

4.4.2.3 Bewegte Schachtel mit Array aufgenommen

Es sollten mittels eines Mikrofonarrays (96 Sensorpunkte) die Änderungen erfasst werden, die durch das bewegte Hindernis hervorgerufen wurden. Da die Sensorzahl des Arrays sehr groß gewählt wurde, wurde das Array nicht tatsächlich mit 96 Mikrofonen aufgebaut. Statt dessen wurden 6 Mikrofone in y-Richtung angeordnet und ausgehend von der Startposition 15 mal in x-Richtung verschoben, was zu den insgesamt $16 \times 6 = 96$ Sensorpositionen führt.

Da die Bewegung der Schachtel (ca. 2 Meter) jeweils nur mit 6 Sensorpositionen aufgezeichnet werden konnte, musste das Hindernis die gleiche Bewegung für jede, der 16 x-Positionen des linearen Arrays, ausführen. Somit wird das 2-dimensionale Mikrofonarray durch 16 Einzelmessungen eines linearen 6er-Arrays realisiert.

Die genauere Beschreibung des Mikrofonarrays bzw. der Durchführung der Messung erfolgt in Punkt 4.5 und 4.6.

1.12 Beschreibung des Mikrofonarrays

Bevor die Messung mit einem Mikrofonarray beschrieben wird, sei zunächst die prinzipielle Wirkungsweise eines Arrays vorgestellt.

Will man ein Schallsignal einer bestimmten Richtung aufnehmen, so verwendet man ein Mikrofon, welches eine Richtcharakteristik besitzt. Damit wird die Richtung der Nutzschallenergie bevorzugt aufgenommen, Störgeräusche aus anderen Richtungen dagegen unterdrückt. Weist jedoch das verwendete Mikrofon Kugelcharakteristik auf, so ist keine Bevorzugung einer Aufnahme­richtung möglich.

Werden mehrere Mikrofone in einem Arrayverbund verwendet, so kann auch über Mikrofone mit Kugelcharakteristik eine bestimmte Raumrichtung bevorzugt aufgenommen werden.

1.12.1 Kontinuierliche Apertur

Die Apertur kann bildhaft als Fenster betrachtet werden, durch welches wellencharakteristische Signalenergie aufgenommen wird. Bei Aufnahme der Signalenergie über einen punktförmigen Sensor wird diese ungerichtet aufgezeichnet. Wird sie jedoch über einen stabförmigen Sensor (z.B. Stabantenne) aufgenommen, kommt es zu einer Richtwirkung des Sensors. Mathematisch wird das Ortsfenster (Geometrie des Sensors) über die Aperturbelegung $w(x)$ ausgedrückt. Der Parameter k_x stellt dabei die in Gleichung (1.6) erwähnte Wellenzahl in x -Richtung, dar.

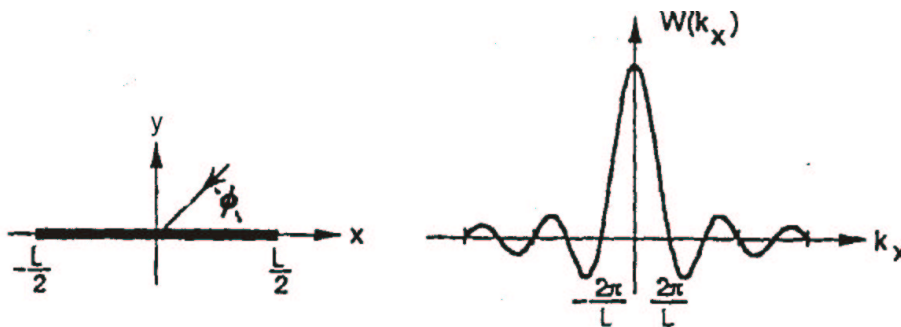


Abb. 4.3: Aperturbelegung und Richtdiagramm [Lit. 14]

$$w(x) = \text{rect}\left(\frac{x}{L}\right) \longleftrightarrow W(k_x) = L \cdot \text{sinc}\left(k_x \frac{L}{2}\right) \quad (4.74)$$

Der linke Teil in Abb. 4.3, entspricht dabei der geometrischen Anordnung des Sensors im Raum. Die rechte Seite stellt die Fouriertransformierte der Aperturbelegung dar. Um dieses Bild zu verstehen, sei zunächst Abb. 4.4 betrachtet.

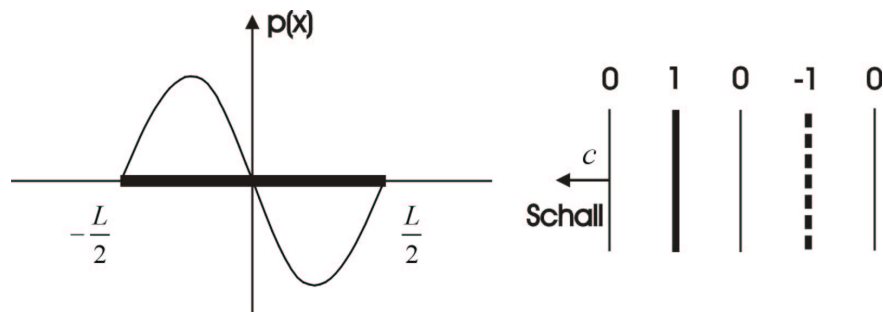


Abb. 4.4: Schalldruckverlauf bei $\phi = 0^\circ$

Trifft eine ebene Schallwelle mit $\lambda = L$ aus der Richtung $\phi = 0^\circ$ auf den linearen Sensor, so ergibt sich der aufgenommene Schalldruck mit:

$$\int_{-L/2}^{L/2} (-\sin x) dx = 0 \tag{4.75}$$

Selbiges ergibt die Integration über Wellenlängen von λ/n $n \in \mathbb{N}$. Diejenigen Frequenzen, welche also das Integrationsergebnis von 0 liefern, entsprechen den Nullstellen der Sinc-Funktion in Abb. 4.3.

In Abb. 4.5 ist nun der Schalldruckverlauf $p(x)$ zu sehen, welcher bei $\phi = 60^\circ$ am kontinuierlichen Sensor anliegt. Bei der Beschreibung des 2-dimensionalen Wellenfeldes wird dementsprechend auch die Wellenzahl k 2-dimensional.

$$k = \begin{bmatrix} k_x \\ k_y \end{bmatrix}$$

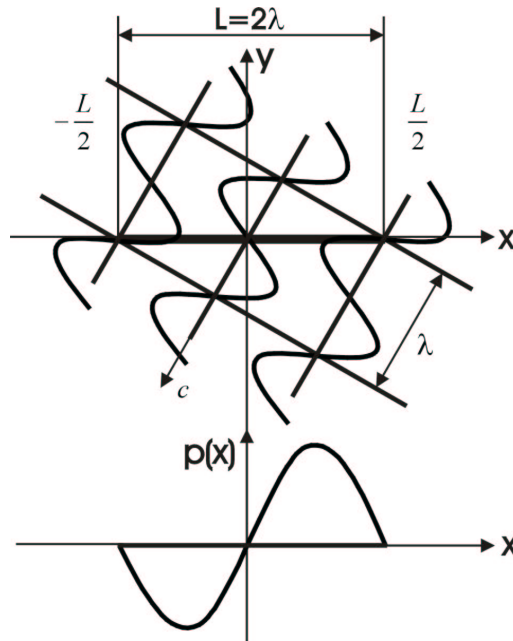
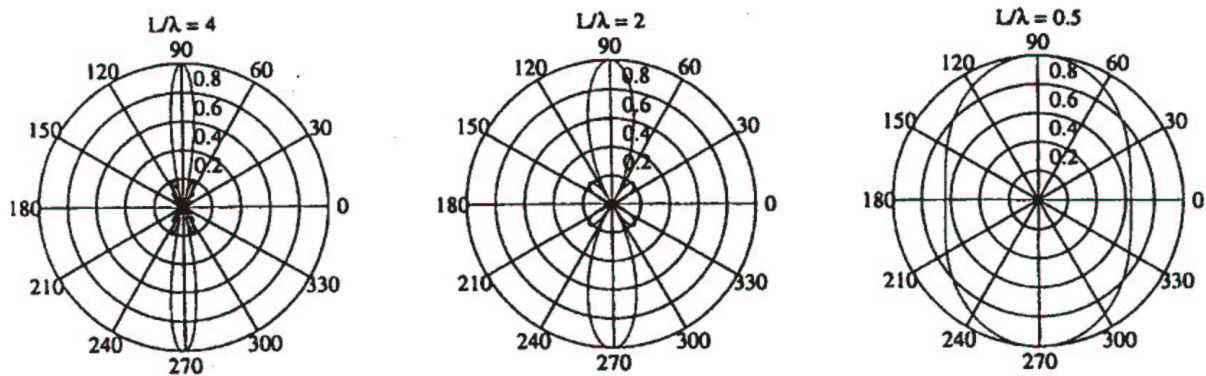


Abb. 4.5: Schalldruckverteilung bei schrägem Einfall

Bei diesem Winkel löscht sich ebenfalls wieder positive und negative Druckhalbwelle aus. Abb. 4.6 entspricht also dem mittleren Richtdiagramm bei Schalleinfall 60° .

Werden nun sämtliche Schalleinfallrichtungen ϕ berücksichtigt, so ergibt sich die winkelabhängige Darstellung der Richtcharakteristik einer linearen Apertur.



dargestellt. Bei höheren Frequenzen wird also die Richtwirkung der linearen Apertur selektiver bezüglich einer Richtung, hat hingegen aber mehr Nebenkeulen. Die Zahl der auftretenden Richtungsmaxima z_{\max} zwischen $\phi = 0 - 90^\circ$ kann wie folgt angegeben werden:

$$z_{\max} = \text{Round}\left(\frac{L}{\lambda}\right) \tag{4.76}$$

Die Zahl der Nullstellen z_0 beträgt:

$$z_0 = \text{Trunc}\left(\frac{L}{\lambda}\right) \quad (4.77)$$

Round und *Trunc* bezeichnen dabei Runden bzw. Abschneiden auf ganzzahlige Werte.

Das Richtdiagramm in Abhängigkeit vom Einfallswinkel ϕ und der Frequenz f wird wie folgt angegeben:

$$W(f, \phi) = \left| \frac{\sin\left(\frac{L \cdot f \cdot \cos \phi}{c} \pi\right)}{\left(\frac{L \cdot f \cdot \cos \phi}{c}\right) \pi} \right| \quad (4.78)$$

1.12.2 Diskrete Apertur

Die beschriebene kontinuierliche Apertur wird nun durch punktförmige Sensoren (Mikrofone) ersetzt, welche den gleichmäßigen Abstand d aufweisen.

Es sei nun eine ungerade Anzahl M von Mikrofonen angenommen. Dabei stellen die einzelnen Sensoren eine Abtastung der Apertur dar.

Die Fourier-Transformation der diskreten Apertur (4.7) führt zu dem, in Abb. 4.6 dargestellten Richtdiagramm, welches nun periodisch in $\frac{2\pi}{d}$ ist.

$$w(x) = \sum_{m=-\frac{M-1}{2}}^{\frac{M-1}{2}} \delta(x - md) \quad (4.79)$$

Durch Fourier-Transformation erhält man:

$$W(k_x) = \sum_{m=-\frac{M-1}{2}}^{\frac{M-1}{2}} \exp(jk_x md) = \frac{\sin\left(\frac{M}{2} k_x d\right)}{\sin\left(\frac{1}{2} k_x d\right)} \quad (4.80)$$

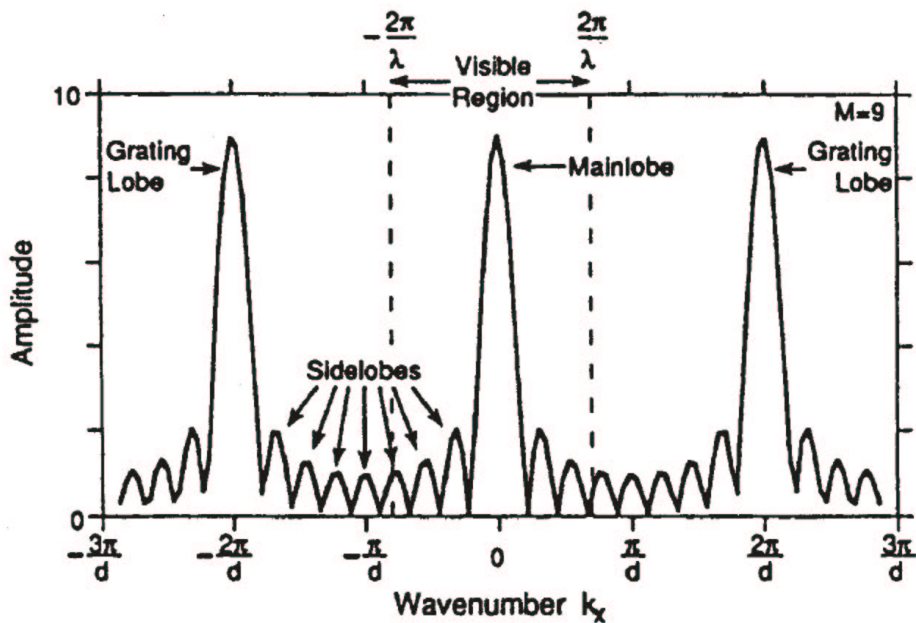


Abb. 4.7: Richtdiagramm eines linearen Mikrofonarrays [Lit. 14]

Um nun Mehrdeutigkeiten des Richtdiagramms in Abb. 4.7 zu verhindern, darf sich der sichtbare Bereich nur von $-\pi/d$ bis $+\pi/d$ bewegen. Dies wird durch die räumliche Abtastbedingung in (4.8) sichergestellt.

$$d \leq \frac{\lambda}{2} \tag{4.81}$$

(4.8) entspricht dabei dem räumlichen Analogon des Shannon'schen Abtasttheorems, um Aliasing zu verhindern.

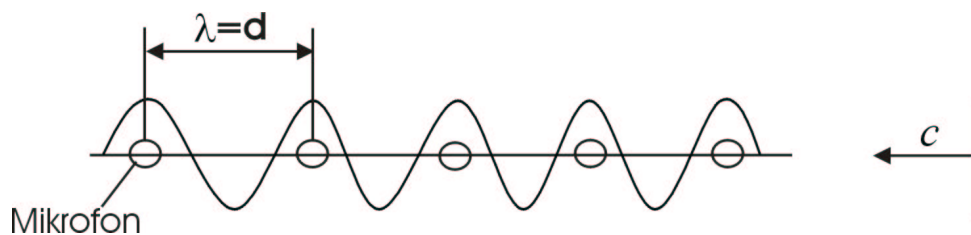


Abb. 4.8: Verletzung der räumlichen Abtastbedingung

Abb. 4.8 zeigt nun eine Verletzung des räumlichen Abtasttheorems. Bei Einfall einer Schallwelle mit $\lambda = d$ aus 0° , erfolgt eine konstruktive Überlagerung sämtlicher Mikrofonensignale. Es zeigt sich also für diese Frequenz das selbe Richtungsverhalten des Arrays, wie eine einfallende Welle aus 90° . Der Grund liegt in der Verschiebung (Aliasing) der Grating-Lobes (Abb. 4.7) in den sichtbaren Bereich.

1.12.3 Beamformer

Das Ausgangssignal eines Arrays erhält man nun durch Summation der (unterschiedlich verzögerten) Signale der einzelnen Mikrofone. Diese Anordnung wird üblicherweise als delay-and-sum Beamformer bezeichnet.

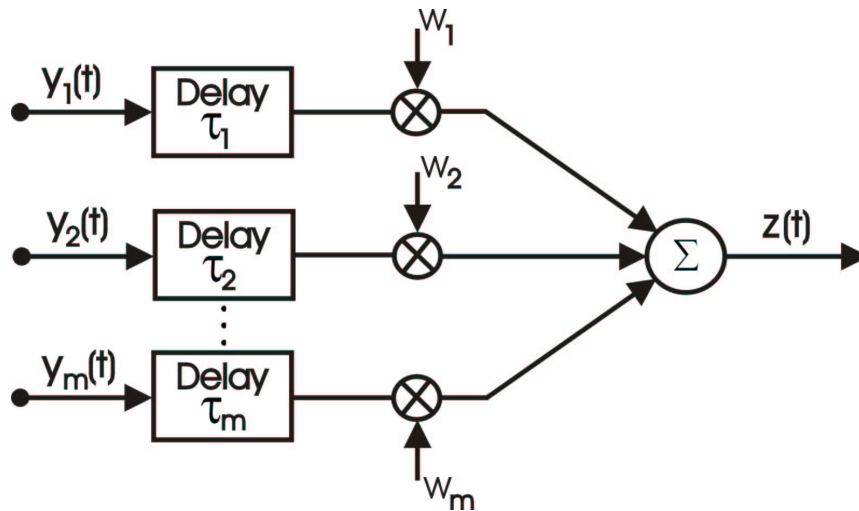


Abb. 4.9: Delay-and-Sum Beamformer

Das Ausgangssignal $z(t)$ ergibt sich mit:

$$z(t) = \sum_{m=1}^M w_m y_m(t - \tau_m) \quad (4.82)$$

Wobei mit der Verzögerungszeit τ_m und den Gewichten w_m die Blickrichtung des Arrays eingestellt werden kann.

Dies hat den Vorteil, dass das Array für verschiedene Blickrichtungen nicht mechanisch umgestellt werden muss.

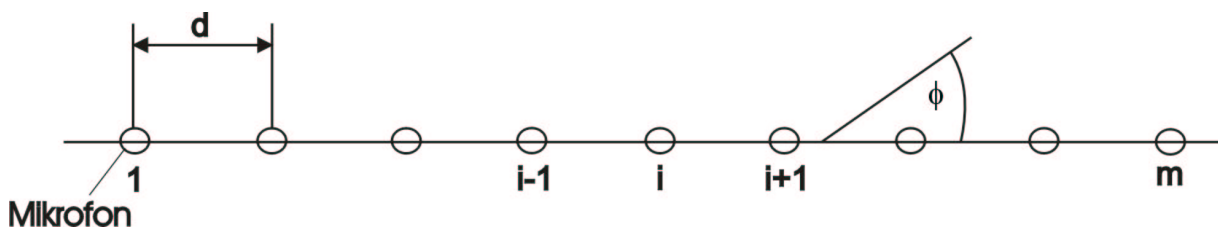


Abb. 4.10: Anordnung der Mikrofone eines linearen Arrays mit konstantem Abstand

Soll das Array nun in Richtung ϕ die maximale Empfindlichkeit aufweisen (die Hauptkeule der Richtcharakteristik richtet sich also in Richtung ϕ), so ergeben sich die benötigten Signalverzögerungen τ_i wie folgt:

$$\tau_i(\phi) = \frac{d \cos \phi}{c} (i-1) \quad (4.83)$$

Es wurde somit gezeigt, dass mit Hilfe eines Arrays eine räumliche Filterung der verschiedenen Einfallsrichtungen ermöglicht wird. Es kann also ein Array-Datensatz durch Verstellen der Verzögerungsglieder später in verschiedenen Schalleinfallsrichtungen analysiert werden.

1.13 Planung des Mikrofonarrays

Wie man anhand von Gleichung (4.8) gesehen hat, hängt also die Geometrie des Arrays mit den Grenzfrequenzen des delay-and-sum Beamformers zusammen.

Dabei wurden die Grenzfrequenzen vor der Messung wie folgt spezifiziert:

- Untere Grenzfrequenz: 100 Hz
- Obere Grenzfrequenz: 4 kHz

Die Grenzfrequenzen wurden zunächst anhand der Breitbandsprachübertragung festgelegt, wobei die obere Grenzfrequenz nachträglich heruntergesetzt wurde, da sich aufgrund der geometrischen Abmessungen eines Mikrofons (Halterung) nur ein gewisser kleinster Mindestabstand realisieren lässt.

Der kleinste Abstand der Mikrofone wurde mit 4,2 cm in y-Richtung realisiert, woraus man eine obere Grenzfrequenz des Arrays von 4083,3 Hz in y-Richtung erhält. Die unteren Grenzfrequenz erhält man über den Zusammenhang, dass die maximale Wellenlänge das Doppelte der Ausdehnung L des Arrays ist.

Die räumliche Ausdehnung des Arrays in y-Richtung mit 174,6 cm führt auf eine untere Grenzfrequenz f_{gu} von 98,2 Hz.

$$f_{gu} = \frac{c}{2L}$$

Bei der unteren Grenzfrequenz ist jedoch die Hauptkeule schon sehr breit (Abb. 4.6, $L/\lambda=0,5$), die räumliche Filterung also nicht mehr sehr selektiv in Blickrichtung des Arrays.

Würde nun dieses Array mit konstantem Abstand d realisiert werden, wäre die benötigte Zahl an Mikrofonen ca. 40 Stück. Dabei würde sich für die untere Grenzfrequenz eine breite bzw. für die obere eine sehr schmale Richtkeule ergeben.

Es zeigt sich also, dass große Wellenlängen im Vergleich zu kleinen, stark überabgetastet werden. Um die unnötige Überabtastung zu reduzieren, wurde das Array nicht mit konstanten Abständen zwischen den Mikrofonen realisiert, sondern die Abstände von Mikrofon zu Mikrofon sukzessiv vergrößert. Dadurch konnte die Anzahl von Mikrofonen reduziert werden. Die Abstände der Mikrofonpositionen sollten außerdem keine gemeinsamen Teiler besitzen, also Primzahlen bzw. irrationale Zahlen sein.

Die y -Abstände der Mikrofone wurden wie folgt berechnet:

$$y_i = a(i) \cdot 4 \text{ cm}$$

$$a = [1 \ 3 \ 7 \ 13 \ 19]$$

Dabei bezeichnet z.B. der Abstand y_5 den Abstand zwischen 5. und 6. Sensor des Arrays. Die realisierten Werte siehe Anhang A3 bzw. B3.

In x -Richtung wurde das Arrays mit einer größeren Sensordichte (Mikrofonpositionen pro Längeneinheit) als in y -Richtung realisiert, weshalb bei der Berechnung der Positionsabstände in x -Richtung eine andere Berechnungsmethode verwendet wurde:

$$x_1 = 4 \text{ cm}$$

$$x_{i+1} = 4 \text{ cm} \cdot 10^{\frac{i}{16}} \quad 2 \leq i \leq 14 \quad (4.84)$$

Während die y -Abstände der Mikrofone tatsächlich aufgebaut wurden, bezeichnen die x -Abstände die Verschiebung des linearen Arrays von einer Teilmessung zur nächsten.

Abb. 4.11 stellt die Aufnahmesituation der Arraymessung dar. Während das lineare Mikrofonarray bei einer Einzelmessung an fixer Position "x" stand, legte die Schachtel bei MLS-Beschallung des Raumes den Weg von ca. 2 Meter zurück. Indem die Mikrofone auf einem fahrenden Wagen montiert waren, konnten die 16 Einzelpositionen der Teilmessungen dadurch sehr präzise eingestellt werden (Positionsfehler < 1 mm)!



Abb. 4.11: Aufnahmesituation der Arraymessung

Sämtliche 96 Sensorpositionen des Arrays sind grundsätzlich nicht direkt mittels SDR vergleichbar, da der kleinste Abstand zwischen 2 Mikrofonen 4,2 cm beträgt. Das SDR liefert sogar negative Werte und man erhält daraus nur mehr die Aussage, dass die 2 betrachteten Impulsantworten im Sinne der Echounterdrückung für den gesamten Frequenzbereich bis 22 kHz nicht mehr ähnlich sind.

Damit also das ganze Array, realisiert durch 16 Einzelmessungen als homogener Datensatz betrachtet werden kann, seien zunächst folgende wichtige Voraussetzungen diskutiert, deren Ergebnisse jedoch erst in Punkt 6.1 zu finden sind.

1.13.1 Stationäre Raumbedingung

Während der 16 Einzelmessungen sollte im Raum nichts verändert werden, abgesehen von der Schachtel, die bei jeder Einzelmessung möglichst die gleiche Bewegung ausführen sollte.

Dies konnte jedoch nicht ganz realisiert werden, da ja auch das 6er-Array jeweils zwischen den Einzelmessungen um eine Position weiterbewegt werden musste, und somit der Messsensor selbst die Messung geringfügig stört. Deshalb wurde auch die Arrayhalterung sehr klein konstruiert, um möglichst eine geringe Beeinflussung des Raumes zu gewährleisten. Zusätzlich wurden vor dem fahrbaren Array-Untergestell Schachteln aufgestellt (siehe Anhang A, Abb. A8), damit zumindest der Direktschall nicht durch die verschiedenen Positionen relevant verändert wurde.

1.13.2 Verstärkungskompensation

Um den Verstärkungsfaktor zu bestimmen, wurde jeweils vor und nach der Arraymessung eine Kalibrierungsmessung durchgeführt.

Um die verschiedenen Verstärkungsfaktoren der einzelnen Mikrofonvorverstärker abzugleichen, wurden dabei die 6 Arraymikrofone vor bzw. nach der Arraymessung mit einem Kalibriergerät gemessen. Die 2. Kalibrierungsmessung wurde ca. 12 Stunden nach der letzten Teilarraymessung (Mikrofonvorverstärker blieb in der Nacht eingeschaltet) durchgeführt, und wurde deshalb auch weniger gewichtet als die erste Kalibrierungsmessung.

1.13.3 Kompensation der verschiedenen Mikrofonfrequenzgänge

Der Frequenzgang der 6 verschiedenen Mikrofone sollte gleich sein. Es wurden 6 qualitativ hochwertige, baugleiche Mikrofone mit Kugelcharakteristik verwendet.

Um nun die Frequenzgänge der Mikrofone zu erhalten, wurde eine Vergleichsmessung durchgeführt. Das Ergebnis dieser Messung lieferte also nicht den absoluten Frequenzgang, sondern die Differenz der Verstärkung als Funktion der Frequenz $|\underline{V}(f)|$ und die Differenz der Phase als Funktion der Frequenz $\arg(\underline{V}(f))$.

Der Messablauf sah folgendermaßen aus. Es wurde das erste Mikrofon in eine Halterung gespannt und dann mit dem MLS Signal beschallt. Die Messdauer wurde dabei mit ca. 1 Minute gewählt, um die Temperaturdrift gering zu halten. Danach wurde das nächste Mikrofon an der gleichen Position eingespannt. Dies war sehr genau möglich, da eine Mikrofonmarkierung als Referenzpunkt diente. Der Fehler kann somit im Zehntelmillimeterbereich angenommen werden.

Da also jedes Mikrofon die Impulsantwort an derselben Position aufgenommen hat, sollte später durch Vergleich die wechselseitige Abweichung der einzelnen Übertragungsfunktionen $\underline{V}(f)$ bestimmt werden können.

1.13.4 Kompensation des Geschwindigkeitsfehlers

Da die Schachtel für jede 6er Arrayposition die gleiche Bewegung ausführen sollte, mussten etwaige Fehlerquellen soweit wie möglich minimiert werden.

Dazu wurden am Boden Klebestreifen angebracht, um den fahrbaren Untersatz bei jeder Messung auf dieselbe Startposition zu stellen. Um verschiedene Geschwindigkeiten der Teilmessungen zu kompensieren, war es notwendig, bei jeder Messung den zurückgelegten Weg und die dafür benötigte Zeit zu messen.

Man kann die daraus errechneten mittleren Geschwindigkeiten in einem Vektor v_m zusammenfassen:

$$v_m = \begin{bmatrix} v_{1m} \\ \vdots \\ v_{16m} \end{bmatrix} \quad (4.85)$$

Daraus werden die relativen Unterschiede der einzelnen Geschwindigkeiten berechnet. Die Korrekturfaktoren erhält man nun indem der kleinste Wert aus v_m durch die Pseudo-Inverse von v_m dividiert wird. Die Geschwindigkeitsunterschiede können durch einen unterschiedlichen Offset O_k bei der RIR-Berechnung kompensiert werden, siehe Kapitel 2.3.2.2. Die kompensierten Offset-Werte werden folgend berechnet:

$$O_k = 65535 \cdot \frac{\min(v_m)}{v_m} \quad (4.86)$$

Damit erfolgt effektiv für alle 16 Teilmessungen die Berechnung der RIR an denselben Ortskoordinaten, die durch die langsamste Messung vorgegeben werden. Bei der Arraymessung betrug die max. Geschwindigkeitsabweichung 3,7 %. Der lineare Geschwindigkeitsfehler kann somit kompensiert werden. Es verbleibt somit noch der Fehler, der durch die Gleichlaufschwankungen der Einzelmessungen verursacht wurde, welcher aber im Vergleich zum linearen Geschwindigkeitsfehler weniger relevant ist und somit vernachlässigt wird.

In den Abbildungen 6.6 ist zu sehen, dass dieser Fehler vor allem im tieffrequenten Bereich ohnehin keine gravierenden Auswirkungen hat.

1.13.5 Protokollierung der Temperaturveränderung

Da eine vollständige Arraymessung ca. 12 Stunden dauert, musste eine Temperaturdrift sogar im °C-Bereich einkalkuliert werden. Um diese später abschätzen und gegebenenfalls im Datensatz kompensieren zu können, wurde ein fixplaziertes Mikrofon aufgestellt, welches zum Zweck der Temperaturprotokollierung verwendet wurde.

Kennt man nämlich von jeder Teilmessung die RIR eines Mikrofons an fixer Position bei identischer Raumaufstellung (nur der Arraywagen wurde bei jeder Teilmessung verschoben).

Dessen Einfluss wird aber aufgrund seiner geringen Ausdehnung vernachlässigt), kann später aus diesen Daten die Temperaturänderungen im Raum berechnet werden. Dadurch lassen sich auch kleine Temperaturveränderungen im 1/100-°C-Bereich nachweisen.

Dieser Effekt wird später im Kapitel 6 noch genauer behandelt.

1.14 Ablauf der Arraymessung

Wie bereits festgestellt, sollte die Temperaturmessung unter stationären Raumbedingungen durchgeführt werden, weshalb sich die Schachtel während der Temperaturmessung noch nicht bewegen durfte.

Die Vorschubmutter des Schlittens wurde aus diesem Grund vor Beginn jeder Messung auf $x = -85,8$ mm gestellt, wobei die Schachtel jedoch bei $x = -3,3$ mm aufgestellt wurde. Während die Temperaturmessung also stattfand, legte die Vorschubmutter den Weg von $x = -85,8$ mm auf $x = -3,3$ mm zurück.

Die Position $x = 0$ mm entspricht der ersten, später aus den Aufnahme­daten generierten Impulsantwort.

Aus den Aufnahme­daten konnten insgesamt 1048 Raumimpulsantworten generiert werden.

In Abb. 4.14 ist nun das Wegdiagramm von Messschlitten 1 zu sehen.

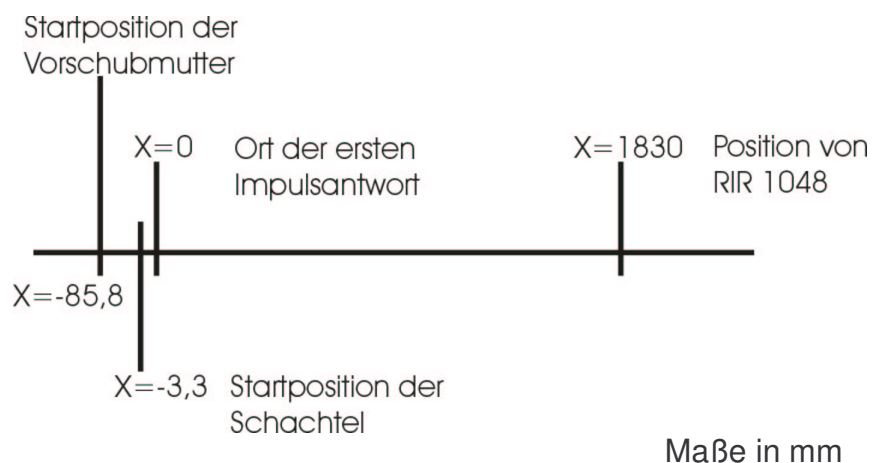


Abb. 4.14: Wegdiagramm Messschlitten 1

1.14.1 Zeitplan einer Teilmessung

Um also das vollständige Mikrofonarray, bestehend aus 96 Sensorpunkten zu erhalten, wurden 16 Teilmessungen mit je 6 Mikrofonen durchgeführt. Dabei dauerte eine Teilmessung ca. 1665 Sekunden.

Der genaue Zeitplan einer Teilmessung ist in Tab. 4.3 zu sehen.

Startzeit [min:sec]	Aktion
------------------------	--------

0:00	Start Messschlitten 1
0:15	Start „Pd“-Ablaufsteuerung
0:23	• Start Temperaturmikrofon-Aufnahme (Stereo File)
0:25	→ Start MLS-Beschallung
Ca. 1:10	Schraubenmutter erreicht Fahrgestell der Schachtel (Schachtel beginnt zu fahren)
1:12	• Ende Aufzeichnung „Temperaturmikrofon“
1:13	➤ Start der Mehrkanal-Arrayaufnahme
27:25	➤ Ende Mehrkanalaufzeichnung
27:26	→ Ende MLS-Beschallung
Ca. 27:42	Ausschalten von Messschlitten 1

Tab. 4.3: Zeitplan während einer Teilarraymessung

1.14.2 Ablauf zwischen den Teilmessungen

Zwischen den Teilmessungen (Messpause) musste beispielsweise die Schachtel wieder auf ihre Startposition zurückgestellt werden (dies erfolgte händisch) bzw. das 6er Array mit Hilfe des Messschlittens 2 auf die nächste Arrayposition plaziert werden.

Der genaue Ablaufplan ist in Abb. 4.15 dargestellt.

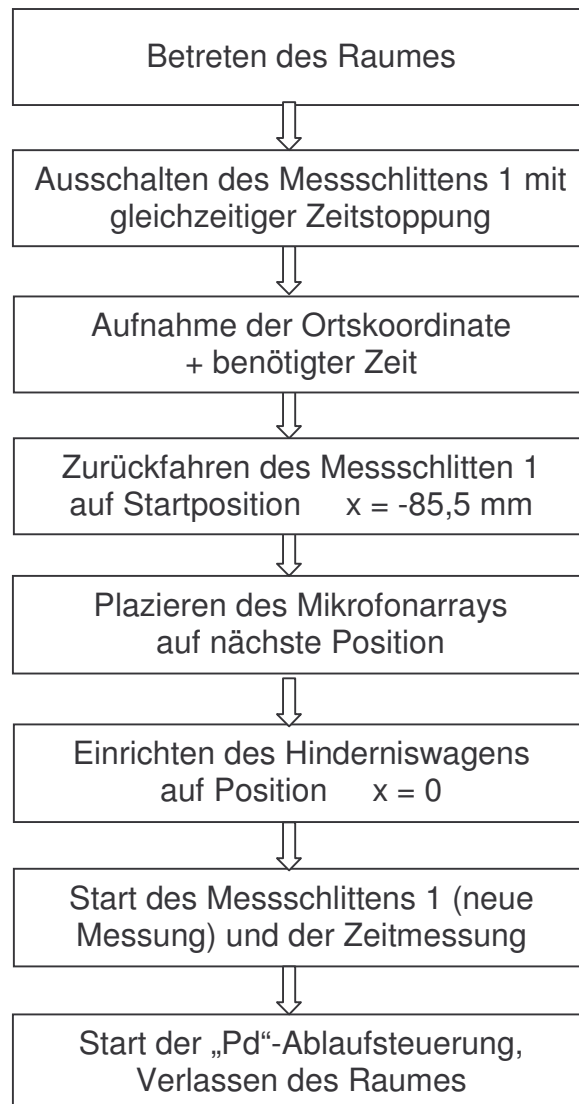


Abb. 4.15: Ablaufplan während den Messpausen

Dieser Ablauf dauerte jeweils ca. 15 Minuten.

Verfahren zur Auswertung der Messdaten

1.15 Darstellung der Raumimpulsantwort

1.1.1 Zeitbereich

Im Zeitbereich liefert eine Impulsantwort direkte bzw. indirekte Aussage über:

- Nachhallzeit
- Größe des Raumes
- Absorptionsgrad des Raumes
- Zeitliches Auftreten der Schallfronten
- Erste Reflexionen und daraus resultierende Kammfiltereffekte

Die Zeitbereichsdarstellung der Impulsantwort eignet sich vor allem im vorderen Teil zur zeitlichen Zuordnung der ersten Reflexionen innerhalb der RIR. Später ist das jedoch durch die zahlreiche Überlagerung von einzelnen Impulsen nicht mehr möglich. Es kann beispielsweise aus starken ersten Reflexionen der Kammfiltereffekt (siehe Abb. 5.1) erklärt werden. Der zeitliche Abfall der Energie der RIR entsprechend $h(t)^2$ gibt Auskunft über die Nachhallzeit.

1.15.1 Frequenzbereich

Durch die Fourier-Transformation (Gl. (2.4)) der Impulsantwort $h(t)$ erhält man die Frequenzbereichsdarstellung, welche auch Übertragungsfunktion genannt wird.

Im Frequenzbereich liefert eine Impulsantwort direkte Aussage über:

- Raumresonanzen bzw. Kammfiltereffekte
- Allgmeindämpfung des Raumes

In Abb. 5.1 sind keine besonders hervortretenden Raumresonanzen zu erkennen. Der Grund liegt in der Geometrie des CUBES (siehe Anhang A1 bzw. A2). Bei der Herleitung der Eigenschwingungen eines Raumes (siehe Kapitel 1.2.1) wurde ein quaderförmiger Raum vorausgesetzt, was für den CUBE jedoch nicht zutrifft. Durch die ungleichförmige Geometrie werden somit starke isolierte Resonanzfrequenzen vermieden. Hingegen ist der Kammfiltereffekt im hochfrequenten Bereich sehr deutlich zu sehen. Die genauere Behandlung erfolgt unter Punkt 5.7.

Über 7 kHz erkennt man bereits die Höhendämpfung des Raumes, welche durch die Auskleidung des Raumes mittels Absorbermaterial hervorgerufen wird. Über 16 kHz

ist ein sehr starker Höhenabfall in Abb. 5.1 sichtbar, welcher ebenfalls auf die Schallabsorption der Wände zurückgeführt wird, wobei in diesem Frequenzbereich bereits die Luftabsorption des Schalls relevant wird.

Abb. 5.1 zeigt die Impulsantwort im Zeit- und Frequenzbereich:

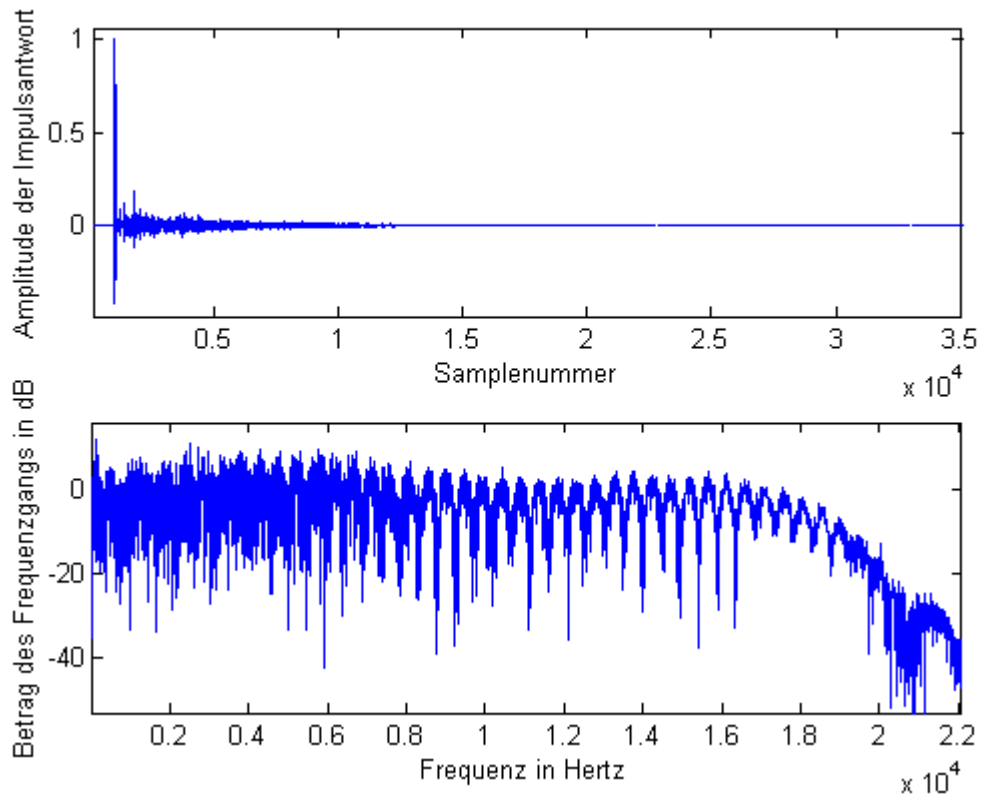


Abb. 5.1: Zeit- und Frequenzbereichsdarstellung einer Raumimpulsantwort

1.16 Definition der Raumimpulsantwort-Matrix M

Falls nicht nur eine, sondern A Raumimpulsantworten der Länge N aufgenommen werden, wird man sie zweckmässigerweise in einer Matrix zusammenfassen. Diese Matrix M wird wie folgt definiert:

$$M = \begin{bmatrix} h_1(n) \\ h_2(n) \\ \vdots \\ h_{(A-1)}(n) \\ h_A(n) \end{bmatrix} \quad (5.87)$$

Hierbei entspricht die zeitliche Abfolge einer Impulsantwort $h_m(n)$ einem Zeilenvektor. Die zeitliche Abfolge der verschiedenen Impulsantworten ist in vertikaler Richtung der Matrix gegeben. In vollständiger Matrixschreibweise wird die Impulsantwort-Matrix M wie folgt angeschrieben:

$$M(m,n) = \begin{bmatrix} h_{1,1} & \dots & h_{1,N} \\ h_{2,1} & \ddots & \vdots \\ \vdots & h_{m,n} & \vdots \\ h_{(A-1),1} & \ddots & \vdots \\ h_{A,1} & \dots & h_{A,N} \end{bmatrix} \quad \begin{matrix} 1 \leq m \leq A \\ 1 \leq n \leq N \end{matrix} \quad (5.88)$$

1.17 Einige Betrachtungen zum SDR

Es soll nun anhand einiger Beispiele von Impulsantwortmanipulationen gezeigt werden, welchen Einfluss diese auf die SDR-Werte mit sich bringen. Es wird von einer Impulsantwort ausgegangen, die 35000 Abtastwerte umfasst.

1.1.1 SDR angewandt auf verschiedene Bereiche der Impulsantwort

Es werden jeweils die identischen Zeitbereiche zweier zeitlich benachbarter Impulsantworten beginnend beim Bereichsindex "y" verglichen. Diese Bereiche der Länge von 1000 Abtastwerte werden vom Anfang bis zum Ende der Impulsantworten durchgestept.

Das SDR wird dabei wie folgt berechnet:

$$SDR = 10 \log_{10} \frac{\sum_{n=y}^{y+999} \{h_1[n]\}^2}{\sum_{n=y}^{y+999} \{h_1[n] - h_2[n]\}^2} \quad [dB]$$

Das SDR wird nun in Abhängigkeit der Zeit angegeben:

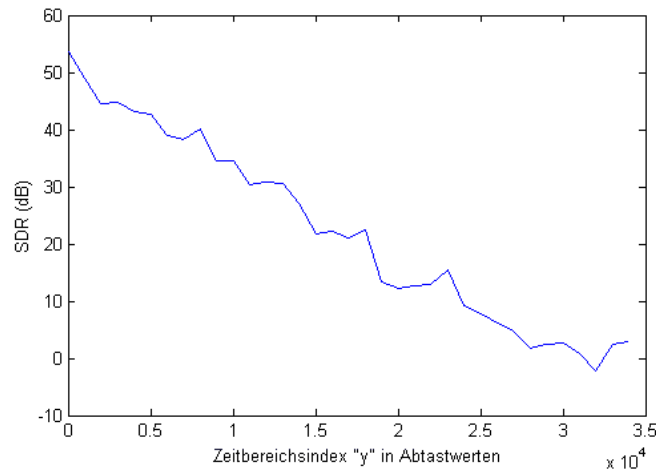


Abb. 5.2: Reproduzierbarkeit einer Impulsantwort in Abhängigkeit der Zeit

Im vorderen Teil der Raumimpulsantwort, kann diese auf über 50 dB reproduzierbar gemessen werden. Im hinteren Teil sinkt das SNR, weshalb auch dieser Teil zweier benachbarter Raumimpulsantworten nicht mehr übereinstimmt. Es ist zu sehen, dass die Reproduzierbarkeit der RIR bis zur Nachhallzeit, die bei ca. 33000 Abtastwerte liegt, praktisch auf null sinkt.

1.17.1 SDR bei Abschneiden beider Raumimpulsantworten

Ausgehend von Gl. (3.6), wird das SDR nicht bis $N - 1$, sondern nur bis Parameter „x“ aufsummiert, welcher von 1000 bis $N - 1$ variiert wird.

Da die Energie exponentiell abnimmt, sinkt das SDR im gesamten Bereich nur um ca. 10 dB, obwohl die Einzelbereiche im hinteren Teil der Impulsantwort bis auf 0 dB SDR abfallen (Abb. 5.2).

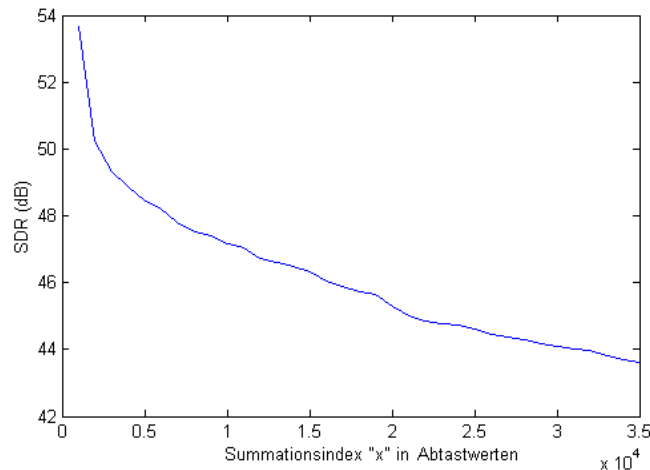


Abb. 5.3: SDR der abgeschnittenen RIR

Die in Abb. 5.2 dargestellten kleinen SDR-Werte im hinteren Bereich der RIR wirken sich, wie in Abb. 5.3 zu sehen ist, nicht besonders negativ aus, da im hinteren Teil einer Impulsantwort die Signalenergie bereits stark abgenommen hat.

1.17.2 SDR bei Abschneiden einer Raumimpulsantwort

Hier soll gezeigt werden, welche Auswirkung es auf das SDR hat, wenn die Impulsantwort $h_2(n)$ abgeschnitten und mit Nullen aufgefüllt wird. Die Referenzimpulsantwort $h_1(n)$ bleibt in ihrer ursprünglichen Form bestehen.

$$h_2(n) = \begin{cases} h_1(n), & \text{wenn } n < z \\ 0, & \text{wenn } n \geq z \end{cases}$$

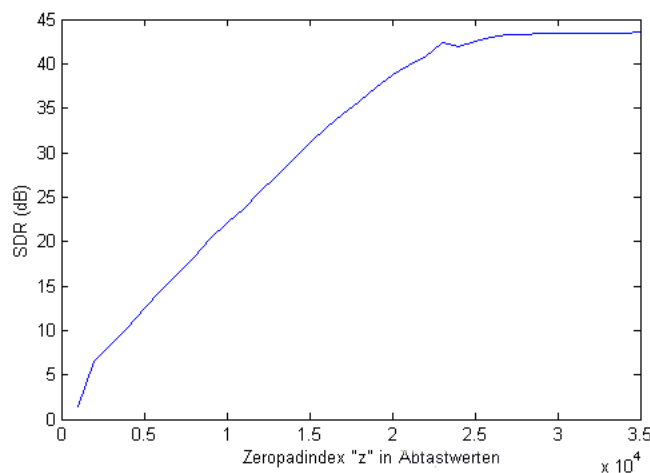


Abb. 5.4: Erreichbares SDR, falls die RIR abgeschnitten wird

Wird die RIR bei ca. 22000 Sample abgeschnitten, ergibt sich keine signifikante Verschlechterung des SDR-Wertes.

Da jedoch die tiefen Frequenzen eine höhere Nachhallzeit aufweisen, kann nicht die allgemeine Aussage getätigt werden, dass RIRs für Echokompensationszwecke bereits bei 2/3 der Nachhallzeit die gesamte relevante Information beinhalten.

1.17.3 Mittelung von Raumimpulsantworten

Hier wird untersucht, in welchem Ausmaß die Mittelung über k Raumimpulsantworten zur Reduktion der Varianz bei der Bestimmung einer RIR, beiträgt.

$$M_{\text{Mittel}}(n) = \frac{1}{k} \sum_{m=60-k+1}^{60} M(m, n) \quad (5.89)$$

Diese gemittelte RIR (jetzt als Referenzimpulsantwort betrachtet) wird nun mit den Raumimpulsantworten im Indexbereich von $m=40$ bis $m=170$ bezüglich SDR verglichen.

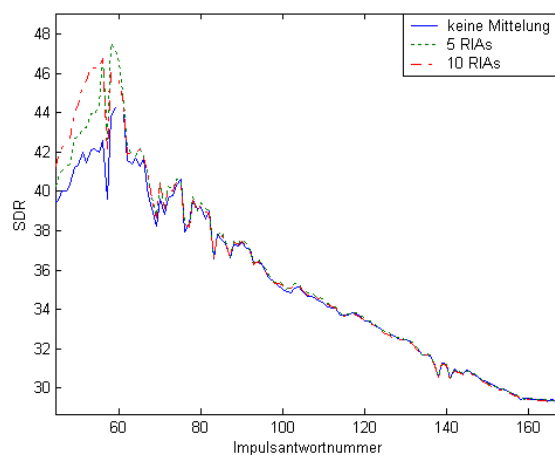


Abb. 5.5: SDR bei Mittelung über $k=1$, $k=5$ bzw. $k=10$ Raumimpulsantworten

In horizontaler Richtung von Abb. 5.5 wird die Impulsantwort-Nummer dargestellt, welche in der Matrixdarstellung $M(m, n)$ dem Parameter m entspricht. Es ist zu erkennen, dass zwar die Mittelung innerhalb jenes Bereichs, wo Mittelung passiert ($m = 51$ bis 60), das SDR um einige dB steigt, jedoch außerhalb des Mittelungsbereichs ergibt sich keine Erhöhung des SDR für $m > 60$.

Die Ursache des Abfalls des SDR wird später zum Teil auf die Temperaturveränderung zurückgeführt und sei zunächst als Drift bezeichnet. Es wird für alle weiteren Betrachtungen keine Impulsantwort-Mittelung vorgenommen, da die Drifterscheinung der RIR-Messung bei weitem die statistische Messwertschwankung übersteigt.

1.17.4 SDR in Abhängigkeit des Abstandes

Dabei soll festgestellt werden, wie sich das SDR im Bezug auf den Abstand des Mikrofons zum Lautsprecher ändert. Dieser bewegt sich von innerhalb bis deutlich außerhalb des Hallradius.

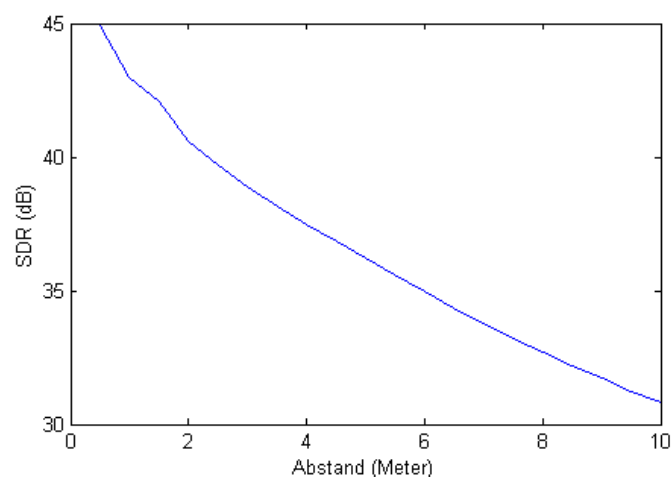


Abb. 5.6: SDR in Abhängigkeit des Abstandes vom Lautsprecher zum Mikrofon

Das Ansteigen der SDR-Werte bei kleiner werdendem Abstand von Mikrofon zu Lautsprecher, erklärt sich durch das Ansteigen des Verhältnisses von Direktschall zu Diffusschall. Der Direktschall hat erstens das größte SNR im Vergleich zu allen späteren Reflexionen und bietet klarerweise auch eine höhere Reproduzierbarkeit als eine schon mehrfach reflektierte Schallfront.

1.17.5 SDR in Abhängigkeit des Beschallungspegels

Die Messung wurde bei einem Abstand zwischen Lautsprecher und Mikrofon von 5 m durchgeführt. Spalte 2 der Tab. 5.1 zeigt dabei das SDR zweier aufeinanderfolgender RIRs.

Pegel des Lautsprecher-Signals [dB]	SDR [dB]
0	37,7
-3	36,6
-6	35,9
-12	33,6
-20	27,5

Tab. 5.4: SDR in Abhängigkeit des Beschallungspegels

Anhand Tab. 5.1 ist zu erkennen, dass sich durch Reduktion des Lautsprecherpegels um 12 dB, das SDR nur um 4,1 dB verschlechtert. Müsste also bei Anwesenheit von Personen im Raum eine RIR gemessen werden, wäre eine Reduktion des MLS-Signalpegels natürlich viel angenehmer für die Anwesenden, und hätte ein nicht viel schlechteres SDR-Verhältnis zur Folge. Da sich jedoch niemand regungslos während der Messung verhält, dürfen dabei nicht allzu präzise Messergebnisse erwartet werden.

1.18 Impulsantwort im Frequenzbereich

Hier werden zwei zeitlich benachbarte RIRs jeweils in den Frequenzbereich transformiert, um diese mittels SDRF zu vergleichen.

Frequenzbereich [Hz]	SDRF [dB]
0 – 100	23,7
101 – 200	45,5
201 – 500	46,4
501 – 1000	50,2
1001 – 2000	55,0
2001 – 3400	54,9
3401 – 8000	50,8
8001 – 16000	46,9
16001 – 22050	46,6

Tab. 5.5: SDRF in verschiedenen Frequenzbereichen

Tab. 5.2 zeigt, dass eine RIR im gesamten Frequenzbereich reproduzierbar gemessen werden kann. Im Frequenzbereich unter 100 Hz weisen 2 benachbarte RIRs die größten Unterschiede auf.

1.19 Impulsantwort im Zeit- und Frequenzbereich

Hier wird die RIR sowohl im Zeit- als auch im Frequenzbereich betrachtet. Dazu wird diese in 6 Zeitabschnitte $B_i(i)$ zu je 5000 Abtastwerte geteilt, in den Frequenzbereich transformiert und dann jeweils in 9 Frequenzbereichen mittels SDR betrachtet.

$$B_i(i) = [5000(i-1)+1 \quad 5000i] \quad 1 \leq i \leq 6$$

Frequenzbereich [Hz]	$B_t(1)$	$B_t(2)$	$B_t(3)$	$B_t(4)$	$B_t(5)$	$B_t(6)$
	SDRF [dB]	SDRF [dB]	SDRF [dB]	SDRF [dB]	SDRF [dB]	SDRF [dB]
0 – 100	30,4	22,2	9,6	5,8	-0,5	-1,4
101 – 200	50,7	41,2	26,7	20,9	12,9	11,2
201 – 500	52,4	39,8	29,1	21,4	10,6	4,5
501 – 1000	56,6	47,2	36,2	26,2	15,9	10,3
1001 – 2000	61,0	51,3	43,1	35,2	25,7	18,2
2001 – 3400	58,0	50,4	45,5	38,6	30,0	22,5
3401 – 8000	52,8	44,7	40,0	32,3	23,2	14,2
8001 – 16000	49,1	36,3	26,6	13,6	4,0	-0,7
16001 – 22050	48,8	30,0	13,1	2,5	-1,3	-0,5

Tab. 5.6: SDRF in verschiedenen Zeit- und Frequenzbereichen

Tab. 5.3 zeigt also die frequenzabhängige Bewertung von Abb. 5.2. Hier ist zu sehen, dass die Impulsantwort im Bereich der Nachhallzeit ($\approx B_t(6)$) im mittleren Frequenzbereich noch reproduzierbar gemessen werden kann. Das SDR leidet jedoch an den geringen SDRF-Werten im hohen und tiefen Frequenzbereich.

1.20 Kenngrößen von Raumimpulsantworten

1.1.1 Nachhallzeit

Die Nachhallzeit erhält man anhand des 60 dB Abfalls der Einhüllenden Impulsantwort. Der exponentielle Energieabfall ist dabei im logarithmischen Maßstab als Gerade zu sehen. In Abb. 5.7 ist dies bei ca. 33000 Abtastwerten der Fall.

Somit beträgt die hier ermittelte Nachhallzeit ungefähr:

$$T_N = N \frac{1}{T_s} = \frac{33000}{44100} = 0,75 \text{ sec} \quad (5.90)$$

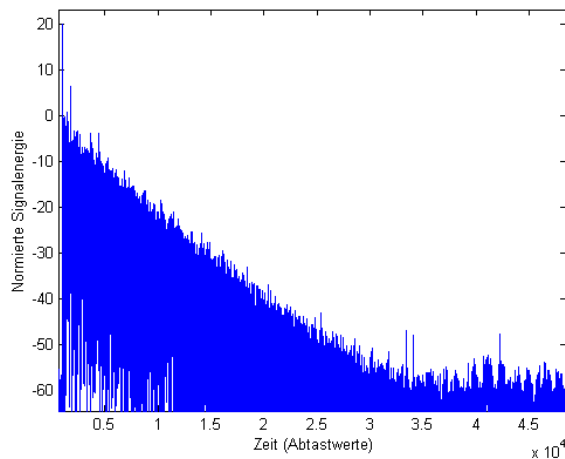


Abb. 5.7: Nachhallzeitbestimmung anhand des Energieabfalls in der Einhüllenden der Impulsantwort

Es lässt sich auch die Dynamik der RIR gut erkennen. Im Bereich des Direktschalls und der ersten Reflexionen gilt die exponentielle Abnahme der Energie noch nicht, da noch keine statistische Energieverteilung im Raum herrscht.

Der Direktimpuls liegt, wie in Abb. 5.7 zu sehen, ca. 80 dB über der Signalenergie, des Rauschteppichs der Impulsantwort Nachhallzeit. Somit muss der Dynamikumfang einer RIR mit 80 dB angegeben werden, was also nahezu dem gesamten CD Dynamikumfang entspricht, sofern noch ein Headroom eingeplant wird.

Der Rauschteppich ist in Abb. 5.7 bei ca. -60 dB zu sehen. Das maximale SNR beim Spitzenwert des Direktschalls beträgt also ungefähr 80 dB.

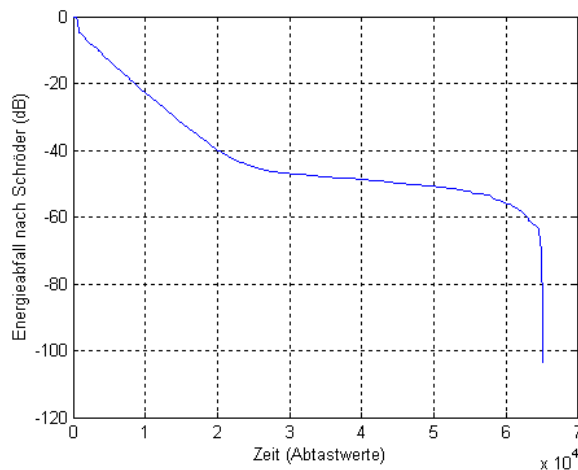


Abb. 5.8: Nachhallzeitbestimmung anhand der Schröder-Rückwärtsintegration

Abb. 5.8 zeigt die Nachhallzeitbestimmung nach „Schröder“ (Gl. 1.22). Aufgrund des Rauschteppichs der Impulsantwort, ist obiges Bild jedoch nur bis zum -40 dB Abfall aussagekräftig. Durch Verlängern der Geraden, erhält man die Nachhallzeit.

1.20.1 Nulldurchgangsrate

Mit dieser Kenngröße wird die Anzahl der Nulldurchgänge pro Sekunde angegeben.

$$N_0 = \frac{N}{T} \quad \left[\frac{1}{s} \right] \quad (5.91)$$

N Zahl der Nulldurchgänge

N_0 Nulldurchgangsrate

Die untersuchte Impulsantwort weist eine Rate von 8962 Nulldurchgänge/Sekunde auf. Dies entspricht einer charakteristischen Signalfrequenz von 4481 Hz entsprechen.

1.20.2 Untersuchung zum Hallradius

Für die Hallradiusberechnung wird meist die Formel von Sabine (siehe Gl. (1.21)) verwendet. Diese ist jedoch stark vereinfachend, und gilt nur für kleine Absorptionsgrade α . Außerdem ist der Hallradius frequenzabhängig.

Für folgende Berechnungen wurde die Impulsantwort in zwei Teile zerlegt, in den Direktschall und den Nachhallteil.

In der weiter unten stehenden Abb. 5.11 ist der Direktschallimpuls im Zeitbereich gut zu erkennen. Er wird auf die Zeitdauer von 980 – 1080 Abtastwerten begrenzt, wobei dies für die verschiedenen Abstände der Abb. 5.9a jeweils ein anderer Bereich ist.

$$h_{dir}(n) = h(n)W_{dir}(n)$$

$$h_{diff}(n) = h(n)W_{diff}(n)$$

$$W_{dir}(n) = 1, 980 \leq n \leq 1080, \text{sonst } 0$$

$$W_{diff}(n) = 1, 1080 < n \leq N, \text{sonst } 0$$

Nun wird das Verhältnis der Energien gebildet und wegen der großen Dynamik logarithmiert:

$$\frac{E_{dir}}{E_{diff}} [dB] = 10 \log_{10} \frac{\sum_{n=0}^{N-1} h_{dir}^2(n)}{\sum_{n=0}^{N-1} h_{diff}^2(n)} \quad (5.92)$$

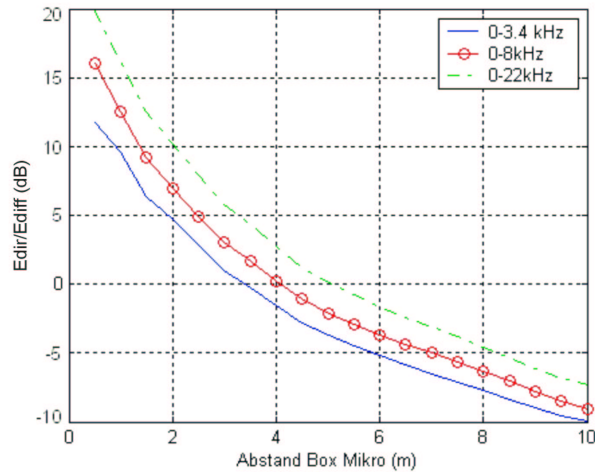


Abb. 5.9a: Verhältnis der Direkt- zur Diffusschallenergie in Abhängigkeit von der Distanz Lautsprecher-Mikrofon

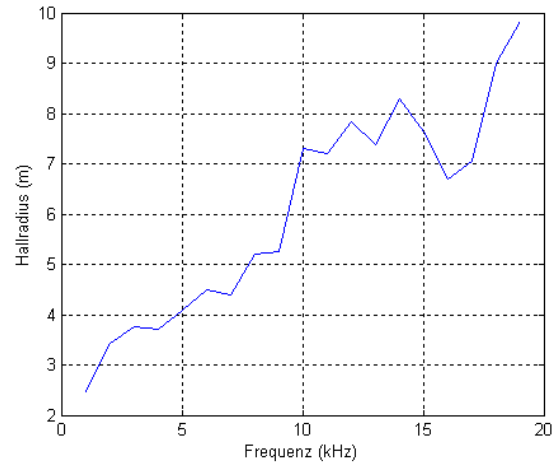


Abb. 5.9b: Hallradius in Abhängigkeit der Frequenz

Bei 0 dB kann in Abb. 5.9a also der Hallradius abgelesen werden. Durch jeweilige Filterung der zeitlich gefensterten Impulsantwortteile mit einem Tschebycheffilter 1. Art, 4. Ordnung, werden dabei 3 Frequenzbereiche analysiert. Bei der Filterung ist wichtig, dass diese mit den gefensterten RIRs passiert. Durch die Filterung entsteht bei tiefen Grenzfrequenzen des Tiefpassfilters eine große Nachschwingzeit, die zu zeitlichem Ineinanderlaufen von Direktschall und ersten Reflexionen führen würde.

Der Grund für die Abhängigkeit des Hallradius von der Frequenz (Abb. 5.9b) liegt im Abstrahlverhalten des Lautsprechers. Strahlt eine Schallquelle als Kugelstrahler, so ist der Schalldruck p proportional dem reziproken Abstand. Bei hohen Frequenzen ist jedoch die Annahme des Kugelstrahlers nicht mehr korrekt, und die Richtcharakteristik der Box wird anisotrop. Liegt nun das Mikrofon in der Hauptkeule, wird der Hallradius scheinbar größer, da für die Richtung der Hauptkeule die $\frac{1}{r}$ Schalldruckabnahme nicht mehr gilt.

1.21 Entstehen von Kammfilter-Effekten

Kammfilter werden aufgrund der grafischen Ähnlichkeit ihres Betragsfrequenzganges mit einem Kamm so genannt.

Der Effekt sei anhand eines Beispiels erklärt:

Betrachtet werde in Abb. 5.10a ein Impuls mit Amplitude a , dessen konstantes Betragsspektrum in Abb. 5.10b zu sehen ist. Wird nun ein zweiter Impuls mit Amplitude b und Verzögerungszeit T_i addiert und das Summensignal in den Frequenzbereich transformiert, so ergibt sich bereits der Kammfiltereffekt, der in Abb. 5.12b dargestellt wird.

Dabei ergibt sich für die Kerbfrequenzen:

$$f_{k,n} = \frac{1}{T_i} n \quad n = 1, 2, 3, \dots \quad (5.93)$$

T_i Zeitintervall zwischen zwei Impulsen

Für die Kerbfrequenzen gilt dann folgende Verstärkung G :

$$G(f_{k,n}) = 20 \log \frac{a-b}{a} \text{ [dB]} \quad \text{für } a > b > 0 \quad (5.94)$$
$$\text{für } a = b \quad G(f_{k,n}) \rightarrow -\infty$$

Die Verstärkung G bei der Kerbfrequenz erhält man aufgrund der destruktiven Interferenz zweier Signale der Frequenz $f_{k,n}$.

Folgendes gilt für die Maxima, die genau zwischen den Kerbfrequenzen liegen:

$$G(f_{k,n}) = 20 \log \frac{a+b}{a} \text{ [dB]} \quad \text{für } a \geq b \geq 0 \quad (5.95)$$

Die Maximalverstärkung erhält aufgrund konstruktiver Interferenz zweier Signale und beträgt maximal 6 dB ($a = b$)

Somit sind die Kerbfrequenzen umso deutlicher zu sehen, je geringer die Pegelunterschiede von a und b sind. Kommen weitere Impulse hinzu, so verliert sich die Regelmäßigkeit des Kammfilters.

In den folgenden Abbildungen sollen nun einige grundlegende Aspekte einer Impulsantwort gezeigt werden. Dabei befindet sich jeweils in den linken Abbildungen die Raumimpulsantwort im Zeitbereich und auf der rechten Seite die dazugehörige Frequenzbereichsdarstellung.

Abb. 5.10a zeigt den bereits erwähnten Impuls und das dazugehörige Betragsfrequenzspektrum.

In Abb. 5.11a ist der Direktschallanteil der RIR zu sehen. Hier wird deutlich, dass wir ja die Impulsantwort des bandbegrenzten Systems LEM bestimmt haben. Die Bandbegrenzung wird einerseits vom Lautsprecher im tiefen, andererseits vom Anti-Aliasing-Filter des A/D-Wandlers im hohen Frequenzbereich verursacht. Durch diese Bandpassfilterung kommt es somit zur Verbreiterung des idealen Impulses.

Die Abb. 5.12a und Abb. 5.12b verdeutlichen das Zustandekommen des Kammfiltereffektes.

In Abb. 5.12a ist der Direktschallimpuls und die erste Reflexion zu sehen. Diese ist stark ausgeprägt, da sie durch den glatten Fußboden des Raumes, der einen geringen Absorptionsgrad aufweist, reflektiert wurde. Dabei stellt jedoch jede Reflexion einer Schallfront an einer Oberfläche keine ideale Verzögerung, sondern eine weitere Filterung dar. Da also die Bodenreflexion eine gefilterte Version des Direktschallanteils darstellt, ist deshalb auch der Kammfiltereffekt nicht gleichmäßig über den gesamten Frequenzbereich zu sehen (unterschiedliche Unterdrückung bei den Kerbfrequenzen, welche nicht auf Unterabtastungseffekte zurückzuführen sind).

In Abb. 5.13a sind weitere Raumreflexionen verdeutlicht. Diese sind aber aufgrund der kleineren Amplitude nicht so stark im Frequenzbereich sichtbar, da die Kerbfrequenzunterdrückung geringer ist.

In Abb. 5.14a wird der Nachhallteil, wo Einzelreflexionen nicht mehr voneinander getrennt werden können, dargestellt. Im Frequenzbereich ist nun deutlich die Höhenbedämpfung des Raumes zu erkennen, die hier ab ca. 6 kHz sichtbar wird.

Im letzten Bilderpaar wird schließlich die gesamte RIR im jeweiligen Bereich dargestellt. Im tiefen Frequenzbereich ist der Kammfiltereffekt nicht mehr zu erkennen, da dieser vom Nachhall (Abb. 5.15b) sozusagen verschmiert wird. Im höheren Frequenzbereich, wo die Nachhallenergie bei 10 kHz schon ca. um 10 dB abgenommen hat (Abb. 5.14b), bleibt die Kammfiltererscheinung deswegen erhalten.

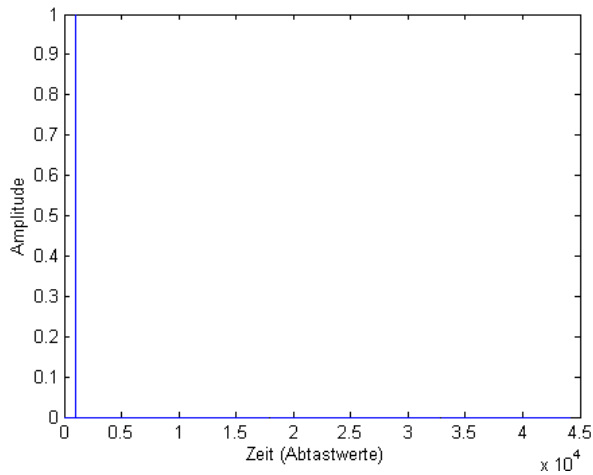


Abb. 5.10a: Idealer Impuls

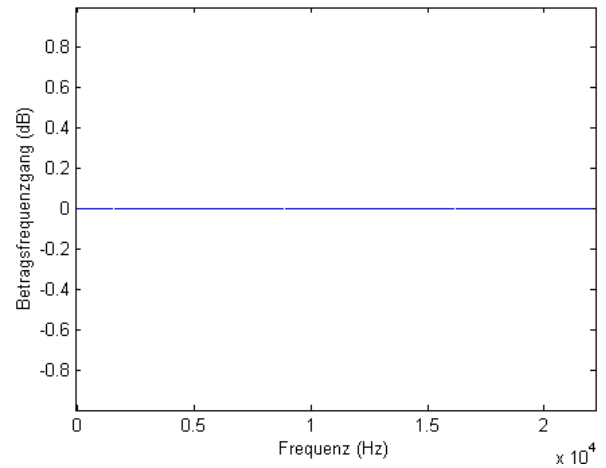


Abb. 5.10b: Konstantes Betragsspektrum

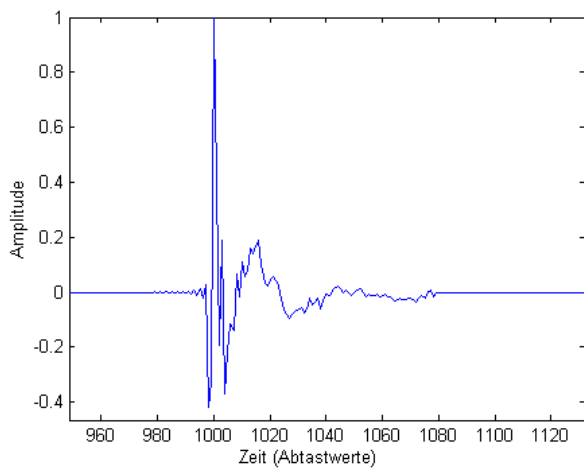


Abb. 5.11a: Direktschallimpuls RIR im Zeitbereich

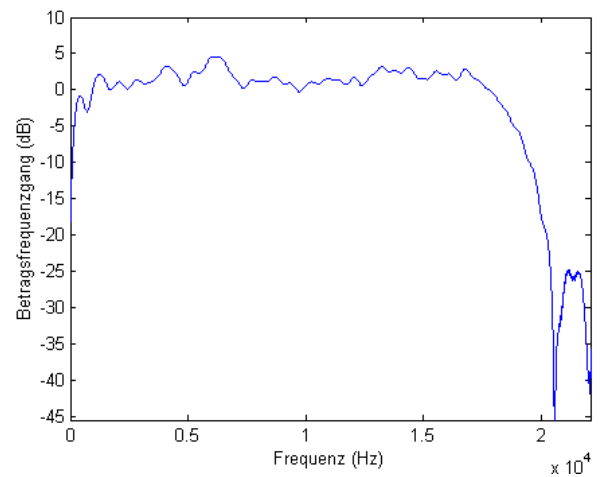


Abb. 5.11b: Direktimpuls Betragsspektrum

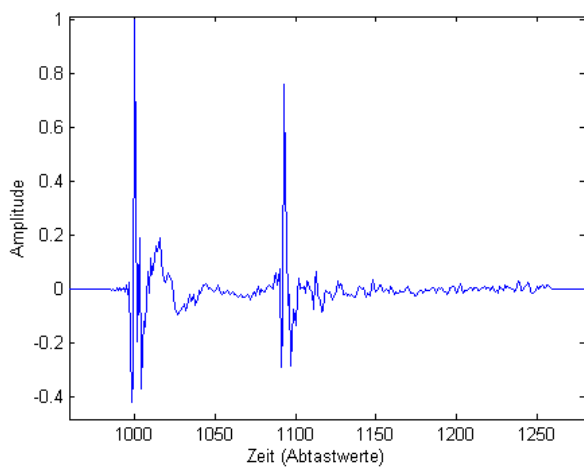


Abb. 5.12a: Direktschallimpuls und Bodenreflexion

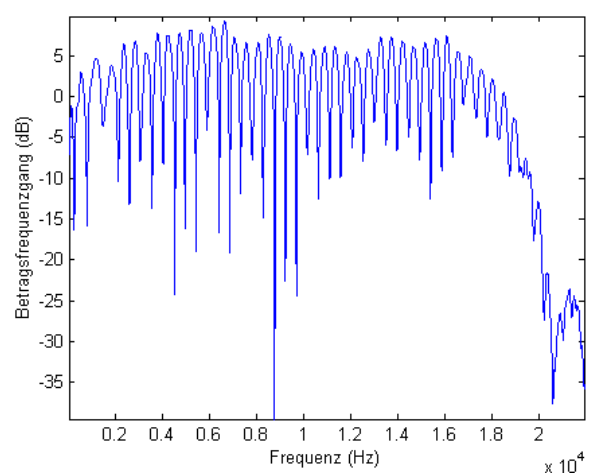


Abb. 5.12b: Kammfiltereffekt

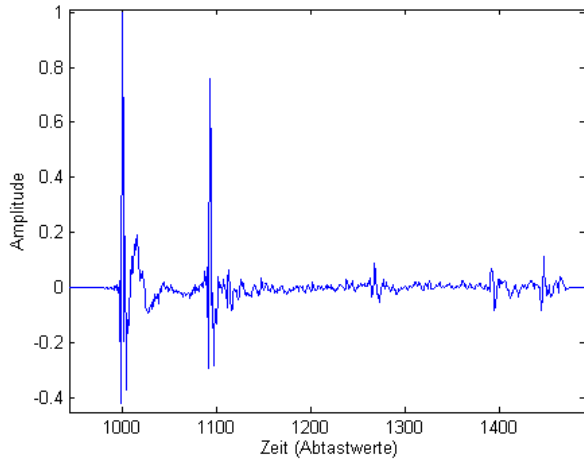


Abb. 5.13a: Weitere Reflexionen

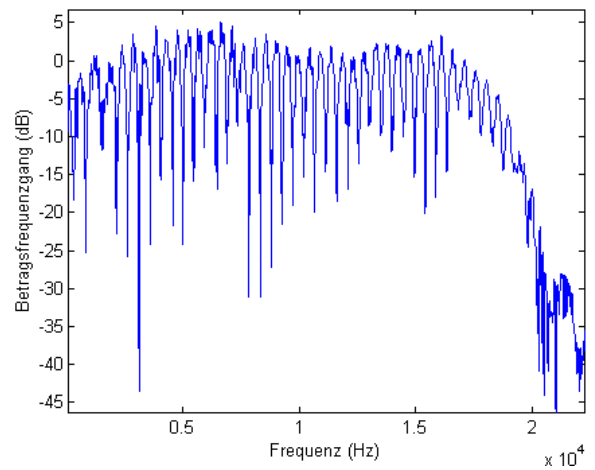


Abb. 5.13b: Kammfiltereffekt wird unregelmäßiger

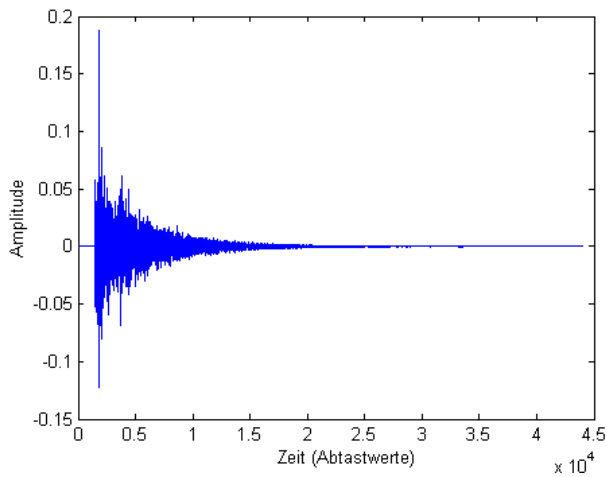


Abb. 5.14a: Nachhallanteil

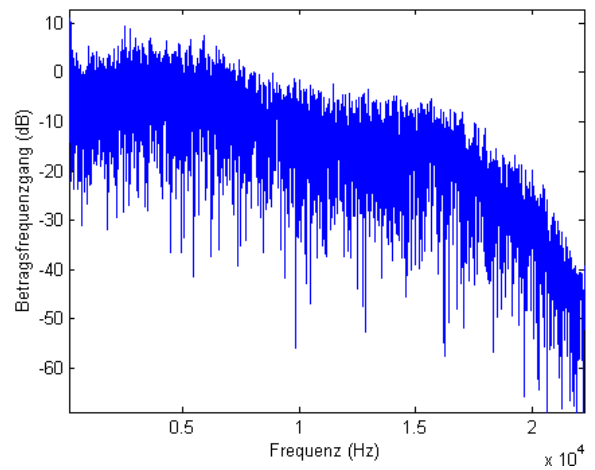


Abb. 5.14b: Betragsspektrum des Nachhalls

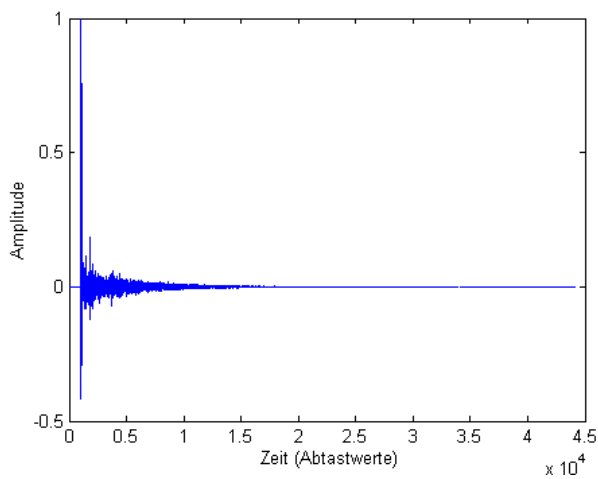


Abb. 5.15a: Gesamtimpulsantwort

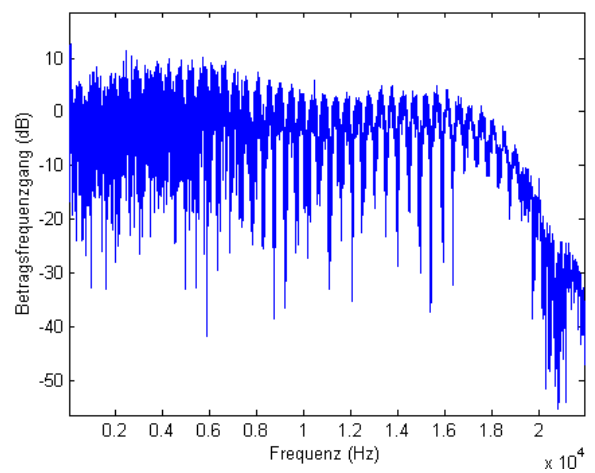


Abb. 5.15b: Gesamtfrequenzgang

1.22 Temperatureinfluss

Besonderen Einfluss auf die Impulsantwort hat die Temperaturveränderung, da die Schallgeschwindigkeit, wie in Kapitel 1 bereits angegeben, von der Temperatur abhängt.

Nimmt man eine gleichmäßige Temperaturerhöhung des Raumes an, steigt also die Ausbreitungsgeschwindigkeit des Schalls überall an. Dies hat für die Impulsantwort die Auswirkung, dass sie in der Zeitachse gestaucht wird, da die Wellenpakete aufgrund erhöhter Ausbreitungsgeschwindigkeit den selben Weg in kürzerer Zeit zurücklegen können.

Nehmen wir beispielsweise bei 20°C eine Temperaturveränderung von 1°C an, so erhöht sich die Schallgeschwindigkeit um 0,6 m/s. Bezogen auf die Ausgangsschallgeschwindigkeit von 343,3 m/s, ergibt sich somit eine prozentuelle Veränderung von 0,174 %. Diese hat, bezogen auf eine Impulsantwortlänge von beispielsweise 20000 Abtastwerten, schon eine enorme Auswirkung.

Diese Stauchung der Impulsantwort kann durch eine lineare, inverse Zeitskalierung der Zeitachse auf die ursprüngliche Länge der RIR zurückgerechnet werden. Es wird gezeigt, dass dadurch Temperaturveränderungen im Raum teilweise kompensiert werden können. Gleichung (1.1) wird nochmals angeschrieben.

$$c(\vartheta) \approx 331,3 + 0,6 \cdot \vartheta \quad \vartheta \dots \text{Lufttemperatur in } ^\circ\text{C}$$

$$h(v + dv, t) = h(v, t - \delta t)$$

$$v \cdot t = (v + dv)(t - \delta t)$$

Wird $dv \cdot \delta t$ vernachlässigt (ist jedoch nur für kleine Skalierungsfaktoren zulässig) und Gl. (1.1) eingesetzt, so erhält man:

$$\delta = \frac{dv}{v} = \frac{0,6 d\vartheta}{v} \quad (5.96)$$

δ differentieller Skalierungsfaktor

Ein positiver differentieller Skalierungsfaktor bedeutet eine Temperaturerhöhung. Die neue Impulsantwortlänge N_{neu} wird somit bei Temperaturzunahme verkürzt.

Die neue Folgenlänge ergibt sich zu:

$$N_{neu} = N(1 - \delta) \quad (5.97)$$

$(1 - \delta)$ Skalierungsfaktor

Die absolute Verkürzung S der Impulsantwort in Abtastwerten erhält man über:

$$S = N(-\delta) \tag{5.98}$$

Die Kompensation der temperaturabhängigen Zeitskalierung erfolgt mit einem Algorithmus gemäß [Lit. 11], in Übereinstimmung mit der physikalischen Wellenausbreitung berücksichtigt dieser Algorithmus nur die differenzielle Zeitverschiebung der Impulsantwortwerte, nimmt aber keine explizite Interpolation vor (implizit aber schon).

$$h'[m] = \frac{1}{N} \sum_{k=0}^{N-1} H'[k, m] e^{j\frac{2\pi k}{N}m} \tag{5.99}$$

$$H'[k, m] = \begin{cases} H[k] e^{-j\frac{2\pi k}{N}m\delta} & 0 \leq k \leq N/2 - 1 \\ H[k] e^{j\frac{2\pi k}{N}m\delta} & N/2 \leq k \leq N - 1 \end{cases} \tag{5.100}$$

Dieser Algorithmus benötigt jedoch einen sehr großen Rechenaufwand, der mit $N \log N + 4N^2$ abgeschätzt werden kann.

1.1.1 Simulierte Temperaturerhöhung

In folgender Tabelle wurde eine Temperaturerhöhung simuliert, indem eine Impulsantwort der Länge $L=20000$ auf $L-S$ Abtastwerte skaliert wurde.

Der differenzielle Zeitskalierungsfaktor wurde von 0,00005 bis 0,0005 variiert, was bei einer Länge von 20000 Abtastwerten auf eine absolute Längenänderung von 1 bis 10 Abtastwerten führt.

Die 3. Spalte zeigt die Auswirkung auf das SDR. Die Zeitskalierung wird mit dem Matlab-Befehl "Resample" durchgeführt. Um die Qualität dieser Operation zu überprüfen, wurde in der 4. Spalte der Tab. 5.4 dieser Vorgang invertiert und noch einmal das SDR relativ zur Ausgangsfolge überprüft.

Differenzieller Zeitskalierungsfaktor	Längenänderung S absolut	Äquivalente Temperaturdifferenz	SDR-Wert	SDR nach inverser Zeitskalierung
δ	Abtastwerte	°C	dB	dB
$5 \cdot 10^{-5}$	- 1	0,029	19,30	58,12
10^{-4}	- 2	0,057	13,39	50,66
$1,5 \cdot 10^{-4}$	- 3	0,086	10,03	48,17
$2 \cdot 10^{-4}$	- 4	0,114	7,75	46,13
$2,5 \cdot 10^{-4}$	- 5	0,143	6,06	44,01
$3 \cdot 10^{-4}$	- 6	0,172	4,75	42,22
$3,5 \cdot 10^{-4}$	- 7	0,200	3,69	40,94
$4 \cdot 10^{-4}$	- 8	0,229	2,81	40,18
$4,5 \cdot 10^{-4}$	- 9	0,257	2,06	39,80
$5 \cdot 10^{-4}$	- 10	0,286	1,42	39,75

Tab. 5.7: Zeitskalierung über Matlab-Befehl "Resample"

In Tab. 5.4 ist zu sehen, dass bereits eine Zeitskalierung von einem Abtastwert bezogen auf 20000 Abtastwerte nur mehr ein SDR von knapp 20 dB liefert. Dies entspricht einer Temperaturdifferenz von nur 0,03°C. Der Resampling-Algorithmus des Matlab-Befehls arbeitet hier mit genügender Genauigkeit, da immerhin SDR-Werte von 40 dB bei 2-maliger Anwendung erreicht werden. Hier ist noch zu erwähnen, dass leider nur von ganzzahligen auf ganzzahlige Folgenlängen skaliert werden kann. Möchte man eine Folge auf nicht ganzzahlige Folgenlängen skalieren, so wird der Algorithmus aus (5.13), (5.14) verwendet.

Diffenzieller Zeitskalierungsfaktor	Längenänderung S absolut	Äquivalente Temperaturdifferenz	SDR-Wert	SDR nach inverser Zeitskalierung
δ	Abtastwerte	°C	dB	dB
$5 \cdot 10^{-6}$	- 0,1	0,0029	40,7439	92,65
10^{-5}	- 0,2	0,0057	34,7238	92,19
$1,5 \cdot 10^{-5}$	- 0,3	0,0086	31,2028	90,57
$2 \cdot 10^{-5}$	- 0,4	0,0114	28,7052	88,00
$2,5 \cdot 10^{-5}$	- 0,5	0,0143	26,7684	85,27
$3 \cdot 10^{-5}$	- 0,6	0,0172	25,1866	82,79
$3,5 \cdot 10^{-5}$	- 0,7	0,0200	23,8497	80,66
$4 \cdot 10^{-5}$	- 0,8	0,0229	22,6923	78,85
$4,5 \cdot 10^{-5}$	- 0,9	0,0257	21,6720	77,31
$5 \cdot 10^{-5}$	- 1	0,0286	20,7599	76,02

Tab. 5.8: Zeitskalierung über Algorithmus (5.13), (5.14)

Jetzt wurden die Raumimpulsantworten von 0,1 bis 1 Abtastwert gestaucht, bezogen auf 20000 Werte.

1.22.1 Reale Temperaturerhöhung

Hier wird nun anhand von realen Messdaten die Kompensation des Temperatureffekts untersucht. Es werden die Daten einer stationären Langzeitmessung verwendet, wobei die Messzeitpunkte der 2 betrachteten Raumimpulsantworten eine Zeitdifferenz von ca. 12,5 min aufweisen und ein SDR von 22,21 dB liefern. Die erste RIR wird dabei als Referenzgröße betrachtet.

Die Temperaturdifferenz in dieser Zeit ist so gering, sodass sie mit einem Thermometer nicht seriös ermittelt werden könnte. Um also die Temperaturdifferenz zu erhalten und diese dann kompensieren zu können, werden die 2 Impulsantworten in einer Zeitbereichsdarstellung übereinander gelegt (siehe Abb. 5.16), um die Richtung der Zeitskalierung, die absolut ca. 1 Abtastwert bezüglich der Gesamtlänge von 32000 Abtastwerte beträgt, zu sehen.

Da in diesem Fall ein Temperaturrückgang festgestellt wurde, wird nun die zweite Impulsantwort versuchsweise um $k \cdot 1/100$ Abtastwerte gestaucht ($k=50$ bis 150). Dabei werden jedoch aus Zeitgründen (Komplexität: $N \log N + 4N^2$) nur die ersten 5000 Abtastwerte gestaucht und mit der Referenz verglichen (dementsprechend war die absolute Zeitskalierung nur $5/32$ des Gesamtwertes).

Alle gestauchten Impulsantworten werden dann auf ihr SDR bezüglich der Referenz-RIR untersucht. Anhand des dabei beobachteten Maximalwertes des SDR kann somit der Zeitskalierungsfaktor ermittelt werden, der dann zum Kompensieren für die gesamte RIR Folgenlänge von 32000 Werten verwendet wird.

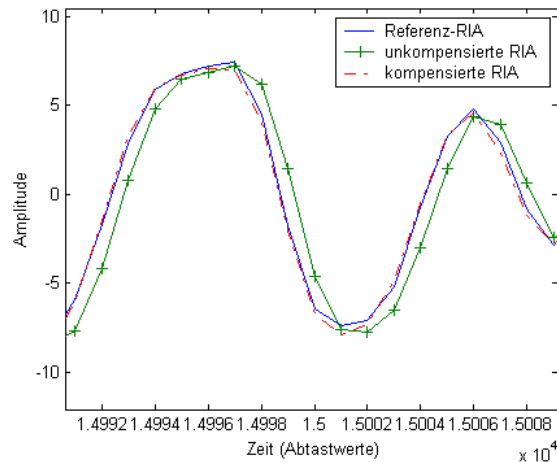


Abb. 5.16: Detailausschnitt von 2 RIRs bei ca. 0,02 °C Temperaturänderung

Der differenzielle Zeitskalierungseffekt δ wurde schließlich mit $-3,65 \cdot 10^{-5}$ ermittelt. Wird dieser nun über (5.13), (5.14) auf die 2. Impulsantwort angewandt, so ergibt sich die in Tab. 5.6 angeführte Verbesserung des SDR, wobei mehrere Frequenzbereiche getrennt analysiert wurden. Es zeigen sich vor allem Verbesserungen im mittleren und hohen Frequenzbereich. Im tiefen Frequenzbereich bei großen Wellenlängen wirkt sich eine Temperaturveränderung also kaum aus, da aufgrund der starken Überabtastung bei tiefen Frequenzen benachbarte Abtastwerte fast identisch sind. Die Kompensation bringt dementsprechend auch wenig Verbesserung.

Frequenzbereich [Hz]	SDRF [dB]	SDRF [dB] kompensiert
0 – 100	26,0	26,0
101 – 200	45,7	46,6
201 – 500	42,8	44,9
501 – 1000	38,8	45,5
1001 – 2000	28,3	33,0
2001 – 3400	24,9	33,8
3401 – 8000	23,5	34,1
8001 – 16000	22,6	30,6
16001 – 22050	17,1	19,5

Tab. 5.9: SDRF-Verbesserung durch Temperaturkompensation

Es zeigt sich somit, dass vor allem der hohe Frequenzbereich das SDR beeinflusst. Das gesamte SDR der beiden Impulsantworten verbesserte sich durch die Temperaturkompensation von 22,21 dB auf 27,45 dB, was deutlich unter dem SDR-Wert zweier benachbarter Impulsantworten liegt (Zeitdifferenz der Messzeitpunkte 1,5 sec). Man kann daher feststellen, dass die inverse lineare Zeitskalierung nicht alle Komponenten der Drifterscheinung in den RIAs berücksichtigt. Es bleiben beispielsweise Temperaturunterschiede (Inhomogenitäten) innerhalb des Raumes unberücksichtigt. Verändert sich nur in einem Teil des Raumes die Temperatur, so ändert sich auch die Schallgeschwindigkeit nur für die Schallfronten, die diesen Bereich durchdringen. Die verschiedenen Schallfronten müssten demnach unterschiedlich temperaturkompensiert werden, wodurch aber die Komplexität des Systems enorm steigen würde. Es können somit nur globale Temperaturdifferenzen kompensiert werden. Außerdem dürfen Drifterscheinungen, die vom Messequipment herrühren, nicht vergessen werden. Hierzu zählen beispielsweise Frequenzjitter des A/D-D/A-Wandlers oder eine geringe Drift der Verstärkungsfaktoren im Aufnahmesystem. Auch auf Veränderungen des Luftdrucks bzw. der Luftfeuchtigkeit sollte nicht vergessen werden, wobei diese im Vergleich zur Temperaturänderung weniger relevant im Bezug auf die Schallgeschwindigkeit sind.

1.22.2 Erwärmungseffekt des Lautsprechers

Für diesen Punkt werden die Messdaten aus Messung 4.4.1.3 verwendet, die einer Langzeitmessung durch Messpausen unterschiedlicher Länge entspricht. Dabei ist in den Abbildungen 5.17 auf der vertikalen Achse die Amplitude des Direktschallimpulses der RIR aufgetragen. Abb. 5.17a (mit Lautsprechersignalpegel von -12dB gemessen) zeigt einen „intuitiven“ Verlauf des zeitvarianten Zeitbins nach jeder der 4 Messpausen. Anhand Abb. 5.17b (Lautsprecherpegel von 0 dB) sieht man jedoch, dass der Zeitbin der Impulsantwort nach jeder Messpause von einem lokalen Maximalwert abfällt, was nicht einer erwarteten Verlängerung der Kurve vor der jeweiligen Pause entspricht.

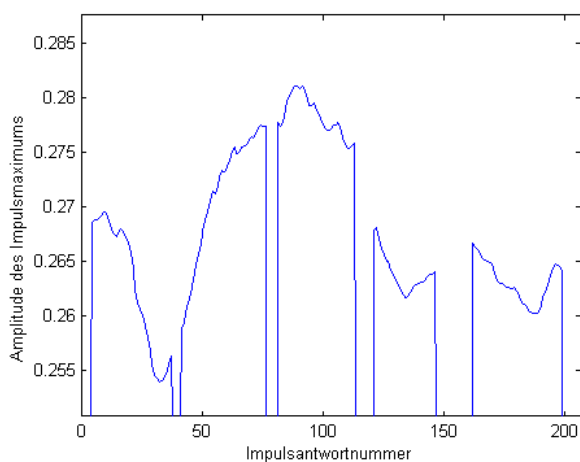


Abb. 5.17a: RIR Spitzenwert über 200 RIRs bei max. Lautsprecherpegel

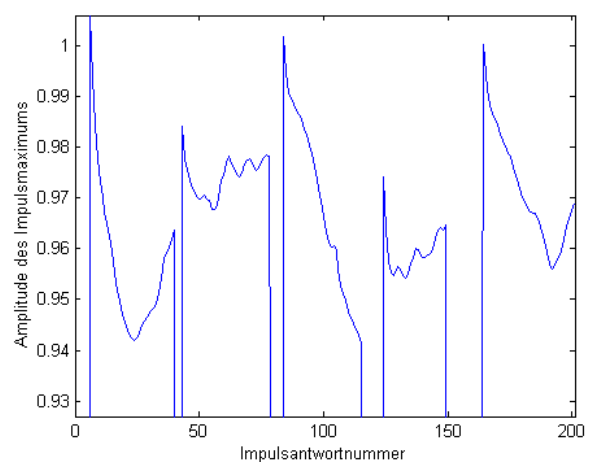


Abb. 5.17b: RIR Spitzenwert über 200 RIRs bei Lautsprecherpegel von -12dB

Die Erklärung dieses unterschiedlichen Verhaltens liegt in der Erwärmung des Lautsprechers während der Beschallung mit hohem Signalpegel. In der Messpause kühlt dieser wieder etwas ab, was also bei Wiederbeginn der Beschallung zu einem transienten Erwärmungsprozess führt, welcher in Abb. 5.17b ca. 5 MLS-Perioden lang mit einer Anfangserhöhung des Zeitbins von bis zu 2 % sichtbar wird.

1.23 Impulsantwortschätzung aus Sprachdaten

Hierfür wurden die Aufnahmedaten aus 4.4.1.4 verwendet.

Um aus den Sprachdaten die RIR zu berechnen, kann nicht derselbe Algorithmus (2.19) wie bei MLS Signalen verwendet werden, da die Sprachsignale im Vergleich zur MLS sehr wohl korrelieren.

Die RIR muss also über die Beziehung (2.23) berechnet werden. Dabei ist jedoch zu beachten, dass nur Blöcke verwendet werden können, die durch Sprachpausen eingegrenzt sind, die länger als die Nachhallzeit sind. Die Zeitdauer von A – B und C – D muss somit größer als die Nachhallzeit des Raumes sein.

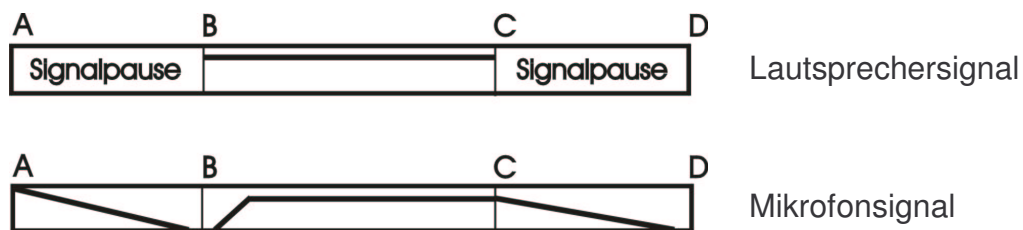


Abb. 5.18: Voraussetzungen zur Berechnung der Impulsantwort bzw. Anregung des Raumes mit Sprechsignalen

Abb. 5.19 zeigt nun das, für die Berechnung verwendete Sprachsignal, welches der Zeitdauer von B – D in Abb. 5.18 entspricht.

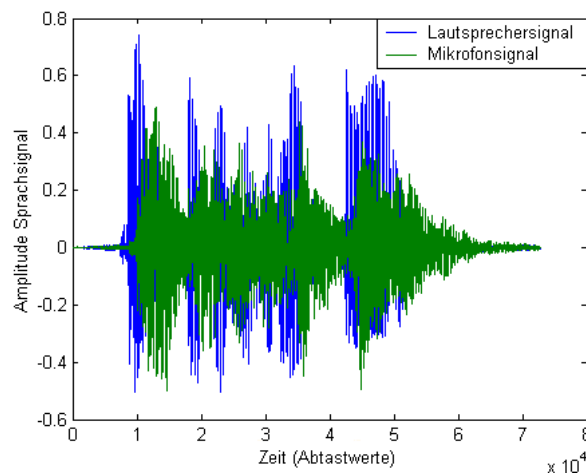


Abb. 5.19: Verwendetes Sprachsignal

Über (2.23) erhält man nun folgenden Frequenzgang der Impulsantwort.

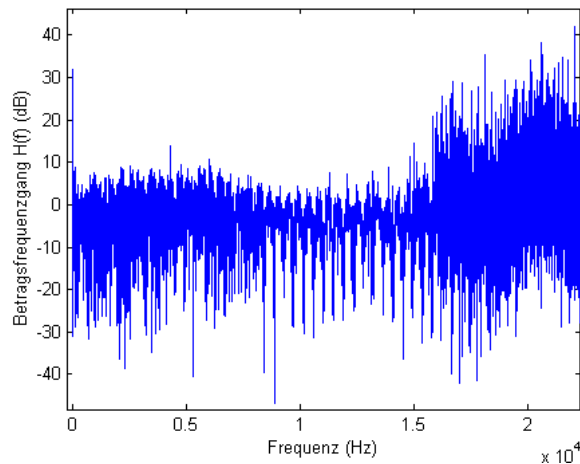


Abb. 5.20: Frequenzgang der berechneten Raumimpulsantwort

Da für das Anregungssignal des Raumes eine MP3-Datei verwendet wurde, welche mit 16 kHz bandbegrenzt ist, kann die Übertragungsfunktion auch nur bis 16 kHz bestimmt werden. Im Bereich über 16 kHz erfolgt die Division zweier kaum ausgesteuerter Rauschsignale zu Frequenzbereich, ohne jede Aussagekraft. Dieser Bereich wird daher durch Filterung (Rechteckfenster über FFT-Block) entfernt. Der Frequenzbereich unter 100 Hz wurde ebenfalls weggefiltert, da auch hier die Raumimpulsantwort nicht reproduzierbar ermittelt werden konnte.

Nun wird die Impulsantwort mit einer aus den MLS-Daten ermittelten Impulsantwort verglichen (der Zeitunterschied der Messzeitpunkte beider RIRs betrug ca. ½ Minute), bei der die gleiche Bandbegrenzung auf 100 – 16000 Hz durchgeführt wurde, um die 2 RIRs prinzipiell vergleichen zu können.

Frequenzbereich [Hz]	SDRF [dB]
101 – 200	30,0
201 – 500	30,6
501 – 1000	26,8
1001 – 2000	23,6
2001 – 3400	20,5
3401 – 8000	27,1
8001 – 16000	17,7

Tab. 5.10: SDRF einer aus Sprachdaten geschätzten Impulsantwort

Für die in Tab. 5.7 berechneten SDRF-Werte, wurden die beiden RIRs bei 20000 Abtastwerte abgeschnitten, da die aus den Sprachdaten berechnete Raumimpulsantwort dahinter nicht mehr reproduzierbar ermittelt werden konnte. Generell kann die RIR jedoch erstaunlich gut anhand der Sprachdaten ermittelt werden.

Es muss jedoch einschränkend folgendes festgestellt werden:

- Der Beschallungspegel des Raumes während der Messung war höher als normaler Gesprächspegel
- Die Messung erfolgte sozusagen „steril“, da dem Messsignal kein Grundgeräuschpegel überlagert war
- Aufgrund des, für die Schätzung der RIR gewählten Algorithmus, müssen immer die erforderlichen Signalpausen vorhanden sein, um eine neue RIR berechnen zu können

Im realen Anwendungsfall des Echokompensators können also nicht zu hohe SDRF-Werte erwartet werden, da vor allem der Grundgeräuschpegel die richtige Bestimmung der Impulsantwort im späteren Zeitbereich erschwert.

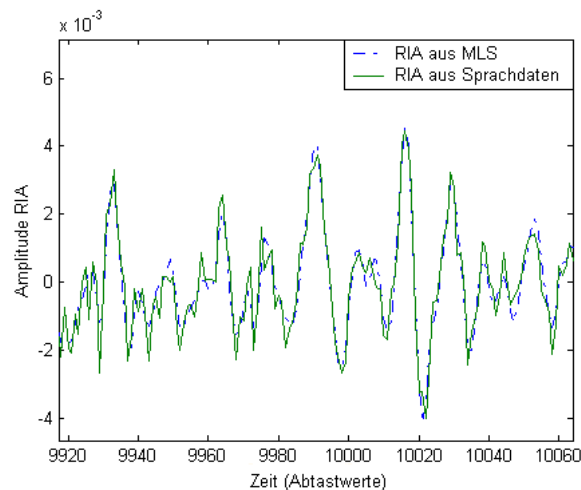


Abb. 5.21: Vergleich RIR aus MLS zu RIR aus Sprachdaten

Abb. 5.21 zeigt, dass die aus Sprachdaten geschätzte Impulsantwort im Bereich um 10000 Abtastwerte noch ziemlich genau ermittelt werden kann.

Messergebnisse bei nichtstationären Raumverhältnissen

1.24 Untersuchung zur Homogenität des Arraydatensatzes

6.1.1 Kompensation der Mikrofonverstärkung

Die aufgenommenen Daten der Kalibrierungsmessung mussten zuerst im Bereich von 1 kHz bandpassgefiltert werden (Bandbreite 400 Hz), um ein störungsfreies Kalibriersignal (1 kHz) zu erhalten.

Messmikrofon	Verstärkungsfaktoren Kalibrierungs- messung 1	Verstärkungsfaktoren Kalibrierungs- messung 2	Endgültige Verstärkungs- faktoren
1	1,111	1,092	1,106
2	1,051	1,052	1,051
3	1,150	1,164	1,153
4	1,117	1,130	1,120
5	1,110	1,132	1,115
6	1,000	1,000	1,000

Tab. 6.11: Verstärkungsfaktoren zur Kompensation der Mikrofonvorverstärkereinstellungen

Da die zweite Kalibrierungsmessung erst 12 Stunden nach Ende der Arraymessung erfolgte, werden die Faktoren nur mit 25 % gewichtet, die Verstärkungsfaktoren der ersten Messung mit 75 %.

Wie aus Tab. 6.1 ersichtlich, weist Mikrofon 5 die maximale Verstärkungsdrift auf. Nimmt man den Verstärkungsdriftverlauf zwischen Messung 1 und Messung 2 linear an, so beträgt der Fehler zwischen erster und letzter Teilarraymessung 1,1 %, was ca. 40 dB SDR entsprechen würde.

Die endgültigen Verstärkungsfaktoren wurden schließlich bei der Impulsantwortberechnung der einzelnen Mikrofonspuren berücksichtigt.

6.1.2 Kompensation der Mikrofonfrequenzgänge

Vergleicht man die Impulsantworten der verschiedenen Mikrofone, auf gleichen Energieinhalt normiert, so zeigen sich SDR Werte von :

Mikrofon-Nummer	SDR [dB]
1	Ref.
2	21,0
3	15,5
4	15,7
5	13,4
6	12,4

Tab. 6.12: SDR-Werte ohne Frequenzgangkompensation

Diese Werte scheinen auf den ersten Blick mit einer Drift behaftet zu sein (mit den Mikrofonen 1 bis 6 wurde zeitlich hintereinander an der jeweils selben Position und dem selben Mikrofonvorverstärker gemessen), jedoch zeigte sich im Frequenzbereich die Erscheinung, dass das SDRF für hohe Frequenzen durchaus höher sein kann als in tieferfrequenten Bereichen. Das weist darauf hin, dass die Mikrofone unterschiedliche Frequenzgänge haben. Somit müsste also das SDR durch Frequenzgangentzerrung noch verbessert werden können.

Dazu soll die Übertragungsfunktion $\underline{V}(f)$ das Verhältnis zweier Impulsantworten angeben:

$$\underline{V}(f)_{1,2} = \frac{\underline{H}(f)_{Mic1}}{\underline{H}(f)_{Mic2}} \quad (6.101)$$

Würde man Mikrofon 1 als Referenz verwenden, so wäre der zeitliche Unterschied zwischen den Messungen größer, als wenn ein Mikrofon, mit dem in der Mitte der Messreihe gemessen wurde, als Referenzgröße dient. Für die Vergleichsfunktionen wird also eine RIR des Mikrofons 4 verwendet. Somit müssen für 6 Mikrofone 5 Vergleichsfunktionen gebildet werden.

In Abb. 6.1 ist beispielsweise $\underline{V}(f)_{4,5}$ in Betrag und Phase dargestellt.

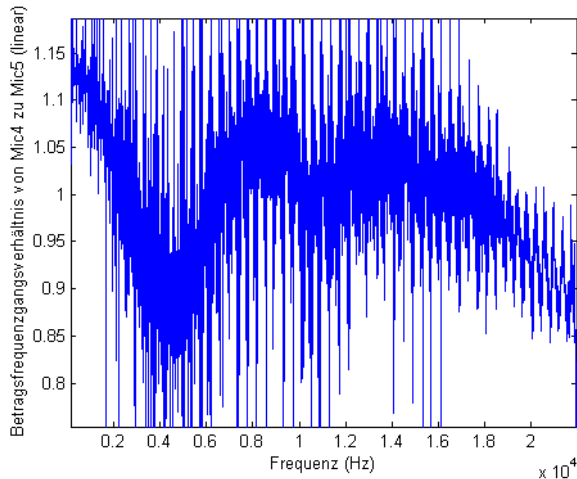


Abb. 6.1a: Betrag $\underline{V}(f)_{4,5}$

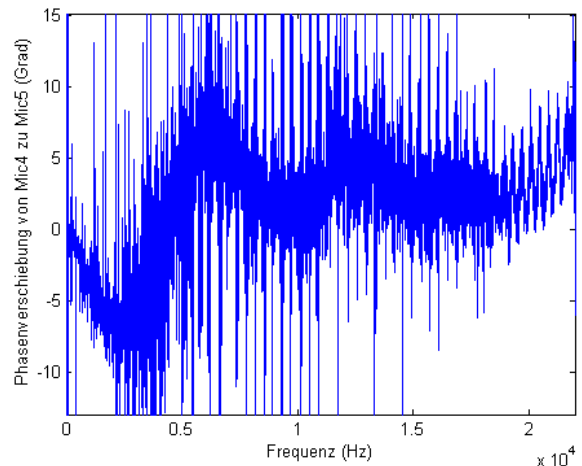


Abb. 6.1b: Argument $\underline{V}(f)_{4,5}$

Die exakte Entzerrung zwischen den Impulsantworten 4 und 5 wäre nicht sinnvoll, da offensichtlich starke zufällige, impulsartige Störungen der gesuchten Verstärkungsfunktion überlagert sind. Man würde zwar, falls Impulsantwort 4 als Referenzimpulsantwort verwendet wird, für diese optimale SDR-Werte erhalten; wird jedoch eine andere Impulsantwort als Referenz gewählt, so wäre das SDR sehr klein.

Es wäre auch kaum plausibel, dass 2 Mikrofone einen Frequenzgang, wie in Abb. 6.1a zu sehen ist, haben. Um also zu einem brauchbaren Ergebnis zu kommen, werden die Verstärkungs- und die Phasenverhältnissfunktionen jeweils tiefpassgefiltert. Um dabei die Einschwingvorgänge der gefilterten Funktionen an den Randbereichen möglichst zu vermeiden, wurde hierfür ein Nullphasenfilter verwendet. Die Filterung wurde mit einem Tschebychef-Tiefpass 4. Ordnung mit Grenzfrequenz 30 Hz durchgeführt und führte zu folgendem Ergebnis:

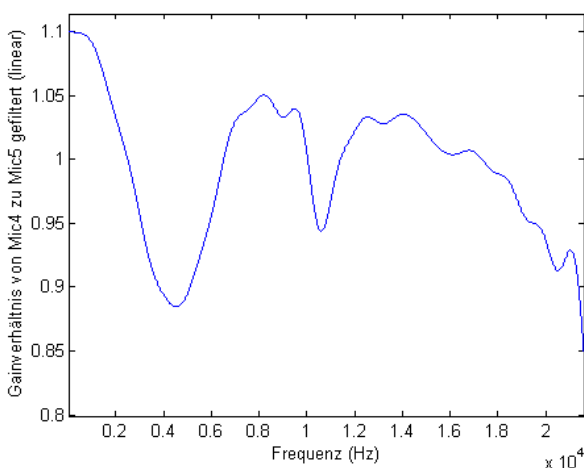


Abb. 6.2a: Betrag $V(f)_{4,5}$ gefiltert

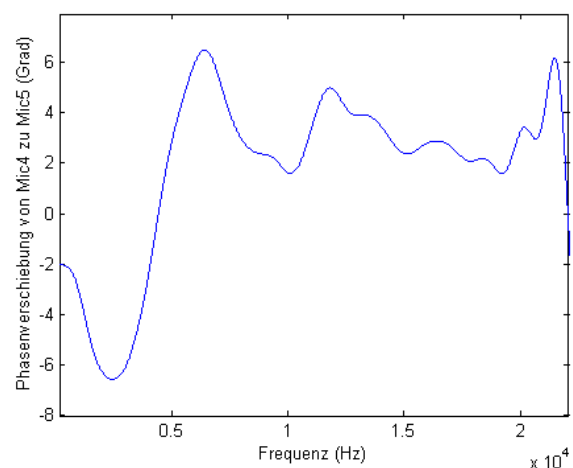


Abb. 6.2b: Argument $V(f)_{4,5}$ gefiltert

In Abb. 6.2 sind deutliche Unterschiede zwischen den 2 Mikrofonen zu sehen, die vor allem im tieffrequenten Bereich zu sehr schlechten SDRF-Werten führen. Werden nun die Impulsantworten der verschiedenen Mikrofone durch die gefilterten Vergleichsfunktionen multipliziert, so ergeben sich bedeutend bessere SDR-Werte als vorher. Als Referenzimpulsantwort soll nun wieder die des ersten Mikrofons gelten.

Es ergibt sich:

Mikrofon-Nummer	SDR [dB]
1	Ref.
2	24,8
3	22,5
4	24,5
5	25,3
6	24,5

Tab. 6.13: SDR-Werte mit Frequenzgangkompensation

1.24.1 Temperaturkompensation

Anhand der in Abschnitt 5.8.2 beschriebenen Temperaturkompensation, werden nun die Impulsantworten des zusätzlichen Temperaturmessmikrofons der Teilarraymessungen auf eine als Referenzimpulsantwort dienende RIR der ersten Teilarraymessung kompensiert, um die Temperaturveränderung während der Messung abschätzen zu können.

Nummer der Teilarraymessung	Differenzielle Zeitskalierungsfaktor δ für die Temperaturkompensation
1	0,000000
2	0,000005
3	0,000015
4	0,000028
5	0,000035
6	0,000048
7	0,000030
8	0,000055
9	0,000083
10	0,000100
11	0,000128
12	0,000143
13	0,000190
14	0,000253
15	0,000245
16	0,000310

Tab. 6.14: Differenzierende Zeitskalierungsfaktoren zum Kompensieren der Temperaturveränderung

Es stellte sich also heraus, dass die Temperatur von der 1. bis zur 16. Messung um $0,177^{\circ}\text{C}$ anstieg. Aufgrund des enormen Rechenaufwandes wurden jedoch die gesamte Datenbank mit den Arraymessungen nicht temperaturkompensiert. Die Ergebnisse in Tab. 6.4 dienen vor allem zum Abschätzen der Auswirkung der Temperaturveränderung auf die Arraymessung. Bei Bedarf können natürlich nachträglich die in der Datenbank vorliegenden Messdaten anhand dieser Tabelle temperaturkompensiert werden.

In Abb. 6.3 wird nun die Auswirkung der Temperaturerhöhung auf die RIR der Kompensationsmessung in den jeweiligen Frequenzbereichen dargestellt.

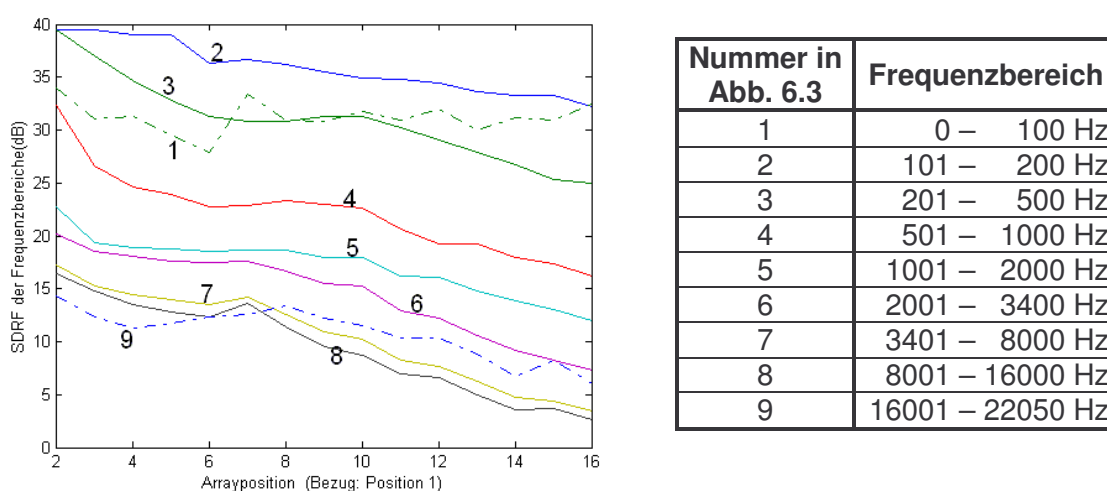


Abb. 6.3: Auswirkung der Temperaturveränderung

Jedoch muss festgestellt werden, dass die Temperaturdrift in nur sehr begrenztem Maße kompensiert werden kann. Es wurde anhand der letzten (16.) Messung versucht, diese in Bezug auf die RIR der ersten Messung zu kompensieren (die beiden RIRs wurden in einem Zeitabstand von ca. 12 Stunden aufgenommen), was jedoch nur auf ein SDR von 12 dB führte. Die Raumimpulsantworten sind also noch mit weiteren, nicht genau identifizierbaren Langzeitdrifterscheinungen beaufschlagt. Diese Drifterscheinungen verursachen somit den größten Fehler des Arraydatensatzes.

1.24.2 Geschwindigkeitskompensation

Die, aus (4.13) erhaltenen Offsetwerte, welche zur Kompensation der unterschiedlichen Geschwindigkeiten des Antriebsmotors dienen, sind in Tab. 6.5 angegeben.

Nummer der Teilarraymessung	Offsetwerte [Abtastwerte]
1	65535
2	64681
3	64223
4	64001
5	63081
6	63885
7	64853
8	64200
9	63906
10	63391
11	63290
12	64108
13	63704
14	65363
15	64668
16	64775

Tab.6.15: Offsetdaten zur Geschwindigkeitskompensation

Diese Offsetwerte wurden bei der Impulsantwortschätzung als Offset berücksichtigt (siehe Abb. 2.10).

1.25 Vier-Quadranten-Darstellung

Wie man in Kapitel 5 schon gesehen hat, liegen die Messergebnisse in der Darstellung $M(m, n)$ vor.

Diese Matrix reicht aus, um Impulsantworten in stationären Räumen zu erfassen. Ändern sich jedoch die geometrischen Gegebenheiten im Raum (z.B. Gegenstand wird verschoben, Fenster wird geöffnet oder Mikrofon wird bewegt), so muss die Ortskoordinate des bewegten Gegenstandes ebenfalls in die Matrix eingehen.

Wir ersetzen nun die Nummer der Raumimpulsantwort mit der Ortskoordinate des bewegten Objektes. Zur Unterscheidung wird der Parameter m in Formel (5.2) durch w ersetzt.

$$M(w, n) = \begin{bmatrix} h_{1,1} & \dots & h_{1,N} \\ h_{2,1} & \ddots & \vdots \\ \vdots & h_{w,n} & \vdots \\ h_{(w-1,1)} & \ddots & \vdots \\ h_{w,1} & \dots & h_{w,N} \end{bmatrix} \quad \begin{array}{l} 1 \leq w \leq W \\ 1 \leq n \leq N \end{array} \quad (6.102)$$

w Parameter für zeitdiskrete Abtastung der Ortskoordinate des bewegten Objekts

Nun sei festgehalten, dass der w -te Zeilenvektor dieser Matrix einer Impulsantwort, bei der sich das Hindernis an der Stelle $x = w \cdot T_{SO}$ befindet, entspricht. T_{SO} stellt dabei die Wegstrecke dar, welche das Hindernis von einer Impulsantwortmessung auf die nächste zurücklegt. Die Spaltenvektoren von $M(w, n)$ stellen dabei den örtlichen Verlauf eines beliebigen Zeitbins $h(n)$ der Impulsantwort in Bezug auf das bewegte Hindernis dar. Die Zeilenvektoren von $M(w, n)$ bezeichnen die jeweilige Impulsantwort, die dem Raum mit dem Hindernis an der Position $x = w \cdot T_{SO}$ zugeordnet werden.

Die Ort-, Zeitdarstellung der Matrix $M(w, n)$ kann nun in 3 weitere Darstellungsformen übergeführt werden:

Ort-, Zeit-, Darstellung	Ort-, Frequenz – Darstellung
Ortsfrequenz, Zeit – Darstellung	Ortsfrequenz, Frequenz – Darstellung

Abb. 6.4: Darstellungsmöglichkeiten der Aufnahmedaten

Da die Matrix $M(w, n)$ in zeitdiskreter Form vorliegt, werden nun die weiteren Größen zur zeitdiskreten Darstellung von Ortsfrequenz und Frequenz hergeleitet:

Als erstes sei das betrachtete Ortsfenster X_o angegeben:

$$X_o = 1830mm \quad \text{betrachtetes Ortsfenster}$$

Die Ortsfrequenzauflösung wird durch Δv angegeben:

$$\Delta v = \frac{1}{X_o} = 0,546 \frac{1}{m} \tag{6.103}$$

Das Abtastintervall T_{SO} des Ortes erhält man folgendermaßen:

$$T_{SO} = \frac{X_o}{N} = \frac{1830mm}{1048} = 0,001746m = 1,746mm \tag{6.104}$$

N Anzahl der Ortsabtastwerte

Die Abtastfrequenz ν_{SO} des Ortes erhält man folgendermaßen:

$$\nu_{SO} = \frac{1}{T_{SO}} = 572,68 \frac{1}{m} \quad (6.105)$$

Der Übergang vom zeitdiskreten in den kontinuierlichen Bereich passiert folgendermaßen:

$$x = w \cdot T_{SO} \quad (6.106)$$

$$\nu = p \cdot \Delta\nu \quad (6.107)$$

$$t = n \cdot T_A \quad (6.108)$$

$$f = k \cdot \Delta f \quad (6.109)$$

T_A bezeichnet das zeitliche Abtastintervall der Raumimpulsantwort ($\frac{1}{f_A}$, $f_A = 44,1 \text{ kHz}$)

Der Übergang zwischen zeitdiskretem Bereich (hier erfolgt die Berechnung) und zeitkontinuierlichem Bereich (grafische Darstellung der Matrix) wird nun vollständig angegeben:

<u>Analoge Notierung</u>		<u>Zeitdiskrete Notierung</u>	
$M(x, t)$	$\underline{M}(x, f)$	$M(w, n)$	$\underline{M}(w, k)$
$\underline{M}(\nu, t)$	$\underline{M}(\nu, t)$	$\underline{M}(p, n)$	$\underline{M}(p, k)$

Um nun die sämtlichen Darstellungsmöglichkeiten zu erhalten (vier Quadranten), sind folgende Übergänge nötig:

$$M(x, t) \rightarrow \underline{M}(x, f)$$

$$M(x, t) \rightarrow \underline{M}(\nu, t)$$

$$M(x, t) \rightarrow \underline{M}(\nu, f)$$

Die Berechnung der Matrix im jeweiligen Quadranten erfolgt durch die im folgenden angegebenen Fourier-Transformationen.

$$\underline{M}(w, n) \Rightarrow \underline{M}(w, k)$$

$$\underline{M}(w, k) = \sum_{n=0}^{N-1} \underline{M}(w, n) e^{\frac{-j2\pi \cdot n \cdot k}{N}} \quad (6.110)$$

$$\underline{M}(w, n) \Rightarrow \underline{M}(p, n)$$

$$\underline{M}(p, n) = \sum_{w=0}^{W-1} \underline{M}(w, n) e^{\frac{-j2\pi \cdot w \cdot p}{W}} \quad (6.111)$$

$$\underline{M}(p, k) = \begin{cases} \sum_{w=0}^{W-1} \underline{M}(w, k) e^{\frac{-j2\pi \cdot w \cdot p}{W}} \\ \sum_{n=0}^{N-1} \underline{M}(p, n) e^{\frac{-j2\pi \cdot n \cdot k}{N}} \end{cases} \quad (6.112)$$

Die Berechnung der Ortsfrequenz-, Frequenzdarstellung $\underline{M}(p, k)$ kann sowohl vom Orts-, Frequenzbereich als auch vom Ortsfrequenz-, Zeitbereich erfolgen.

1.26 Darstellung einzelner Vektoren der Matrix M

Die nun in 4 Matrizen vorliegenden Daten aus Punkt. 4.4.2.3 (Arraydatensatz) sollen ausschnittsweise grafisch dargestellt werden, um einen Überblick des Datensatzes zu erhalten. Für diese Analyse werden nur die Daten des Mikrofons 1 an Position 1 dargestellt.

6.1.1 Orts – Zeit – Darstellung

Als erstes wird nun die Matrix in der Orts-Zeit-Darstellung betrachtet. Es werden dabei einige Spaltenvektoren der Matrix gezeigt.

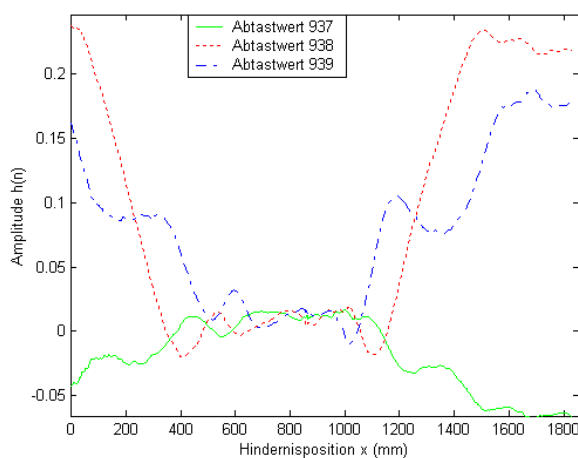


Abb. 6.5a: Verlauf der Zeitbins 937 - 939

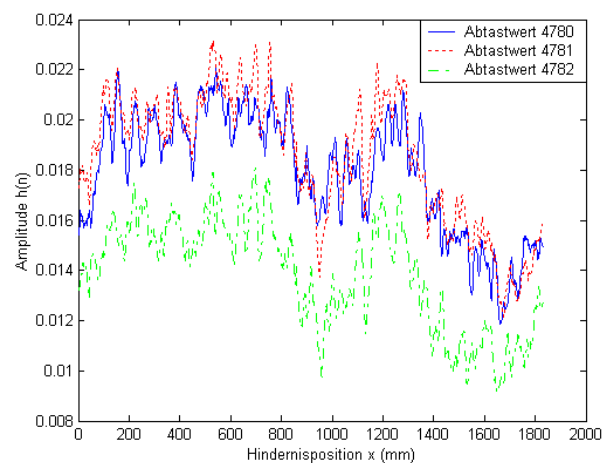


Abb. 6.5b: Verlauf der Zeitbins 4780 - 4782

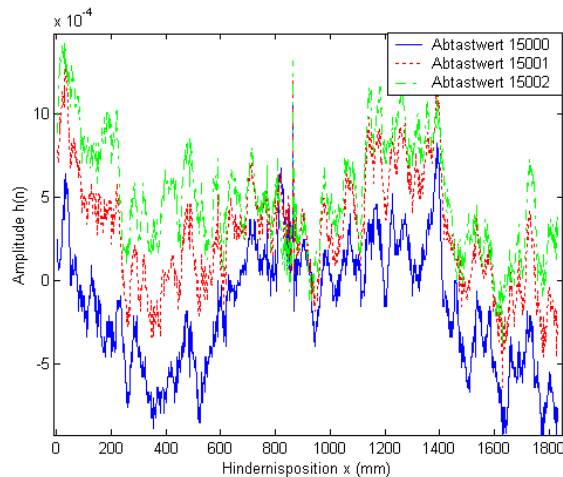


Abb. 6.5c: Verlauf der Zeitbins 15000 - 15002

Ausgehend von $M(w, n)$ sind in Abb. 6.5 verschiedene Abtastwerte der Impulsantwort über dem Ort w dargestellt. Abb. 6.5a. zeigt nun beispielsweise die Abtastwerte

937 – 939, welche die erste Bodenreflexion der Impulsantwort des Mikrofons 1 an Position 1 der Arraymessung darstellen. Es ist dabei die Abschattung der ersten Bodenreflexion durch die Schachtel deutlich zu erkennen. Dieses Bild sollte nun mit Abb. A3 (Anhang A3) verglichen werden.

Interessant ist beispielsweise, dass der Abtastwerte 938, welcher das Maximum der Bodenreflexion darstellt, in einem Bereich von ca. 1,4 m durch die Schachtel, die jedoch nur 81,5 cm lang ist, beeinträchtigt wird. Durch die Schrägstellung des Hinderniswagens zur Direktverbindung Lautsprecher Mikrophon, kommen zwar noch 39 cm dazu, es bleiben jedoch noch ca. 20 cm übrig, wo die Schachtel die Reflexion beeinflusst, obwohl die Direktverbindung Lautsprecher-Boden-Mikrophon nicht durch das Hindernis unterbrochen wird.

Die Abbildungen 6.5 b und c zeigen Abtastwerte vom mittleren Teil der RIR, wo der Einfluss des bewegten Hindernisses nicht mehr erkennbar ist. Durch die Vielzahl von überlagerten Einzelreflexionen kann die Bewegung der Schachtel nicht mehr im Verlauf eines dargestellten Abtastwerts erkannt werden. Je später der Abtastwert der Impulsantwort liegt, desto höher wird dabei seine Ortsfrequenz. Das zugehörige Ortsfrequenzspektrum wird später noch genauer analysiert.

6.1.2 Orts-, Frequenzdarstellung

Es werden nun einige Spaltenvektoren der Matrix $\underline{M}(w,k)$ in folgender Abbildung gezeigt. Die Kurven stellen dabei den Verlauf des jeweils im Bild angegebenen Frequenzbins über den zurückgelegten Weg der Schachtel dar.

In Abb. 6.6a ist der tiefe Frequenzbereich um 50 Hz zu sehen. Man sieht, dass die tiefen Frequenzen relativ konstant bleiben, also nicht durch die Bewegung der Schachtel verändert wurden. Weiters ist noch eine größere Varianz der Kurve zu sehen, welche aufgrund der schlechteren Reproduzierbarkeit der RIR im Bereich unter 100 Hz nicht verwundert. Bei höheren Frequenzbins werden die Kurven zunehmend glatter; es steigt auch die Ortsfrequenz der Frequenzzeiger. In Bild 6.6b sieht man bereits sehr hochfrequente Signale, die aber sehr glatt ist und dadurch eine gute Reproduzierbarkeit der Messergebnisse im hohen Frequenzbereich verdeutlichen.

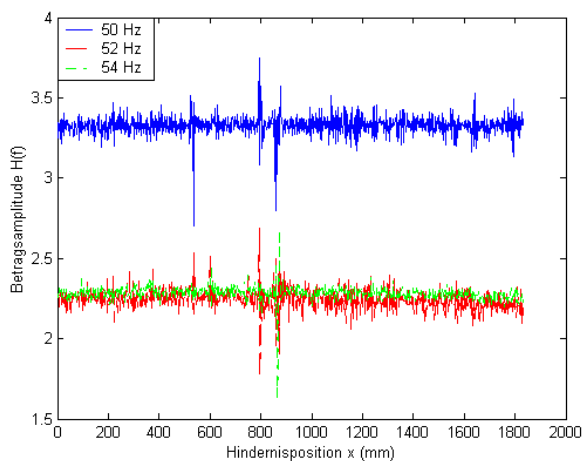


Abb. 6.6a: Frequenzbins im Bereich von 50 Hz

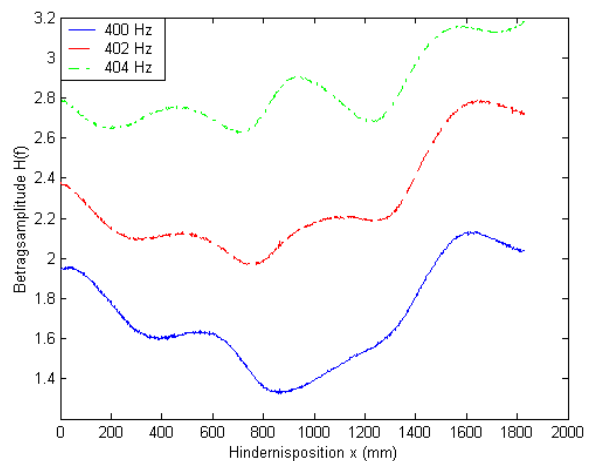


Abb. 6.6b: Frequenzbins im Bereich von 400 Hz

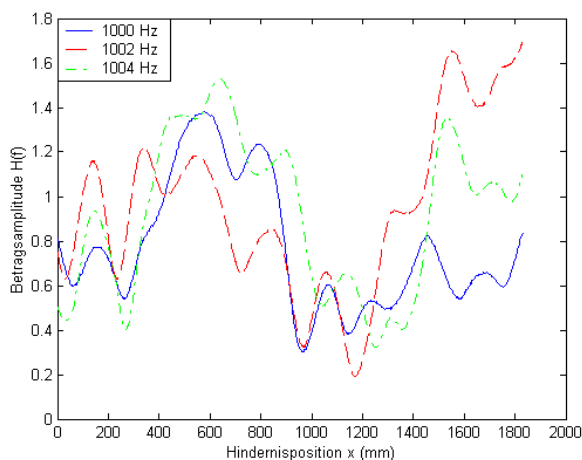


Abb. 6.6c: Frequenzbins im Bereich von 1000 Hz

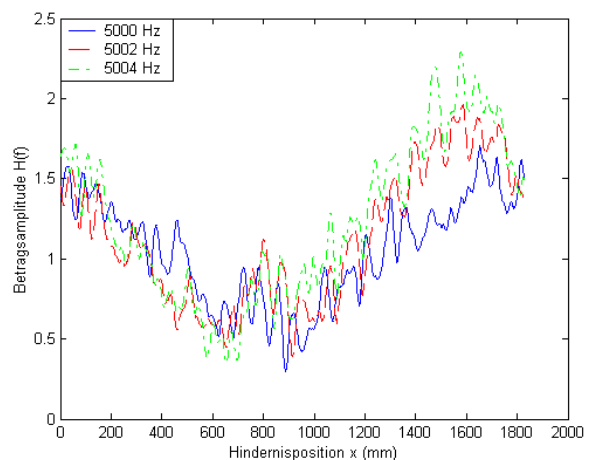


Abb. 6.6d: Frequenzbins im Bereich von 5000 Hz

6.1.3 Ortskurven

In Abb. 6.7. sind nun Ortskurven von komplexen Frequenzbins zu sehen.

In Abb. 6.7a zeigt sich wieder eine große Varianz der Kurve, sie kann also nicht im Detail verfolgt werden. In den Abbildungen 6.7b, 6.7c ist der Weg der Ortskurve jedoch gut ersichtlich. Bei hohen Frequenzen (Abb. 6.7d) wird der Verlauf der Ortskurve aufgrund höherer Ortsfrequenzen wieder schwieriger zu erkennen, ist jedoch noch immer sehr glatt.

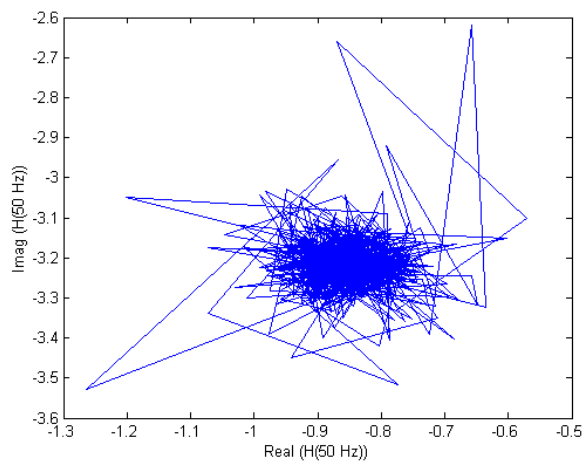


Abb. 6.7a: Ortskurve bei 50 Hz

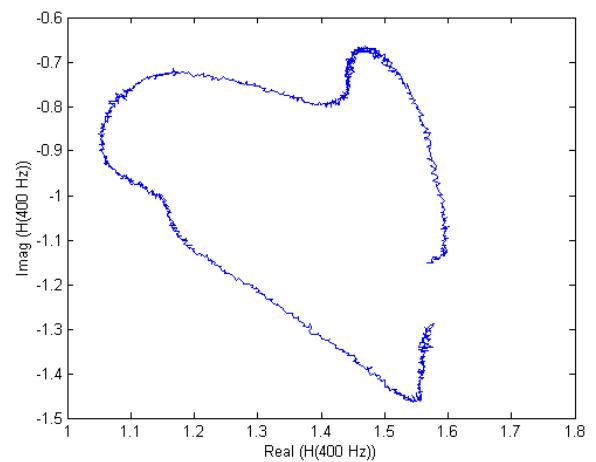


Abb. 6.7b: Ortskurve bei 400 Hz

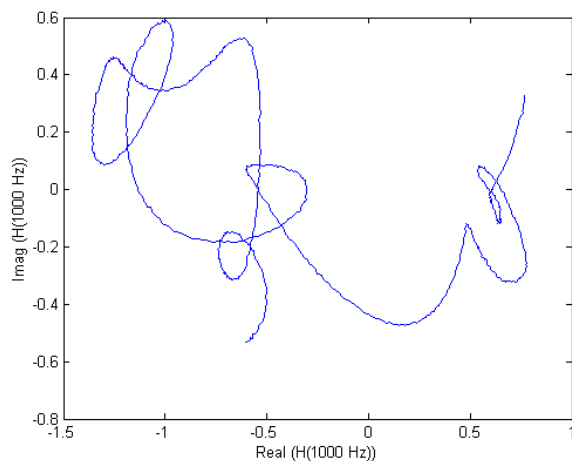


Abb. 6.7c: Ortskurve bei 1000 Hz

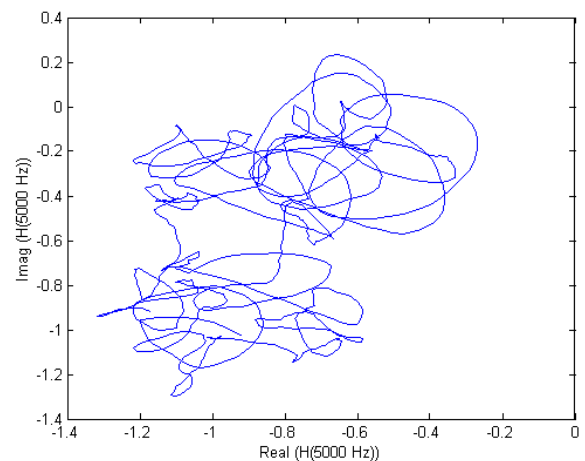


Abb. 6.7d: Ortskurve bei 5000 Hz

1.27 Grafische Vier-Quadranten-Darstellung der Matrix M

Es wird nun die Vier-Quadranten-Darstellung der Matrix M präsentiert. Dabei bezieht sich die Farbgebung des nebenstehenden Farbbalkens immer auf einen logarithmischen Maßstab. Abb. 6.8 zeigt als erstes die Matrix $M(x,t)$.

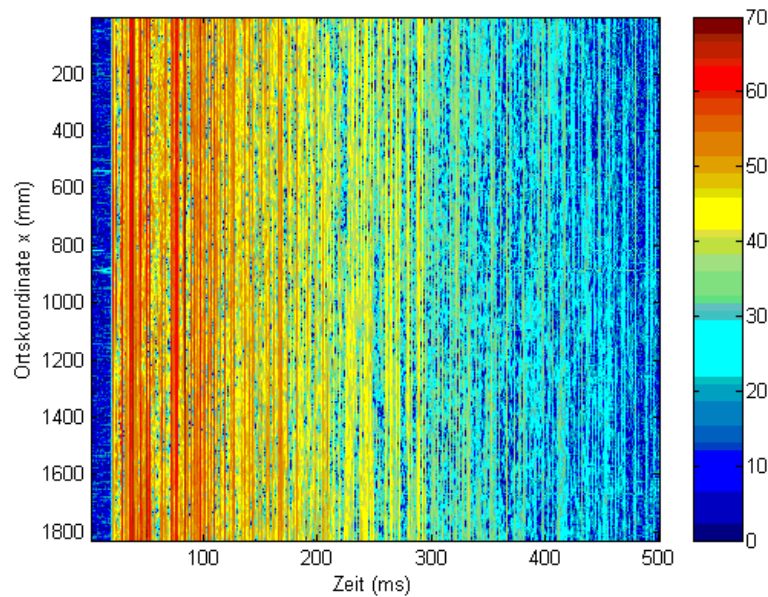


Abb. 6.8: Orts-, Zeitdarstellung

Anhand der Gesamtdarstellung der Matrix $M(x,t)$ in Abb. 6.8 lassen sich keine besonderen Auswirkungen der bewegten Schachtel auf die verschiedenen Raumimpulsantworten ablesen. Es wird nun ein Detailausschnitt dieses Bildes gezeigt.

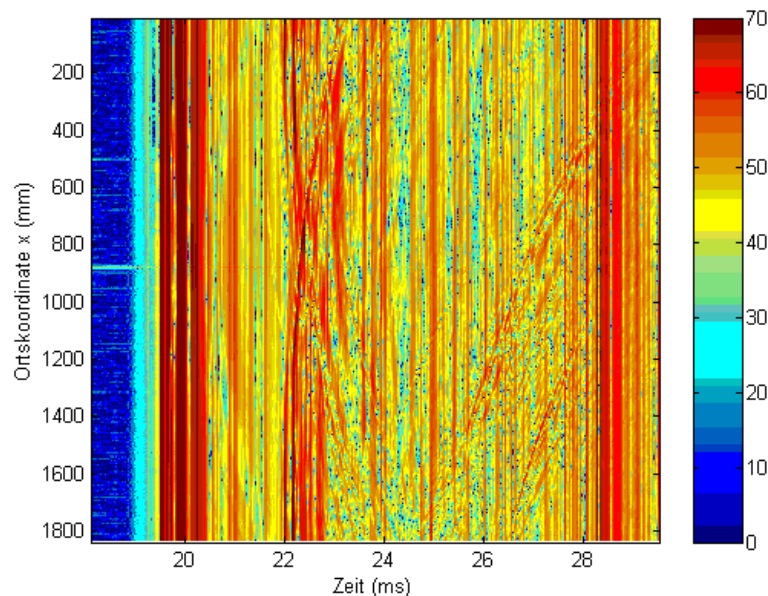


Abb. 6.9: Detailausschnitt aus Abb. 6.8

Im Bereich der ersten Reflexion ist nun die Auswirkung des bewegten Objekts erkennbar. In späteren Zeitbereichen der Impulsantwort verlieren sich jedoch deutlich erkennbare Strukturierungen aufgrund des bewegten Objektes.

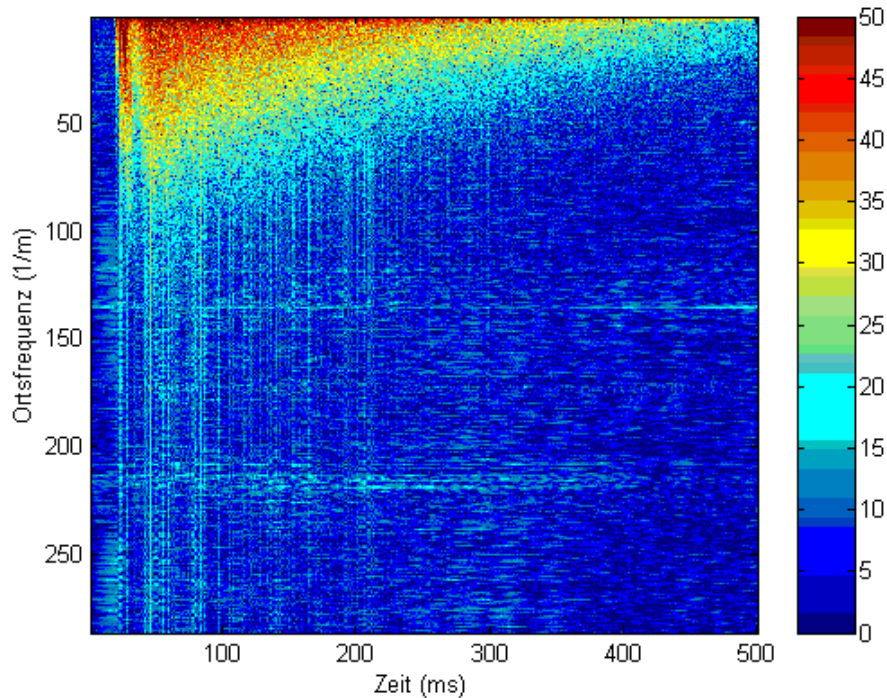


Abb. 6.10: Ortsfrequenz - Zeitdarstellung

Die Ortsfrequenz-, Zeitdarstellung zeigt ein Ortsspektrum, das im früheren Zeitbereich der Raumimpulsantworten ein breiteres Ortsspektrum aufweist, im späteren scheint es schmalbandiger zu werden. Dies ergibt sich jedoch nur aufgrund der Energieabnahme der Impulsantwort über der Zeit.

Um nun eine genauere Aussage über das Spektrum der Ortsfrequenz-, Zeitdarstellung zu tätigen, wird der exponentielle Energieabfall der Raumimpulsantwort kompensiert.

$$h_{komp} = h(n) e^{\alpha n} \quad (6.113)$$

α wird wie folgt berechnet:

$$h(T_n) = 0,001 = e^{\alpha T_n} \quad (6.114)$$

wobei der Schalldruck (Impulsantwortamplitude) bei der Nachhallzeit T_n um -60 dB gesunken ist.

$$\alpha = \frac{\ln 0,001}{T_n} \quad (6.115)$$

Eine derart kompensierte RIR ist in Abb. 6.11 dargestellt.

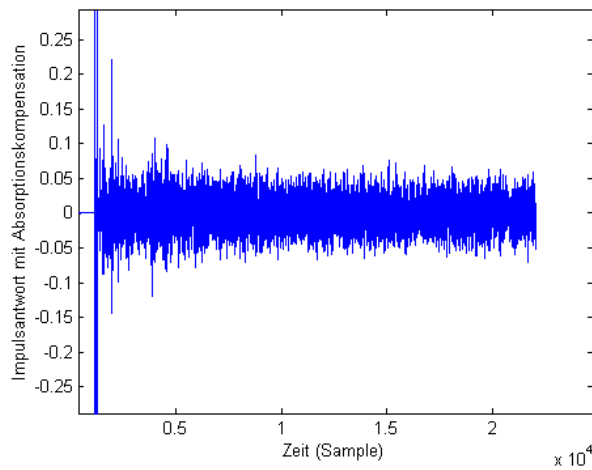


Abb. 6.11: Raumimpulsantwort nach Kompensation des Energieabfalls

Diese Kompensation wurde nun auf alle RIRs der Orts-, Zeitmatrix angewandt, die danach in die Ortsfrequenz-, Zeitdarstellung übergeführt wurde (siehe Abb. 6.12).

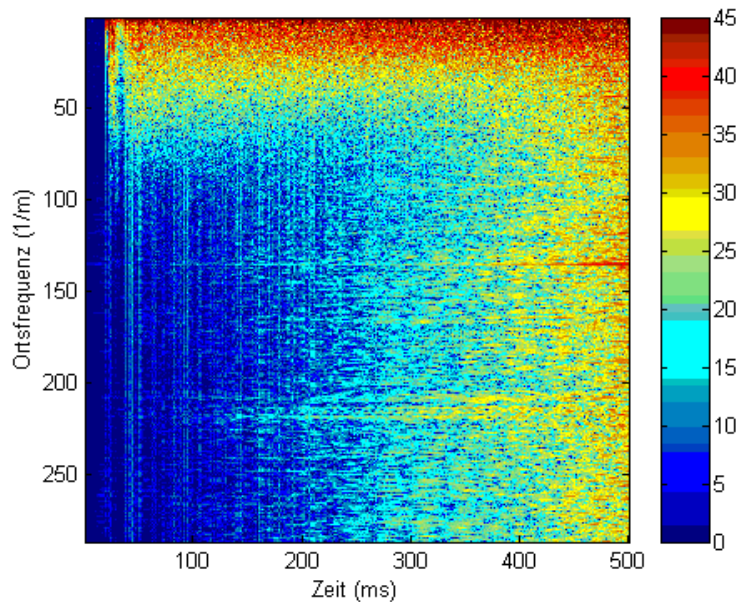


Abb. 6.12: Ortsfrequenz-, Zeitdarstellung
(Kompensation der Absorptionwirkung des Raumes)

Nun kann das tatsächliche Spektrum der einzelnen Zeitwerte der Impulsantwort abgeschätzt werden. Im Bereich innerhalb von $\nu = 25 \frac{1}{m}$ Ortsfrequenz, besitzt das Ortsfrequenzspektrum weißen Charakter. Nach ca. 10000 zeitlichen Abtastwerten, beginnt sich das Ortsfrequenzspektrum auf den ganzen Ortsfrequenzbereich auszudehnen.

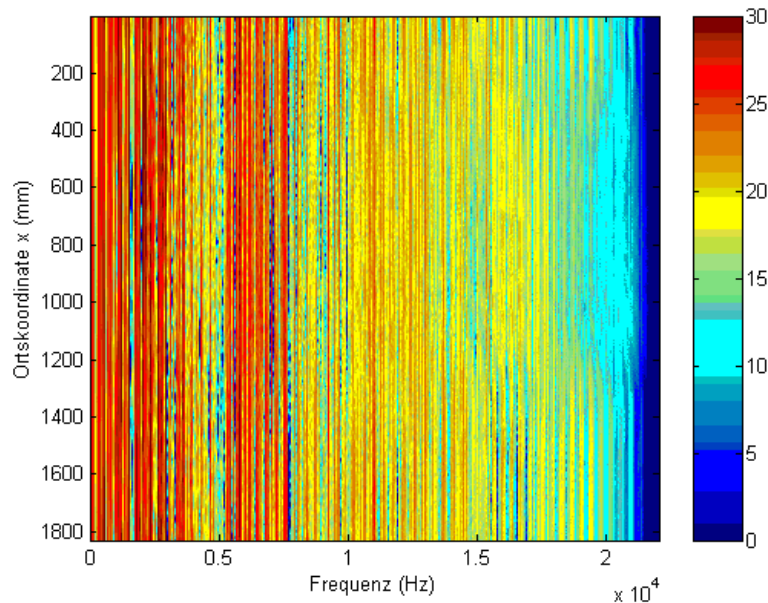


Abb. 6.13: Orts-, Frequenzdarstellung

In der Orts-, Frequenzdarstellung der Matrix M erkennt man im oberen Frequenzbereich ab ca. 15 kHz einen leichten Einbruch des Spektrums, welcher ungefähr der Länge entspricht, wo die Direktverbindung der ersten Bodenreflexion zwischen Lautsprecher und Mikrofon von der Schachtel unterbrochen wurde. Da die Nachhallzeit bei 20 kHz schon deutlich kürzer als 0,75 Sekunden wird, wirkt sich das Wegfiltern der starken ersten Bodenreflexion auf das Frequenzspektrum der Impulsantwort deshalb deutlich sichtbar aus. Im Vergleich zu tieferen Frequenzen, wird durch das Wegfiltern der ersten Bodenreflexion der Impulsantwort also mehr Energie entzogen.

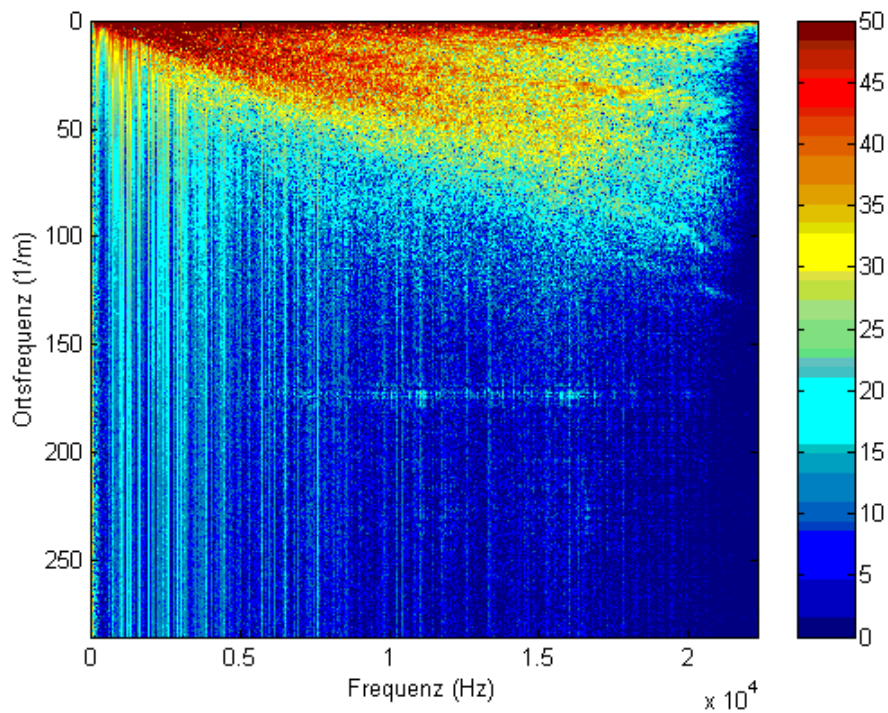


Abb. 6.14: Ortsfrequenz-, Frequenzdarstellung

Abb. 6.14 zeigt nun das Ortsspektrum der Frequenzzeiger von den Abbildungen 6.6. Dabei wird eine deutliche Bandbegrenzung der Frequenzzeiger sichtbar. Die Frequenz der Bandbegrenzung wird wie folgt berechnet:

$$v_{\max} = \frac{2f}{c} \tag{6.116}$$

Für diese empirisch gefundene Bandobergrenze wird im nächsten Abschnitt ein theoretischer Erklärungsversuch entwickelt.

1.28 Synthetischer Betragsfrequenzzeiger

Es soll nun anhand des Strahlenmodells eine Erklärung für die spektrale Zusammensetzung der Frequenzzeiger ausgesucht werden. Abb. 6.15 zeigt die Überlagerung zweier Sinusschwingungen gleicher Frequenz, welche wieder eine Sinusschwingung gleicher Frequenz ergibt.

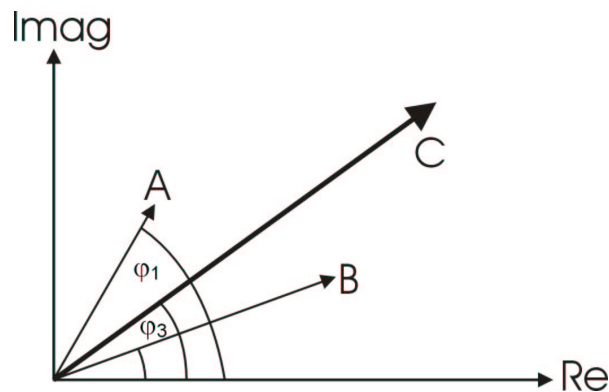


Abb. 6.15: Komplexe Zeigerdarstellung der Überlagerung zweier gleichfrequenter Sinus Signale

$$\underline{x}_1(t) = A e^{j(\omega t + \varphi_1)}$$

$$\underline{x}_2(t) = B e^{j(\omega t + \varphi_2)}$$

$$\underline{y}(t) = \underline{x}_1(t) + \underline{x}_2(t)$$

$$\underline{y}(t) = C e^{j(\omega t + \varphi_3)} = A e^{j(\omega t + \varphi_1)} + B e^{j(\omega t + \varphi_2)} \quad \text{wobei } A, B, C > 0$$

Um nun φ_3 zu ermitteln, kann die Zeitabhängigkeit unberücksichtigt bleiben.

$$C e^{j\varphi_3} = A e^{j\varphi_1} + B e^{j\varphi_2}$$

$$\text{wobei: } C = \left[\left(A \operatorname{Re} \{ e^{j\varphi_1} \} + B \operatorname{Re} \{ e^{j\varphi_2} \} \right)^2 + \left(A \operatorname{Im} \{ e^{j\varphi_1} \} + B \operatorname{Im} \{ e^{j\varphi_2} \} \right)^2 \right]^{\frac{1}{2}}$$

$$\varphi_{3,0} = \arctan \frac{\left(A \operatorname{Im} \{ e^{j\varphi_1} \} + B \operatorname{Im} \{ e^{j\varphi_2} \} \right)}{\left(A \operatorname{Re} \{ e^{j\varphi_1} \} + B \operatorname{Re} \{ e^{j\varphi_2} \} \right)}$$

Durch konjugiert komplexe Erweiterung der komplexen Exponentialschwingung, gelangt man zum reellen Zeitsignal $y(t)$. Es muss jedoch beachtet werden, dass der arctan nur Winkel φ im Bereich von $-\frac{\pi}{2} < \varphi_3 < \frac{\pi}{2}$ liefert.

Alle Winkel außerhalb des Hauptintervalls der arctan-Funktion erhält man mit :

$$\varphi_3 = \varphi_{3,0} + K\pi$$

Obige Überlegungen können nun auf beliebig viele Einzelschwingungen ausgeweitet werden. Die Summe aus beliebig vielen Sinusschwingungen gleicher Frequenz, ist also wieder eine Sinusschwingung.

$$y(t) = \frac{C}{2} \left(e^{j(\omega t + \varphi_3)} + e^{-j(\omega t + \varphi_3)} \right)$$

Es wird nun eine Beschallung des Raumes mit einem Sinus-Ton der Frequenz f_0 angenommen. Ein Mikrofon befindet sich dabei gemäß Abb. 6.16 an Punkt P_1 . Das Mikrofonsignal besteht also aus einer Überlagerung von zahlreichen Schallstrahlen unterschiedlicher Reflexionsordnungen mit der Frequenz f_0 , und, aufgrund verschiedener Reflexionsgrade, verschiedener Amplituden. Es werden dabei sämtliche Schallstrahlen (von Direktschall bis zur n-ten Reflexion am Ende der Impulsantwort) berücksichtigt.

$$\underline{H}(j\omega) = \frac{y(t)}{\underline{x}(t)} \quad \text{wenn } \underline{x}(t) = e^{j\omega t} \quad (6.117)$$

Durch die Ausgangs-, Eingangsbeziehung (6.17) könnte für diese Frequenz die Übertragungsfunktion von Lautsprecher zu Mikrofon bestimmt werden, da im Raum ein stationärer Zustand herrscht.

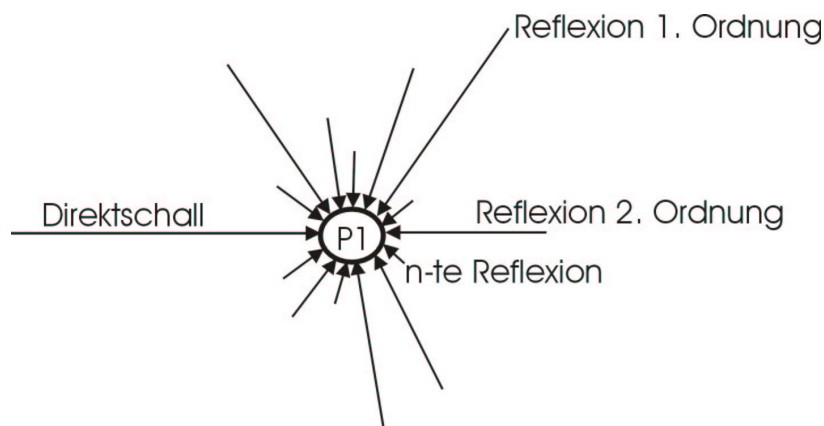


Abb. 6.16: Gesamtsignal durch Überlagerung vieler Einzelstrahlen

Es gibt klarerweise mehrere Reflexionen pro Ordnung, wie in Abb. 1.4 zu sehen ist. Es muss auch die Amplitude einer Reflexion höherer Ordnung nicht zwingend kleiner sein, als die einer Reflexion 1. Ordnung. In Abb. 6.16 repräsentiert nun jeder Pfeil eine Sinusschwingung mit Frequenz f_0 . Insgesamt werden k verschiedene einfallende Schallstrahlen angenommen.

Diese Abbildung wird nun in das nächste Bild übergeführt, wobei wegen der Übersichtlichkeit nur 2 einfallende Schallwellen dargestellt werden.

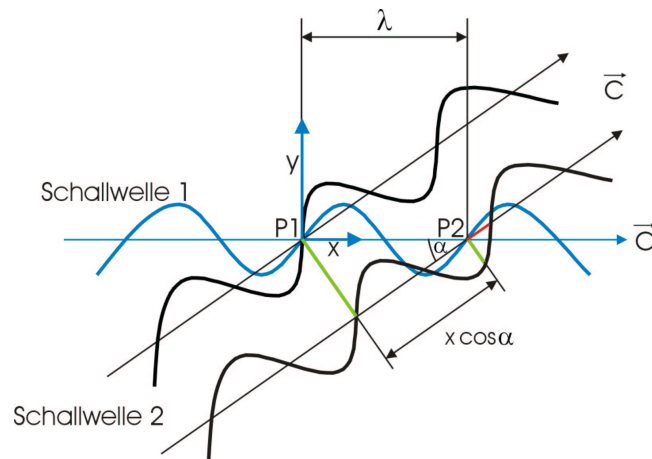


Abb. 6.17: Schräger Schalleinfall aus Winkel α

Außerdem sei erwähnt, dass die Wellen in obiger Abbildung nur wegen der leichteren grafischen Darstellung als Transversalwelle gezeichnet wurden, diese jedoch tatsächlich eine Longitudinalwelle ist. Weiters wird für folgende Betrachtungen vorausgesetzt, dass es sich bei der Schallausbreitung um ebene Wellen handelt ($\overline{P_1P_2} \ll$ Abstand Lautsprecher-Mikrofon).

Wie in Abb. 6.17 ersichtlich, findet in Punkt P_1 eine Überlagerung zweier Wellen unterschiedlicher Ausbreitungsrichtung statt. Dabei sei die Phasenverschiebung in P_1 als null angenommen, womit also in diesem Punkt konstruktive Interferenz herrscht. Bewegt man nun das Mikrofon zu P_2 , so sieht man eine Phasenverschiebung der beiden Wellen, welche der roten Strecke entspricht. Den Maximalwert des Gesamtschalldrucks erhält man, indem man die beiden Wellen, wie in Abb. 6.17 dargestellt, mit entsprechender Phasenverschiebung überlagert. Diese Überlegungen werden nun für sämtliche einfallende Schallwellen übertragen.

Definiert man ein Koordinatensystem in P_1 , ergibt sich die Abhängigkeit der Phasenverschiebung $\varphi_{v,1}(x)$ sämtlicher einfallender Schallwellen, in Bezug auf Welle 1 folgendermaßen:

$$\varphi_{v,1}(x, f_0) = \frac{x}{\lambda_0} (\cos \alpha - 1) 2\pi + \varphi_{v,1,0} \quad (6.118)$$

$$\lambda_0 = \frac{c}{f_0}$$

- $\varphi_{v,1}$ Gesamtphasenverschiebung der Welle v auf Welle 1
 $\varphi_{v,1,0}$ Phasenverschiebung der Welle v auf Welle 1 an der Stelle $x=0$

Aus der Phasenverschiebung kann nun für jede Stelle „ x “ die Gesamtüberlagerung aller Schallwellen, $P_{ges}(x)_{\max}$ ermittelt werden.

$$\underline{P}_v(x, f_0) = a_v e^{\varphi_{v,1,0} + \varphi_{v,1}(x)} \quad (6.119)$$

$$P_{ges}(x, f_0)_{\max} = \left| \sum_{v=1}^N \underline{P}_v \right| \quad (6.120)$$

- a_v Amplitude der einfallenden Schallwelle v
 $P_{ges}(x, f_0)_{\max}$ Schalldruckverlauf über x

$P_{ges}(x, f_0)_{\max}$ entspricht nun dem bereits in Abb. 6.6. dargestellten Betrags des Frequenzzeigers für die jeweilige Frequenz.

Diese Überlegungen werden nun in Matlab in einem Beispiel implementiert.

Signalfrequenz: 5460 Hz
 $\lambda = 62,8$ mm

Nun wurden in Matlab 20 verschiedene Signale dieser Frequenz mit zufälligen Einfallswinkeln im Raum ($0 - 180^\circ$) und verschiedenen Amplituden ($0,5 - 3,5$) überlagert. Ein derart synthetischer Signalausschnitt wird in folgendem Bild dargestellt, welcher dem Betragfrequenzgang in Abhängigkeit von der Ortskoordinate x für die Frequenz 5460 Hz entspricht.

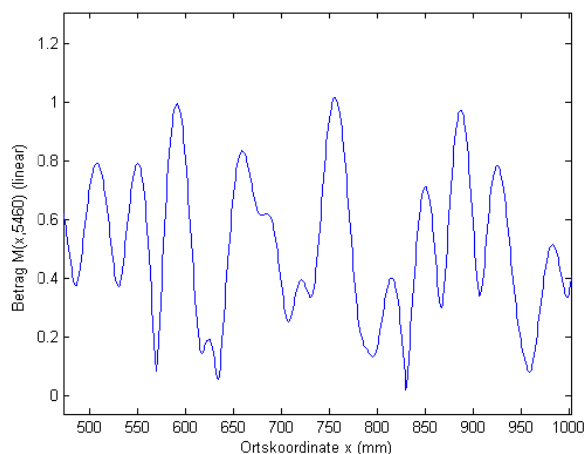


Abb. 6.18: Betrag des Frequenzbins für 5460 Hz über den Ort

Abb. 6.18 stellt nun den Spaltenvektor der Matrix $\underline{M}(x, 5460 \text{ Hz})$ dar. Durch Fourier-Transformation erhält man nun den Betragsortsfrequenzgang des aus 20 Einzelschwingungen bestehenden Gesamtsignals. Auch hier ist nun die Bandbegrenzung deutlich sichtbar, welche bereits in Abb. 6.14 zu sehen war. Es gilt wieder die Beziehung aus (6.16).

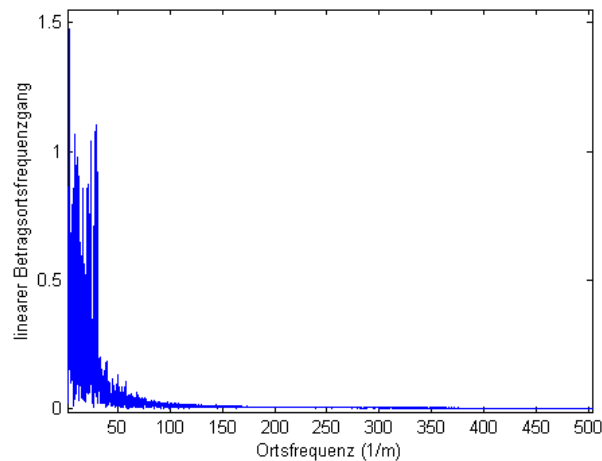


Abb. 6.19a: Ortsfrequenzgang des Signals aus Abb. 6.18

Der Betragsortsfrequenzgang wird noch in ein Leistungsdichtespektrum übergeführt, wobei das örtliche Fenster mit 1 m, die Überlappung mit 90 % realisiert wurde. Das geschätzte Leistungsdichtespektrum wurde anhand von 89 Einzelrealisationen ermittelt, welche durch Hanning-Fenster gewichtet waren.

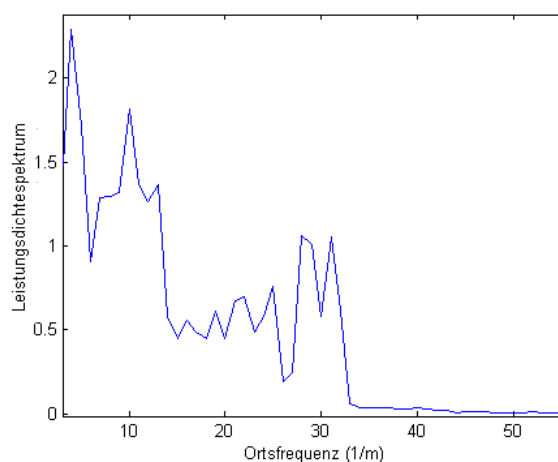


Abb. 6.19b: Geschätztes (Orts-)Leistungsdichtespektrum

An der spektralen Zusammensetzung des Betragsortsfrequenzganges findet auch durch die Mittelung mittels geschätzte Leistungsdichtespektrum keine Veränderung statt.

Bei den bisherigen Überlegungen in Punkt 6.5, wurde nicht ein Hindernis im Raum bewegt wie bei der Arraymessung, wo die Bandbegrenzung nach Gl. (6.16) beobachtet wurde, sondern das Mikrofon selbst bewegte sich durch ein statisches (synthetisch simuliert) Schallfeld.

Es wird nun anhand der Messdaten aus 4.4.2.2 die Ortsfrequenz-, Frequenzdarstellung gezeigt, dass (6.16) sowohl für bewegte Objekte im Raum (nicht stationärer Fall) als auch für ein bewegtes Mikrofon im konstanten Schallfeld (stationärer Fall) gilt.

Anhand Abb. 6.20 sieht man, dass (6.16) auch für ein örtlich abgetastetes Schallfeld gilt.

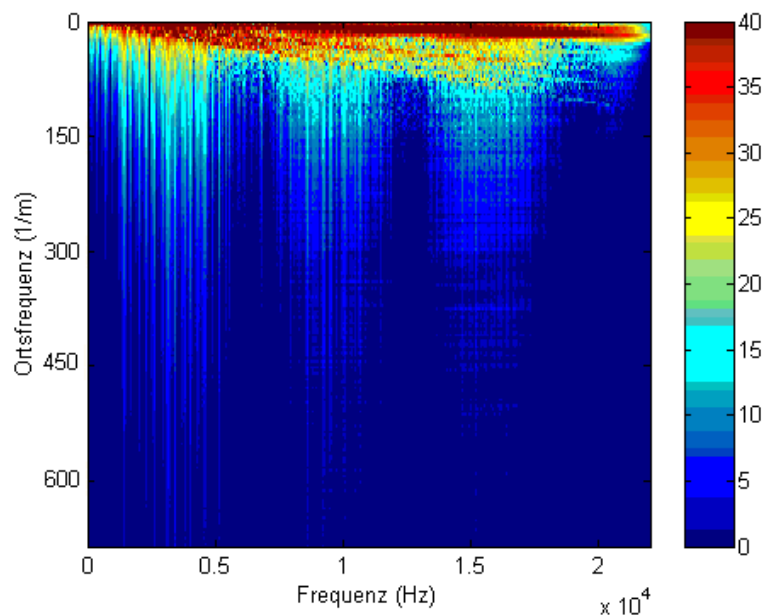


Abb. 6.20: Ortsfrequenz-, Frequenzdarstellung bei bewegtem Mikrofon im statischen Schallfeld

Es wurde somit anhand der deterministischen Überlagerung vieler Sinusschwingungen gleicher Frequenzen gezeigt, dass der Betragsortsfrequenzgang anhand der Beziehung von (6.16) bandbegrenzt wird.

Die Idee der linearen Prädiktion ist folgende:

Es sei ein Signal $s(n)$ mit stationären Abschnitten angenommen. Innerhalb dieses Bereichs ist die Zusammensetzung des Spektrums also konstant, genauso wie die Autokorrelation. Mittels linearer Prädiktion wird nun versucht, den aktuellen Abtastwert des Signals $s(n)$ aus den letzten p -Abtastwerten ($s(n-1)$ bis $s(n-p)$) vorauszusagen.

Gegeben sei die allgemeine Zustandsgleichung eines rekursiven Digitalfilters:

$$y(n) = \sum_{k=0}^{k=L} b_k x(n-k) + \sum_{k=1}^N a_k y(n-k) \quad (6.121)$$

Bei der üblichen Anwendung stellt dabei $x(n)$ das Eingangs- und $y(n)$ das Ausgangssignal dar.

Nun wird (6.21) in die übliche Notation eines Prädiktors übergeführt.

$$s(n) = \sum_{k=1}^p b_k s(n-k) + Gu(n) \quad (6.122)$$

$s(n)$ wird also durch seine letzten Abtastwerte plus einem Eingangssignal $Gu(n)$ berechnet.

$\sum_{k=1}^p b_k s(n-k)$ stellt nun den deterministischen Anteil von $s(n)$ dar.

$Gu(n)$ stellt den Anteil dar, welcher nicht aus der Vergangenheit linear extrapoliert werden kann.

Die zu (6.22) gehörige Übertragungsfunktion wird wie folgt angegeben:

$$H(z) = \frac{S(z)}{U(z)} = \frac{G}{1 - \sum_{k=1}^p b_k z^{-k}} \quad (6.123)$$

$H(z)$ hat also die Struktur eines IIR-Filters.

Nun sei angenommen, dass $Gu(n)$ in Gl. (6.22) vernachlässigbar klein ist. $s(n)$ in (6.22) wird nun durch $s_p(n)$ angenähert.

$$s_p(n) = \sum_{k=1}^p \alpha_k s(n-k) \quad (6.124)$$

(6.24) kann also folgend interpretiert werden.

Sind die letzten Ausgangswerte eines Signals bekannt, so kann daraus der neue aktuelle Ausgangswert $s_p(n)$ vorhergesagt werden. Gleich $s_p(n)$ dem ursprünglichen Signal $s(n)$ für alle n , so ist dies der Idealfall des Prädiktors. Im Allgemeinen wird sich jedoch ein Prädiktionsfehler $e(n)$ ergeben.

$$e(n) = s(n) - s_p(n) = s(n) - \sum_{k=1}^p \alpha_k s(n-k) \quad (6.125)$$

Damit nun die Prädiktion ein gutes Ergebnis liefert, muss also $e(n)$ klein sein. Aus mehreren Möglichkeiten $e(n)$ zu minimieren und daraus die Prädiktionskoeffizienten zu erhalten, wird die Autokorrelationsmethode ausgewählt, welche auch als stationärer Ansatz bezeichnet wird [Lit. 15, 16].

Dabei wird von der Minimierung des mittleren quadratischen Fehlers ausgegangen.

$$E_n = \sum_m e_n^2(m), \quad E \text{ à Minimum} \quad (6.126)$$

Dabei bezeichnet E_n den Fehler im Bereich um n .

$$\begin{aligned} &= \sum_m [(s_n(m) - s_{p,n}(m))]^2 \\ &= \sum_m \left[s_n(m) - \sum_{k=1}^p \alpha_k s_n(m-k) \right]^2 \end{aligned}$$

Nun werden die partiellen Ableitungen 0 gesetzt.

$$\frac{\partial E}{\partial \alpha_i} = 0 \quad i = 1 \dots p$$

Man erhält daraus p Gleichungen, um p Unbekannte α_k zu erhalten.

Nach weiteren Umformungen gelangt man zu folgendem Gleichungssystem:

$$\begin{bmatrix} R(0) & R(1) & \dots & R(p-1) \\ R(1) & R(0) & \dots & R(p-2) \\ \vdots & \ddots & \ddots & \vdots \\ R(p-1) & R(0) & & \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \vdots \\ \alpha_p \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(p) \end{bmatrix} \quad (6.127)$$

$$R \cdot \vec{\alpha} = \vec{R} \quad (6.128)$$

Dieses Gleichungssystem ist auch unter „Yule-Walker-Gleichungssystem“ bekannt. Die Matrix R besitzt Töplitzstruktur.

$$R_n(l) = \sum_{m=0}^{L-1-l} s_n(m) s_n(m+l) \quad (6.129)$$

L bezeichnet dabei die zur Autokorrelation verwendete Länge der Signalausschnitte von s .

(6.27) kann nun über den "Levinson-Durbin Algorithmus" gelöst werden [Lit. 15, 16].

Mit dem bis jetzt gezeigten Prädiktionsansatz kann aus den letzten aktuellen Ausgangswerten eine Vorhersage um einen Zeitschnitt in die Zukunft erfolgen.

Soll nun weiter in die Zukunft prädiziert werden, so wird (6.24) wie folgt erweitert:

$$s_p(n+D) = \sum_{k=0}^{p-1} \alpha_k s(n-k) \quad (6.130)$$

Für $D=1$ entspricht (6.30) nun wieder (6.24).

(6.30) kann auch als Gleichungssystem angeschrieben werden:

Beispiel: $D=5$ $p=6$ $L=20$

$$s(1) = \alpha_0 s(-4) + \alpha_1 s(-5) + \dots + \alpha_{p-2} s(-8) + \alpha_{p-1} s(-9)$$

$$s(2) = \alpha_0 s(-3) + \alpha_1 s(-4) + \dots + \alpha_{p-2} s(-7) + \alpha_{p-1} s(-8)$$

$$s(L-1) = \alpha_0 s(L-1-D) + \alpha_1 s(L-1-D-1) + \dots + \alpha_{p-2} s(L-1-D-(p-2)) + \alpha_{p-1} s(L-1-D-(p-1))$$

$$s(L) = \alpha_0 s(L-D) + \alpha_1 s(L-D-1) + \dots + \alpha_{p-2} s(L-D-(p-2)) + \alpha_{p-1} s(L-D-(p-1))$$

(6.131)

Es gibt nun 3 Parameter, die in (6.30) eingehen:

- D Prädiktionshorizont
- p Ordnung des Prädiktors
- L Lernbereich des Prädiktors

Der Lernbereich des Prädiktors entspricht nun der Anzahl der Gleichungen in (6.31). Es bestehen nun folgende Zusammenhänge der Parameter D , p und L .

Ist $L=p$, so gibt es eine eindeutige Lösung für Gleichungssystem (6.31), wobei dies nicht dem üblichen Prädiktorbetrieb entspricht.

Im Normalfall ist $L > p$, wodurch (6.31) überbestimmt wird und es keine exakte Lösung des Gleichungssystems gibt. In diesem Fall bestimmt man die Lösung nach dem Prinzip des kleinsten Fehlerquadrates.

L bezeichnet nun jenen Bereich, innerhalb dessen die Prädiktionskoeffizienten Gültigkeit haben sollen.

Wird L größer als der stationäre Zeitabschnitt eines Signals ist, so steigt der Prädiktionsfehler, da die Prädiktionskoeffizienten nicht für den gesamten Bereich L optimal sind.

Es soll untersucht werden, ob, in welchem Horizont, und mit welcher Qualität sich die Raumimpulsantwort mittels linearer Prädiktion vorhersagen lässt. Die Prädiktion soll dabei in der Darstellung $M(w, n)$ (örtliche Prädiktion der zeitlichen Abtastwerte) und in der Darstellung $\underline{M}(w, k)$ (örtliche Prädiktion der Frequenzbins) erfolgen.

Wie in der Vier-Quadranten-Darstellung bzw. der Zeigerdarstellung verschiedener Frequenz- bzw. Zeitwerte ersichtlich ist, dürfte die Prädiktion im Frequenzbereich bei tiefen Frequenzbins am besten funktionieren. Es wurden dabei die Messdaten aus 4.4.2.1 verwendet. Diese Messdaten werden deshalb verwendet, da den Zeit- bzw. Frequenzbins bei dieser Messung ein geringeres Störsignal überlagert ist, als bei der Arraymessung, sprich der zeitliche Verlauf der Zeit- bzw. Frequenzbins der Messdaten aus 4.4.2.1 ist glatter. Die Matrix M besitzt bei diesen Messdaten 750 Zeilen.

Hier wird nun versucht, die einzelnen Frequenzzeiger aus den Abbildungen 6.6, vorherzusagen.

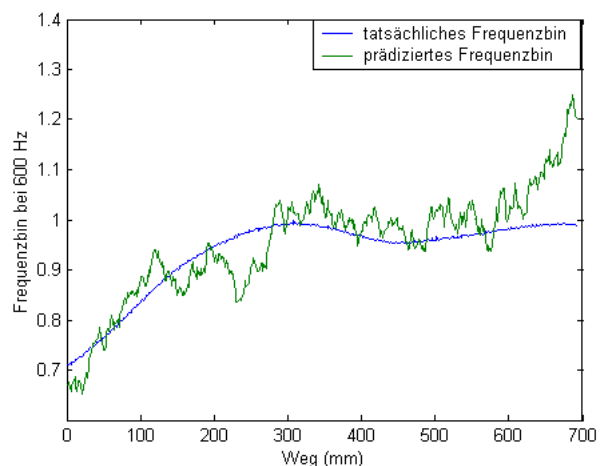
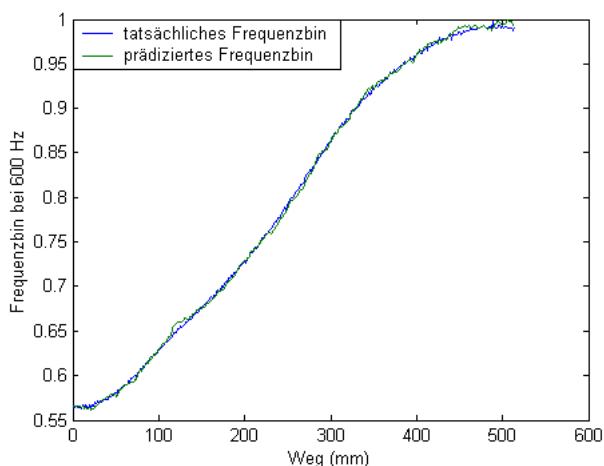


Abb. 6.21a: Prädiktion bei $D = 28$ mm

Abb. 6.21b: Prädiktion bei $D = 208$ mm

Abb. 6.21a. zeigt 2 Frequenzbins bei 600 Hz, welche unter verschiedenem Prädiktorhorizont ermittelt wurden. Der Prädiktorhorizont D von Abb. 6.21a beträgt dabei 20 Abtastwerte (28 mm), in Abb. 6.21b beträgt dieser 150 Abtastwerte (208 mm). (Das Abtastintervall des Ortes beträgt bei diesem Datensatz 1,39 mm.)

Der Lernbereich L umfasst bei beiden Abbildungen den Bereich bis zur Ortskoordinate $x=485$ mm. Darüber hinaus befindet sich das prädizierte Signal. Bezieht man jeweils den Prädiktionshorizont auf die Wellenlänge des Frequenzbins, so erhält man den normierten Prädiktorhorizont d_N .

$$d_N = \frac{D}{\lambda}$$

Abb. 6.21a wurde bei $d_N=0,048$ ermittelt, Abb. 6.21b. bei $d_N=0,36$.

Etwas verwirrend in Abb. 6.21a und 6.21b. ist, dass beide Bilder nicht denselben Ausschnitt des Signals zeigen. Der Grund liegt in der unterschiedlichen Länge des Prädiktorhorizontes D .

Will man nun den Prädiktorfehler in Abhängigkeit des Prädiktorhorizonts abschätzen, sollte jedoch immer der gleiche Prädiktionsbereich als Basis der Beurteilung dienen, da nicht davon ausgegangen werden kann, dass örtlich unterschiedliche Signalbereiche jeweils mit gleicher Qualität prädiziert werden können. Soll also ein Signal mit verschiedenen Parametern (D , p , L) vorhergesagt werden, um optimale Werte für den Prädiktor zu finden, müssen sich die prädizierten Signale immer auf denselben Ortsbereich beziehen.

Anhand der nächsten Abbildung wird dies verdeutlicht.

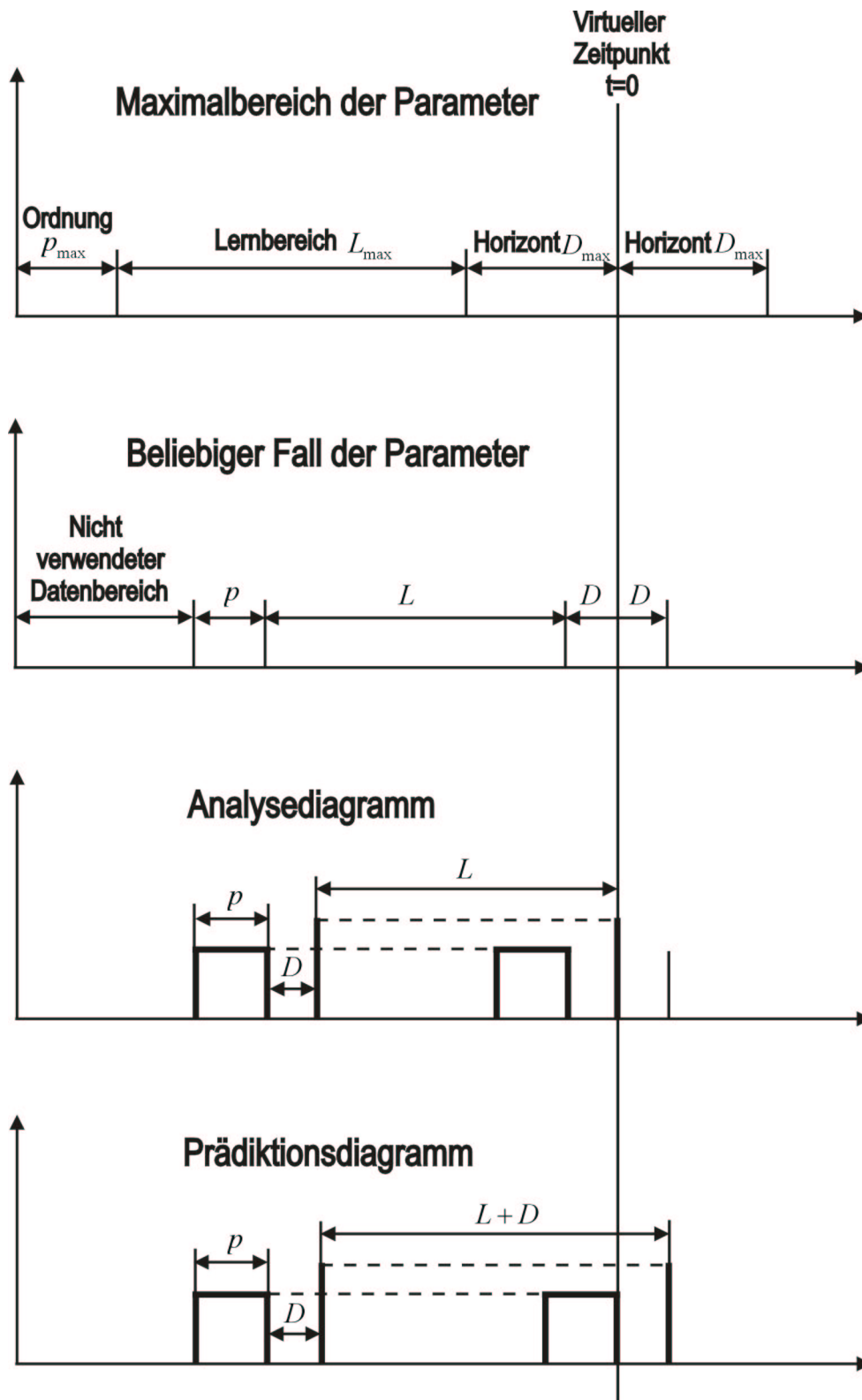


Abb. 6.22: Prädiktionsdiagramm

Wie bereits erwähnt, liegen die Frequenzbins $\underline{M}(w,k)$ vor, wobei für w 750 Abtastwerte vorliegen, was ca. 1,04 Metern entspricht. Es wird nun davon ausgegangen, dass die Matrix $\underline{M}(w,k)$ nur bis $w=600$ bekannt ist. Diese Abtastwerte liegen nun im negativen Bereich der Zeitachse in Abb. 6.22.

Somit können zur Berechnung der Prädiktionskoeffizienten auch nur max. 600 Abtastwerte verwendet werden. Die aus diesem Bereich resultierenden Prädiktorkoeffizienten werden nun verwendet, um die Frequenzbins von $w=601$ bis $w=750$ zu berechnen, welche im positiven Bereich der Zeitachse (scheinbar in der Zukunft) liegen. Diese werden dann mit dem ursprünglichen Signal von $w=601$ bis $w=750$ verglichen. Der maximale Prädiktionshorizont entspricht somit 150 Sample bzw. 208 mm.

Im obersten Diagramm sind nun die maximalen Werte der Parameter (D, p, L) zu sehen. Nur für diesen Fall werden alle 750 Werte der Frequenzbins zur Berechnung benötigt.

Der allgemeine Fall ist im zweiten Diagramm dargestellt, wo zu Beginn ein nicht verwendeter Datenbereich des Signals auftritt, damit der prädizierte Bereich genau bei $w=601$ beginnt.

Das nächste Diagramm zeigt nun die, zur Ermittlung der Prädiktionskoeffizienten verwendeten Datenblöcke, welche das Gleichungssystem aus (6.31) ergeben.

Das letzte Diagramm erklärt die Berechnung des prädizierten Signals im Zukunftsbereich. Durch Multiplikation des Datenblockes der Länge p mit den Prädiktorkoeffizienten, ergibt sich ein Abtastwert y_p , welches sich im positiven Bereich der Zeitachse befindet.

Im folgenden sollen nun die Parameter (D, p, L) so bestimmt werden, dass die Prädiktion gute Ergebnisse erzielt. Als Bewertung des Prädiktors wird dabei das SDR des ursprünglichen zum prädizierten Signal herangezogen.

Es zeigt sich jedoch, dass aus der Prädiktion eines einzelnen Frequenzzeigers keine Aussage über die Qualität des Prädiktors getätigt werden kann, da die SDR-Werte in Abhängigkeit vom Prädiktionshorizont D stark variieren (benachbarte Frequenzzeiger haben beispielsweise im Bereich von $D=20$ bis 30 einen unterschiedlichen Prädiktionsgewinn). Aus diesem Grund, wird eine Mittelung für den Prädiktionsgewinn über jeweils 50 Frequenzbins (Auflösung 2 Hz) vorgenommen, um eine globale Aussage bezüglich des Prädiktionsgewinns in einem Frequenzbereich (z.B. 1000 – 1100 Hz) tätigen zu können. Die Mittelung darf jedoch nicht anhand der SDR-Werte durchgeführt werden, sondern passiert folgendermaßen:

$$SDR = 10 \log_{10} \frac{\sum_{k=1}^{50} \sum_{n=D_1}^{D_2} \{y_k[n]\}^2}{\sum_{k=1}^{50} \sum_{n=D_1}^{D_2} \{y_k[n] - y_{p,k}[n]\}^2} \quad [dB] \quad (6.132)$$

k Parameter für Frequenzbin

D_1, D_2 Grenzen innerhalb derer, das SDR von y und y_p berechnet wird

In der nächsten Tabelle beträgt beispielsweise für die 7. Spalte $D_1=30$ und $D_2=40$.

Ordnung des Prädiktors p	Prädiktorhorizont in Sample												
	5	10	15	20	25	30	40	50	65	80	100	120	150
2	42,8	37,2	33,5	31,0	29,0	27,2	24,5	22,3	19,7	17,7	15,7	14,1	12,5
4	45,2	40,0	36,5	34,0	32,0	30,2	27,3	24,8	21,7	19,2	16,8	15,0	13,1
6	47,8	43,1	39,9	37,4	35,4	33,4	30,3	27,5	23,8	20,7	17,7	15,5	13,4
8	49,8	45,4	42,5	40,4	38,2	36,0	32,3	29,2	25,1	21,6	18,2	15,8	13,5
10	51,3	47,1	44,2	42,2	39,9	37,6	33,4	30,1	25,7	22,0	18,3	15,9	13,5
16	52,9	48,7	45,7	43,1	40,7	38,0	33,9	30,4	25,8	22,1	18,3	15,8	13,5
30	53,4	49,5	47,1	45,0	42,8	40,2	35,7	32,1	27,3	23,3	19,3	16,7	14,1
50	53,4	49,5	47,9	46,8	44,8	42,4	38,6	34,6	29,9	25,4	21,3	18,0	14,6
70	53,1	49,6	48,0	46,8	44,7	42,3	38,1	33,9	29,2	25,1	21,2	17,7	14,2
100	52,7	49,1	47,5	46,1	44,0	41,4	37,4	33,4	29,2	25,2	21,1	17,3	13,6

Tab. 6.16: Prädiktionsergebnisse im Bereich von 400 – 500 Hz

Tab. 6.6. zeigt nun die Qualität der Prädiktion in Abhängigkeit von der Ordnung und des Horizonts des Prädiktors. Der Lernbereich beträgt dabei 350 Abtastwerte, der betrachtete Frequenzbereich geht von 400 – 500 Hz. Es wurde für jeden SDR-Wert der Tabelle Gleichung (6.32) angewandt.

Es zeigt sich erwartungsgemäß, dass die Prädiktion umso besser funktioniert, je kleiner der Prädiktorhorizont ist. Bei der Prädiktorordnung kann man $p \approx 30$ als Anhaltspunkt für eine Sättigung des bei weiteren Ordnungssteigerung erzielten Prädiktionsgewinns betrachten.

Im folgenden werden nun die Prädiktionsergebnisse für Frequenzbins bis 2500 Hz grafisch dargestellt, wobei der Prädiktorhorizont in mm angegeben ist. Hier ist wichtig zu erwähnen, dass der Farbbalken der Abbildungen 6.23 nur den Dynamikbereich von 20 bis 45 dB umfasst. Ein Prädiktionsgewinn von <20 dB wird demnach mit dem Farbwert der unteren grafischen Darstellungsgrenze (blau) dargestellt. Der Grund dieser Maßnahme liegt darin, dass der hauptsächlich interessierende Bereich (Prädiktionsgewinn zumindest 20 dB) detaillierter dargestellt werden kann.

Die Auswertung wurde dabei für 3 verschiedene Lernbereiche von $L=100, 200, 350$ Abtastwerte angewandt.

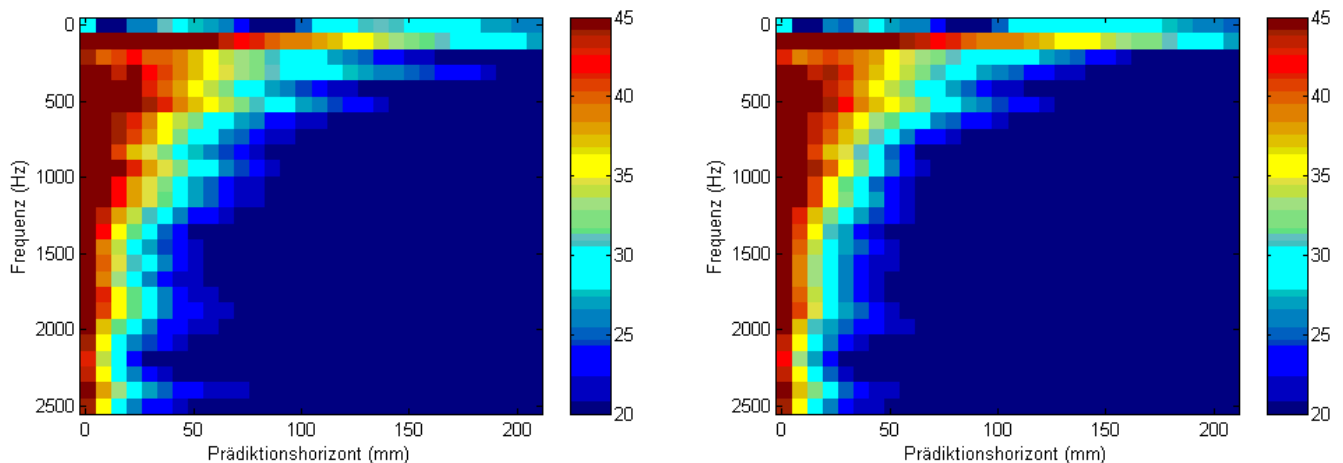


Abb. 6.23a: absoluter Prädiktionsgewinn
Prädiktionsgewinn bei $L=350$ und $p=30$
 $p=30$

Abb. 6.23b: absoluter
bei $L=200$ und

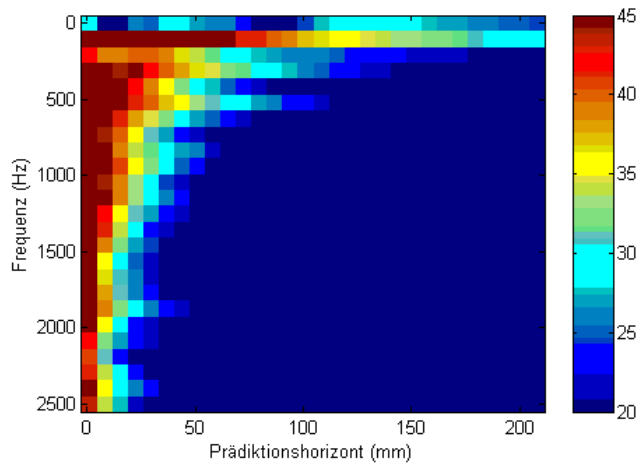


Abb. 6.23b: absoluter Prädiktionsgewinn bei $L=100$ und $p=30$

$w=600$. Da wir den zukünftigen Verlauf der verschiedenen Frequenzbins nicht wissen, nehmen wir also den letzten Wert an der Stelle $w=600$ als den in der Zukunft gültigen, an, sprich die Frequenzbins werden im Bereich $w=601$ bis 750 als konstant extrapoliert. Da wir in Wirklichkeit jedoch den tatsächlichen Verlauf kennen, kann das SDR aus den tatsächlichen, zu den als konstant angenommenen Frequenzbins, berechnet werden. Dieses SDR wird nun, von den in Abb. 6.23 dargestellten SDR-Werten, subtrahiert. Das Ergebnis bezeichnet nun den zusätzlichen Prädiktionsgewinn. Wird der zusätzliche Prädiktionsgewinn negativ (aufgrund der eingeschränkten Dynamik entspricht das der tiefblauen Farbe entsprechend 0 dB), ist es somit besser, die Frequenzbins konstant zu lassen, also den Prädiktor p-ter Ordnung durch ein einfaches Halteglied zu ersetzen.



Abb. 6.24a: zusätzlicher Prädiktionsgewinn bei $L=200$ und $p=30$ und $p=30$

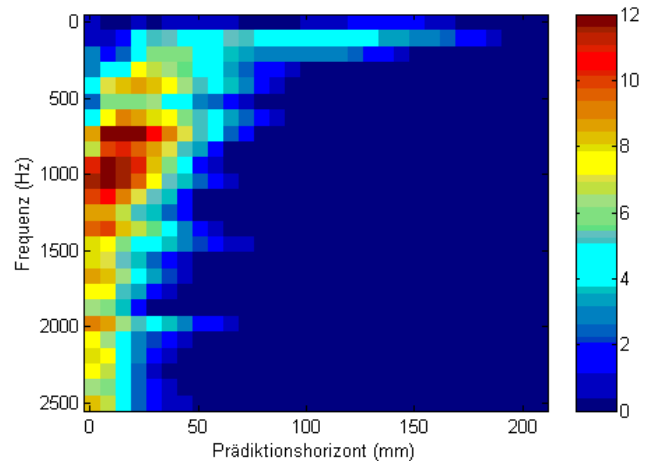


Abb. 6.24b: zusätzlicher Prädiktionsgewinn bei $L=350$ und $p=30$

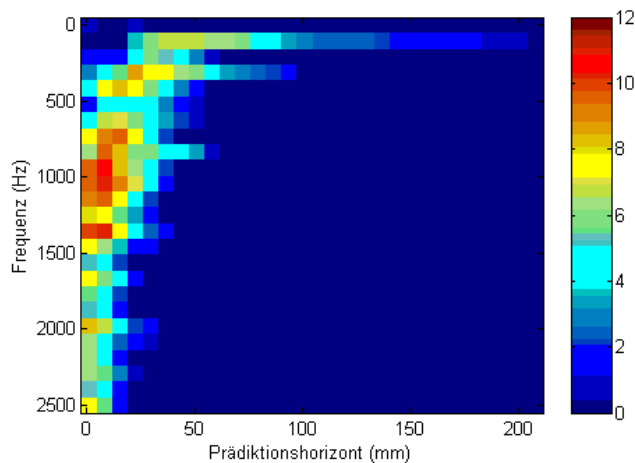


Abb. 6.24c: zusätzlicher Prädiktionsgewinn bei $L=100$ und $p=30$

In den Abbildungen 6.24 ergibt sich also ein maximaler zusätzlicher Prädiktionsgewinn von ca. 12 dB, jedoch nur in einem sehr kleinen Horizont. Auch hier zeigt sich das Verhalten, dass bei hohen Frequenzen weniger weit prädiziert werden kann.

1.30.1 Prädiktion im Zeitbereich

Im Zeitbereich wurden jeweils 150 Bins analog (6.22) gemittelt, wobei diese aus dem Zeitbereich um 1000, 6000, 11000, 16000 und 21000 stammen. Der oberste Querbalken in Abb. 6.25 zeigt beispielweise den Prädiktionsgewinn im Zeitbereich von 926 – 1075 Abtastwerten.

Folgende Abbildung zeigt, dass die Prädiktion dabei viel schlechter als im analysierten Frequenzbereich funktioniert.

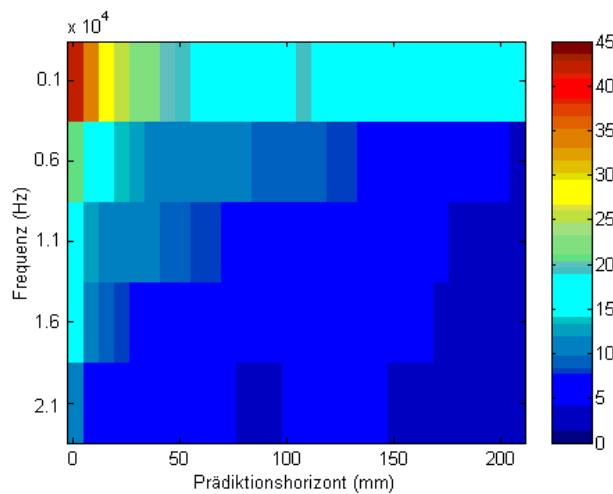


Abb. 6.25: Prädiktionsergebnis im Zeitbereich

Abschließend kann für die Prädiktion folgendes zusammengefasst werden:

Die Prädiktion funktioniert am besten bei der Vorhersage der Frequenzbins für den tiefen Frequenzbereich. Hier ist der erreichbare Prädiktionshorizont absolut gesehen am größten.

Generell kann die Aussage gemacht werden, dass in einem Prädiktionshorizont von ca. 5 – 10 % der Wellenlänge brauchbare Prädiktionsergebnisse erwartet werden können, was bei hohen Frequenzen zu einem absoluten Prädiktionshorizont von wenigen Millimetern führt (siehe Abb. 6.3, 6.4). Für die Echounterdrückung ist dieser Horizont jedoch viel zu gering.

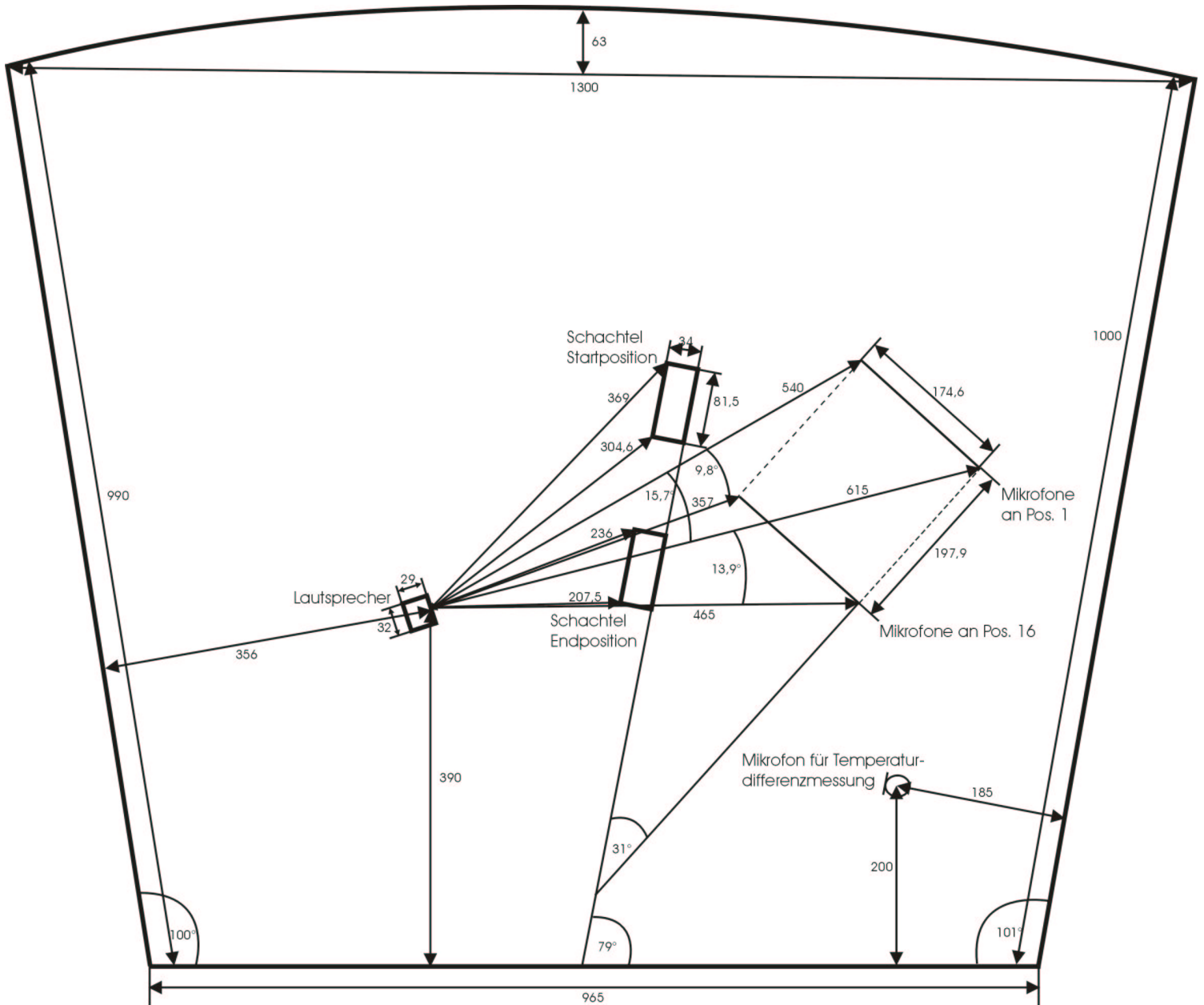
Der hohe Prädiktionserfolg des einfachen Haltegliedes, rechtfertigt den auch schon bisher sehr erfolgreichen Einsatz rekursiver Adaptionsalgorithmen (LMS, RLS) in der Echounterdrückung. Der darüber hinaus gehende Einsatz von linearen Prädiktoren für die zeitliche Änderung der Koeffizienten des adaptiven Kompensators dürfte nur im tiefen Frequenzbereich etwas bringen, sollte dort aber noch näher untersucht werden (z.B. Einsatz nicht linearer Prädiktion).

Weiters kommen im realen Anwendungsfall noch folgende Erschwernisse hinzu:

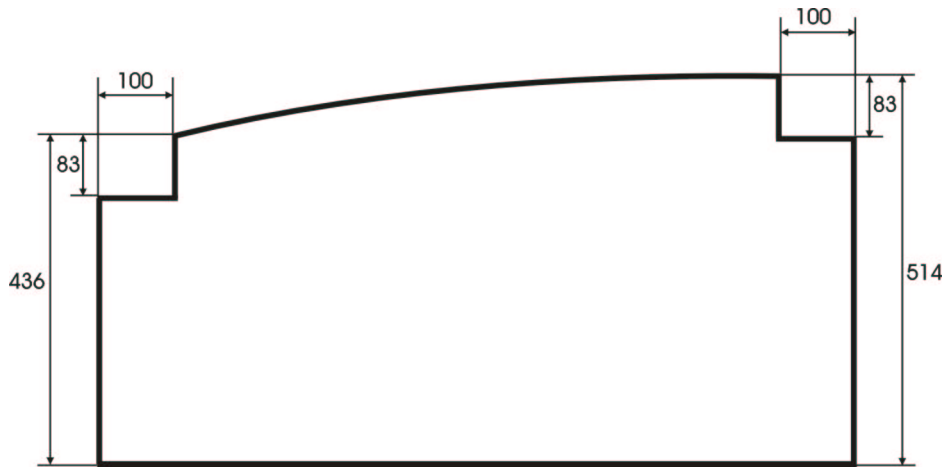
- Geschwindigkeit des bewegten Objektes ist nicht konstant
- Geschwindigkeit des bewegten Objektes ist viel größer als die Geschwindigkeit der bewegten Schachtel
- Für die Messung der RIR stehen i.a. keine Studiogeräte zur Verfügung

Anhang A

A1 Grundriss CUBE



A2 Aufriss CUBE

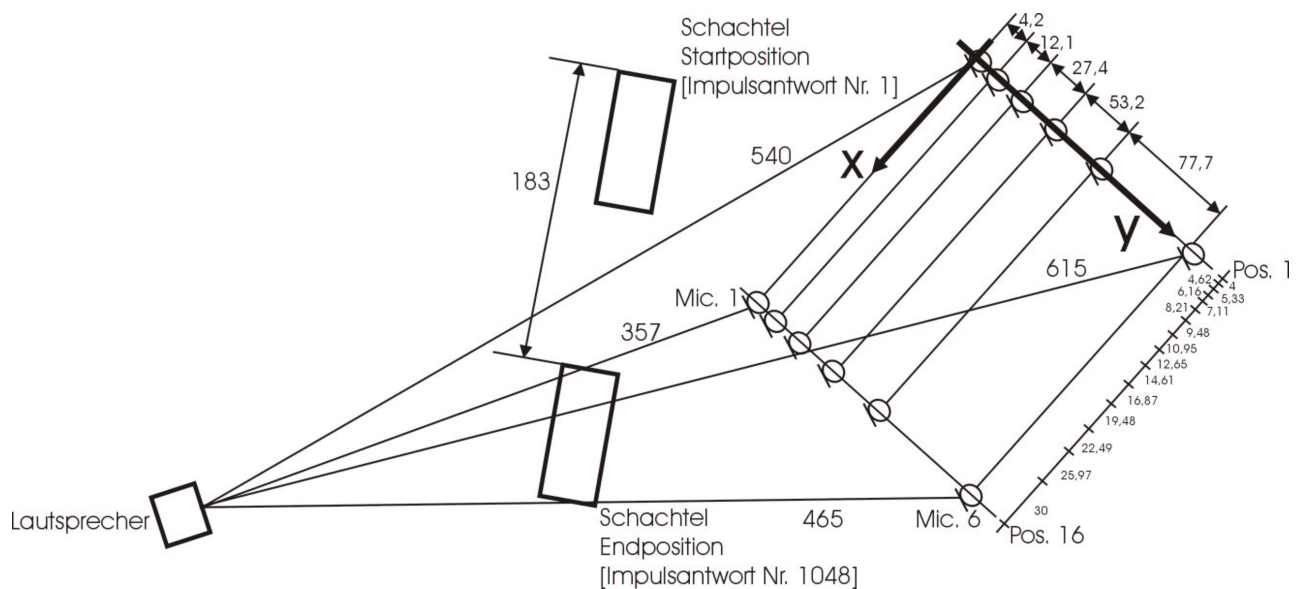


Maße in cm, Maßstab 1:100

Abb. A.2: Aufriss CUBE

Die mittlere Höhe des Cubes wird mit $\frac{4,36 + 5,14}{2} = 4,75 \text{ m}$ abgeschätzt.

A3 Detailaufstellung der Arraymessung



Maße in cm

Mikrofonabstände nicht maßstäblichgezeichnet

Abb. A.3: Detailaufstellung der Arraymessung

A4 Höhe der Mikrofone

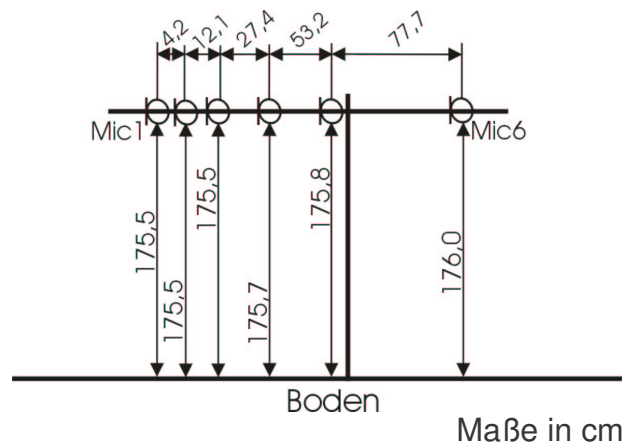


Abb. A.4: Höhe der Arraymikrofone

Anhand dieser Abbildung wird die Höhe der Arraymikrofone verdeutlicht.

A5 Abmessung der Lautsprecheraufstellung

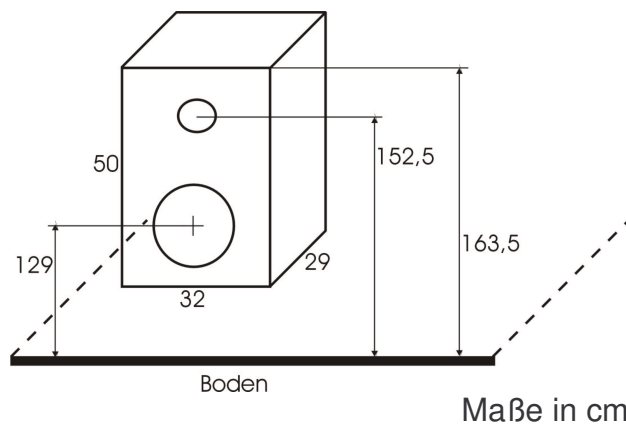


Abb. A.5: Aufstellung Lautsprecher

Der Lautsprecher steht bei der Arraymessung in angegebener Höhe auf einem Ständer. Das dazugehörige Datenblatt ist in Anhang C2 zu sehen.

A6 Abmessungen des bewegten Hindernisses

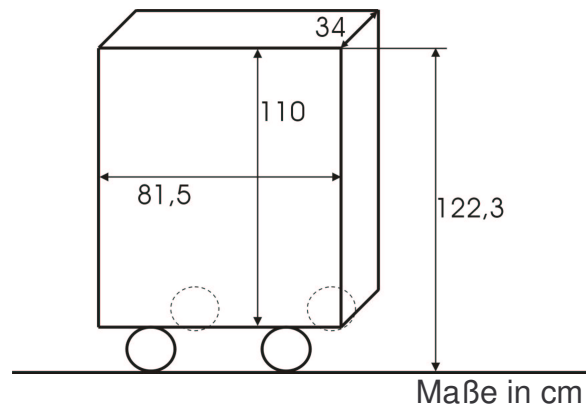


Abb. A.6: Abmessungen des bewegten Hindernisses (Schachtel)

Als Hindernis wurde eine leere Kartonschachtel verwendet.

A7 Fotos der Messungen im CUBE



Abb. A.7: Messschlitten 2 mit fahrbarem Array

Diese Abbildung zeigt das Array. Bei der Konstruktion wurde darauf geachtet, den Messsensor so klein wie möglich zu halten, damit dieser selbst (an verschiedenen Positionen) den Raum so wenig wie möglich beeinflusst.

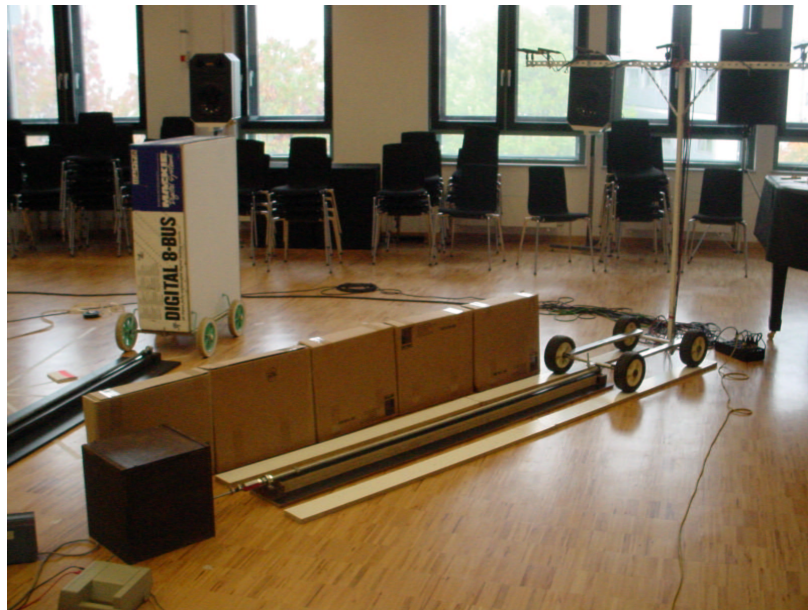


Abb. A.8: Originalaufstellung bei der Messung

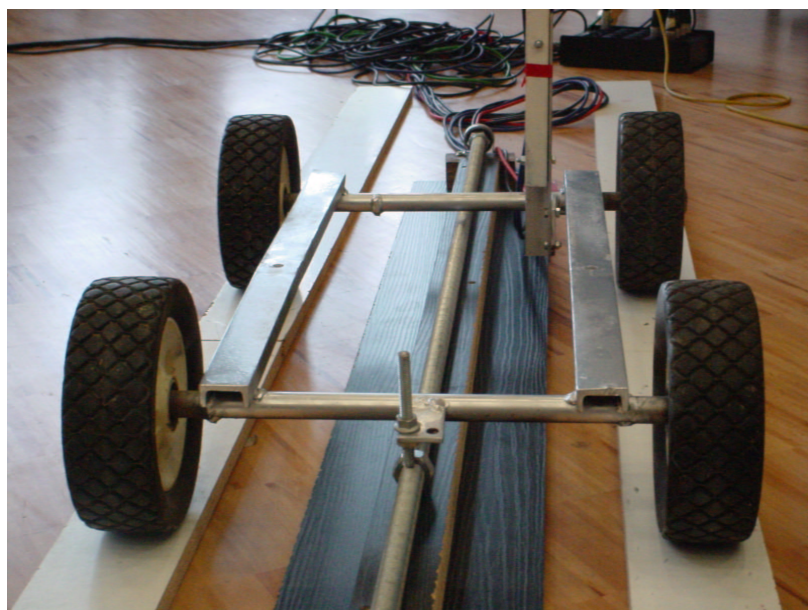


Abb. A.9: Grossaufnahme Messschlitten 2 mit fahrbarem Untersatz des Arrays

In Abb. A.9 zeigt den Antrieb des Arrayuntersatzes über den als Spindeltrieb wirkenden Messschlitten 2. Die Gewindemutter zieht dabei den Schlitten mit loser Kopplung. Diese Kopplung hat den Vorteil, dass leichte Exzentriken der Gewindestange nicht auf den fahrbaren Untersatz übertragen werden.

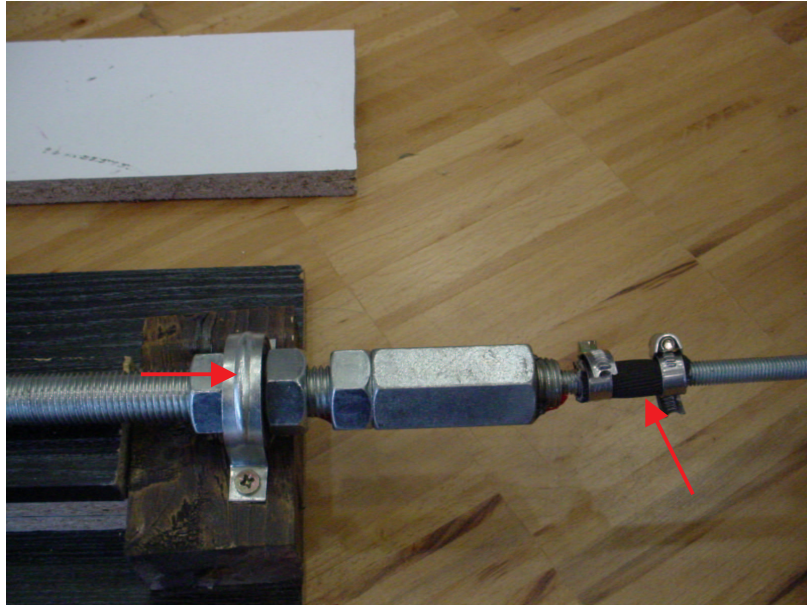


Abb. A.10: Lagerung und Antrieb der Messschlitten

Der linke Pfeil markiert dabei das Lager der Gewindestange. Der rechte Pfeil zeigt auf die Gummikopplung des Messschlittens an den Antriebsmotor.

Diese elastische Kopplung hat 2 Vorteile:

1. Schalldämmung → verringerte Schallweiterleitung des Antriebsmotors an den Spindeltrieb
2. Ausgleichen der Exzentrik des Antriebsmotors

Anhang B

B1 x-Positionen des Arrays

Teilmessung	Abstand zweier Arraypunkte in x-Richtung [cm]	Umdrehungen Messschlitten 2 [U]	x-Position Messschlitten 2 [cm]
Mess1	0,00	0,00	0,00
Mess2	4,00	16,00	4,00
Mess3	4,62	18,48	8,62
Mess4	5,33	21,32	13,95
Mess5	6,16	24,64	20,11
Mess6	7,11	28,44	27,22
Mess7	8,21	32,84	35,43
Mess8	9,48	37,92	44,91
Mess9	10,95	43,80	55,86
Mess10	12,65	50,60	68,51
Mess11	14,61	58,44	83,12
Mess12	16,87	67,48	99,99
Mess13	19,48	77,92	119,47
Mess14	22,49	89,96	141,96
Mess15	25,97	103,88	167,93
Mess16	30,00	120,00	197,93

Tab. B.17: x-Koordinaten der einzelnen Teilarraymessungen

In der 3. Spalte sind die Umdrehungen der Gewindestange von Messschlitten 2 gegeben, um welche das Array von Messung zu Messung weiterbewegt werden musste.

B2 Weg-, Zeitdaten Messschlitten 1

Teilmessung	zurückgelegter Weg - Messschlitten 1 [cm]	Zeit [min]	Durchschnitts- geschwindigkeit [mm/s]
Mess1	195,75	27:44,3	1,176
Mess2	198,50	27:45,7	1,192
Mess3	200,00	27:46,4	1,200
Mess4	200,25	27:42,7	1,204
Mess5	203,00	27:41,3	1,222
Mess6	200,25	27:39,7	1,207
Mess7	197,75	27:43,8	1,189
Mess8	200,00	27:45,8	1,201
Mess9	200,50	27:42,3	1,206
Mess10	201,75	27:39,2	1,216
Mess11	202,00	27:38,6	1,218
Mess12	200,00	27:43,4	1,202
Mess13	201,00	27:41,2	1,210
Mess14	196,50	27:46,3	1,179
Mess15	198,50	27:45,4	1,192
Mess16	198,00	27:43,9	1,190

Tab.B.18: Durchschnittsgeschwindigkeit der einzelnen Teilarraymessungen

B3 Koordinaten sämtlicher Arraysensorpositionen

Position Array- schlitten	Mikrofon 1		Mikrofon 2		Mikrofon 3		Mikrofon 4		Mikrofon 5		Mikrofon 6	
	x	y	x	y	x	y	x	y	x	y	X	y
1	0,0	0,0	0,0	4,2	0,0	16,3	0,0	43,7	0,0	96,9	0,0	174,6
2	4,0	0,0	4,0	4,2	4,0	16,3	4,0	43,7	4,0	96,9	4,0	174,6
3	8,6	0,0	8,6	4,2	8,6	16,3	8,6	43,7	8,6	96,9	8,6	174,6
4	14,0	0,0	14,0	4,2	14,0	16,3	14,0	43,7	14,0	96,9	14,0	174,6
5	20,1	0,0	20,1	4,2	20,1	16,3	20,1	43,7	20,1	96,9	20,1	174,6
6	27,2	0,0	27,2	4,2	27,2	16,3	27,2	43,7	27,2	96,9	27,2	174,6
7	35,4	0,0	35,4	4,2	35,4	16,3	35,4	43,7	35,4	96,9	35,4	174,6
8	44,9	0,0	44,9	4,2	44,9	16,3	44,9	43,7	44,9	96,9	44,9	174,6
9	55,9	0,0	55,9	4,2	55,9	16,3	55,9	43,7	55,9	96,9	55,9	174,6
10	68,5	0,0	68,5	4,2	68,5	16,3	68,5	43,7	68,5	96,9	68,5	174,6
11	83,1	0,0	83,1	4,2	83,1	16,3	83,1	43,7	83,1	96,9	83,1	174,6
12	100,0	0,0	100,0	4,2	100,0	16,3	100,0	43,7	100,0	96,9	100,0	174,6
13	119,5	0,0	119,5	4,2	119,5	16,3	119,5	43,7	119,5	96,9	119,5	174,6
14	142,0	0,0	142,0	4,2	142,0	16,3	142,0	43,7	142,0	96,9	142,0	174,6
15	167,9	0,0	167,9	4,2	167,9	16,3	167,9	43,7	167,9	96,9	167,9	174,6
16	197,9	0,0	197,9	4,2	197,9	16,3	197,9	43,7	197,9	96,9	197,9	174,6

Tab. B.19: Koordinaten der Mikrofone in mm

Koordinatenursprung siehe Detailaufstellung CUBE (Anhang A3)

Anhang C

C1 Frequenzgang Studiolautsprecher Genelec 1032 A, No. 780

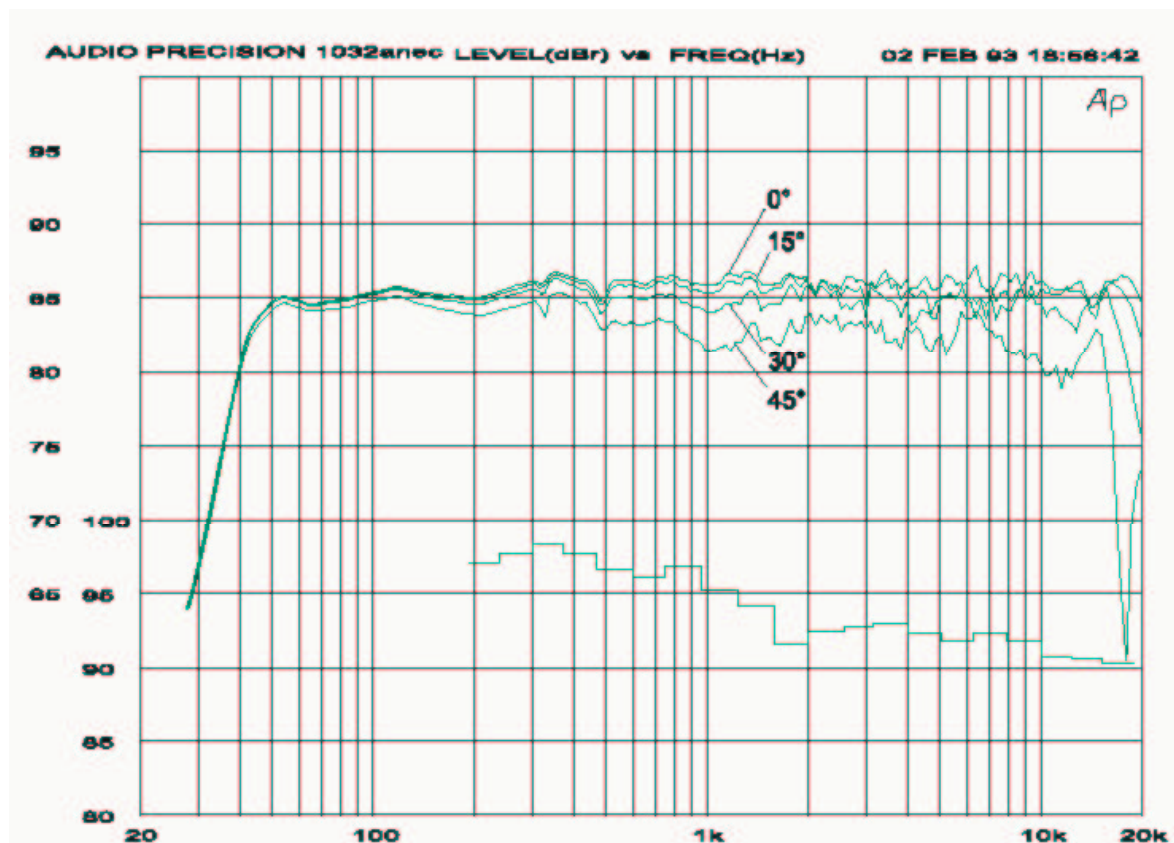


Abb. C.1: Frequenzgang Genelec 1032 A

Die oberen Kurven stellen den Betragsfrequenzgang des Lautsprechers dar. Dabei wird dieser für mehrere Abstrahlwinkel dargestellt. Die untere Kurve stellt die Energieabstrahlung im jeweiligen Terzband dar. Aufgrund der zunehmenden Richtcharakteristik im hochfrequenten Bereich, sinkt daher die Gesamtenergieabstrahlung des Lautsprechers. Die Skalierung wird in dB angegeben.

C2 Datenblatt Genelec

System Specifications	Amplifier Section	Crossover Section
Lower cut-off frequency, -3 dB: ≤ 40 Hz	Bass amplifier output power with an 4 Ohm load: 180 W	Input connector: XLR female Pin 1 gnd Pin 2 + Pin 3 -
Upper cut-off frequency, -3 dB: ≥ 22 kHz		Input impedance: 10 kOhm balanced
Free field frequency response of system: 42 Hz – 21 kHz ($\pm 2,5$ dB)	Treble amplifier output power with an 8 Ohm load: 120 W	Input level for 100 dB SPL output @ 1 m: variable from +6 to -6dBu
Maximum short term sine wave acoustic output on axis in half space, averaged from 100 Hz to 3 kHz: @ 1 m ≥ 113 dB SPL @ 0,5 m ≥ 119 dB SPL	Long term output power is limited by drivers unit protection circuitry. Slew rate: 80 V/ μ s	Input level for maximum short term output of 113 dB SPL @ 1 m: variable from +19 to +7dBu
Maximum long term RMS acoustic output in same conditions with IEC weighted noise (limited by driver unit protection circuit): @ 1 m ≥ 103 dB SPL @ 0,5 m ≥ 109 dB SPL	Amplifier system distortion at normal output: THD $\leq 0,05\%$ SMPTE-M $\leq 0,05\%$ CCIF-IM $\leq 0,05\%$ DIM 100 $\leq 0,05\%$	Subsonic filter below 40 Hz: 18 dB/octave
Maximum peak acoustic output per pair on top of console. @ 1 m from the engineer with music material: ≥ 124 dB	Signal to noise ratio, referred fo full output: Bass ≥ 100 dB Treble ≥ 100 dB	Ultrasonic filter above 25 kHz: 12 dB/octave
Self generated noise level in free field @ 1 m on axis: ≤ 10 dB (A)	Mains voltage: 100/200 V or 115/230 V	Crossover frequency, Bass treble: 1,8 kHz
Harmonic distortion at 90 dB SOL @ 1 m on axis: Freq. 50..100 Hz $< 1\%$ > 100 Hz $< 0,5\%$	Voltage operation range nominal $\pm 10\%$	Crossover acoustic slopes: 24 – 32 dB/octave
Drivers: Bass 250 mm (10") cone Treble 24 mm (1") metal dome Both drivers are magnetically shielded	Power consumption: Idle 50 W Full output 200 W	Treble tilt control operating range in 2 dB steps: +2 to -4 dB & MUTE
Weight: 21,7 kg (48 lb)		Bass roll-off control operating range in 2 dB steps: 0 to -8 dB @ 40 Hz
Dimensions: Height 495 mm Width 320 mm Depth 290 mm		Bass tilt control operating range in 2 dB steps: 0 to -6dB & MUTE
		The 'CAL' position is with all tone controls set to 'off' and input sensitivity control maximum.

Tab. C.20: Datenblatt Genelec 1032 A

C3 Mikrofonvorverstärker

Mikrofonvorverstärker	MIDIMAN	Audio Buddy
Mikrofoneingänge	Frequenzgang	5Hz – 22kHz, +0/-1dB
	Klirrfaktor (THD+N)	< 0.009%
	Rauschabstand	>69 dB
	Gain	max. 58 dB

Tab. C.21: Technische Daten Mikrofonvorverstärker

C4 Verwendete Mikrofone

Eigentum von	Modell	Charakteristik	Aufstellung	Inventarnummer
ORF	Schöps CMC 6--U	Kugel	Mic1	15830
			Mic2	442352
			Mic3	15825
			Mic4	15828
IEM	AKG C 480 B	Niere	Mic5	16701
			Mic6	16702
	AKG C 480 B	Niere	Temperaturmessmic.	574-4/13/6-99

Tab. C.22: Verwendete Mikrofone

C5 Technische Daten der Arraymikrofone

Verstärker-Typ	Speisung	Stromaufnahme	Impedanz	untere Grenzfrequenz (-3 dB)
CMC 6U:	12V - 48V Phantom	8mA (12V)	25 Ohm (12V)	20 Hz
	automatische Umschaltung	4mA (48V)	35 Ohm (48V)	
Maximale Ausgangsspannung:	1V (bei 1 kHz und 1 kOhm Lastwiderstand)			
Kleinster empfohlener Lastwiderstand:	600 Ohm (bei einem kleineren Wert reduziert sich insbesondere die maximale Aussteuerbarkeit)			
Länge:	116mm (incl. 3mm Kapselgewinde)			
Durchmesser:	20 mm			
Gewicht:	68 g			
Oberfläche:	matt-grau (g) oder Nickel (ni)			

Tab. C.23: Technische Daten Arraymikrofon Schöps CMC 6--U

C6 Technische Daten des AD/DA-Konverters

AD/DA Konverter	RME ADI-8DS
Dynamik AD	117 dBA
THD AD	< -110 dB
THD+N AD	< -104 dB
Übersprechdämpfung	> 130 dB
Dynamik DA	112 dBA unmuted
THD DA	< -104 dB
THD+N DA	< -102 dB
Übersprechdämpfung	> 110 dB
Frequenzgang AD/DA	-0.1 dB: 5Hz bis 21.5kHz

Tab. C.24: Technische Daten des AD/DA-Konverters RME ADI-8DS

Quellenverzeichnis

[Lit. 1]

Digitale Audiotechnik 2, Pflüger, 1999

[Lit. 2]

Psychoakustik, Zwicker, 1993

[Lit. 3]

Physikalische & technische Akustik, Neumann, 1979

[Lit. 4]

Die wissenschaftlichen Grundlagen der Raumakustik, Band 1, Cremer, Hirzel, Stuttgart, 1978

[Lit. 5]

Akustik, Hönig, 1994

[Lit. 6]

Elektroakustik AK, Graber, 1998

[Lit. 7]

Handbuch der Tonstudioteknik, Dickreiter, Sauer Verlag KG, München 1997

[Lit. 8]

Audio Analyzer 2012, Application Examples, Brüel & Kjaer

[Lit. 9]

Audio Analyzer 2012, Reference Manual, Brüel & Kjaer

[Lit. 10]

Transfer-Function Measurement with Maximum-Length Sequences, Douglas D. Rife and John Vanderkooy, 1989

[Lit. 11]

Compensating of room acoustic transfer functions affected by change of room temperature, Michiaki Omura, Motohiko Yada, Hiroshi Saruwatari, Shoji Kajita, Kazuya Takeda and Fumitada Itakura, Nagoya, Japan, 1999
kajita@nuee.nagoya-u.ac.jp

[Lit. 12]

URL: www.digidesign.de

[Lit. 13]

URL: www.iem.kug.ca.at/pdb, www.-cra.ucsd.edu/~msp/software.html

[Lit. 14]

URL: www-nt.e-technik.uni-erlangen.de/~strobels/Seminars/Datenfusion/vortrag3.pdf,
28.12.2001

[Lit. 15]

Advanced Signal Processing and Digital Noise Reduction, Vaseghi, 1996

[Lit. 16]

Digital Signal Processing, Haddad, Parson, 1991

[Oppenheim, Schafer]

Oppenheim, Schafer: "Digital Signal Processing", Prentice-Hall, 1975

[Oppenheim, Schafer]

Oppenheim, Schafer: "Zeitdiskrete Signalverarbeitung", Oldenbourg, München Wien, 1999