

# Umsetzung eines Programms zur Sensibilisierung der auditiven Wahrnehmung, Teil: Räumliche Abbildung

Bachelorarbeit aus Aufnahmetechnik 1, SE

Josef Kulmer

Betreuung: Dr. Alois Sontacchi  
Graz, 22. Juni 2010



**UNIVERSITÄT**  
FÜR MUSIK UND  
DARSTELLEND KUNST  
G R A Z - A U S T R I A

institut für elektronische musik und akustik



## Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

Graz, am 22. Juni 2010

Josef Kulmer

## Abstract

During the seminar “Aufnahmetechnik 1” a software program was created to sensitize the awareness of auditory perception. It consists of three parts: timbre, dynamics and spatial sound. This work relates to the item spatial hearing with the aim to create a training tool, which simulates the acoustics of real environments and trains the evaluation of spatial sound. Following attributes for practicing are chosen: perception of distance from sound source to listener; perception of room size and sound of the reverberation. The simulation of the virtual space is realized through a reverberation on the principle of Jean Marc Jot. The software is written in MATLAB and in PureData.

## Zusammenfassung

Im Rahmen des Seminars „Aufnahmetechnik 1“ wurde ein Software Programm für die Sensibilisierung der auditiven Wahrnehmung erstellt. Es besteht aus drei Teilen: Klangfärbung, Dynamik und räumliche Abbildung. Diese Arbeit behandelt den Teil räumliche Abbildung mit dem Ziel eines Trainingstools, welches die Akustik von realen Umgebungen virtuell nachbildet und die Bewertung von räumlichen Klängen trainiert. Als zu übende Attribute wurden gewählt: Wahrnehmung der Distanz von Hörerort zu Schallquelle; Wahrnehmung der Raumgröße und Klang des Nachhalls. Die Simulation des virtuellen Raumes wird über einen Nachhall nach dem Prinzip von Jean Marc Jot realisiert. Die Software ist in MATLAB und in PureData geschrieben.

# Inhaltsverzeichnis

0	Einleitung .....	6
1	Kurzer Exkurs in die Raumakustik.....	7
1.1	Subjektive Begriffe .....	7
1.1.1	Halligkeit (Liveness) .....	7
1.1.2	Wärme .....	7
1.1.3	Brillanz .....	7
1.1.4	Einfallswerte ASW (Spaciousness: Apparent Source Width).....	8
1.1.5	Einhüllung LEV (Spaciousness: Listener Envelopment) .....	8
1.1.6	Klarheit, Durchsichtigkeit .....	8
1.1.7	Präsenz, Intimität .....	8
1.2	Objektive Messgrößen .....	9
1.2.1	Nachhallzeit .....	9
1.2.2	Anfangsnachhallzeit (Early Decay Time EDT).....	11
1.2.3	Bassverhältnis BR .....	12
1.2.4	Dissipation .....	12
1.2.5	Hallradius $r_H$ .....	13
1.2.6	Richtentfernung $r_R$ .....	15
1.2.7	Deutlichkeitsmaß $C_{50}$ .....	15
1.2.8	Schwerpunktszeit $t_s$ .....	16
1.2.9	Hallmaß H .....	17
1.2.10	Seitenschallgrad LE, LF .....	17
1.2.11	Raumeindrucksmaß R.....	18
1.2.12	Klarheitsmaß $C_{80}$ .....	19
1.3	Zusammenfassung .....	19
2	Simulation eines Raumes .....	20
2.1	Direktanteil des Schalls .....	21
2.2	Frühe Reflektionen (Early Reflections).....	21
2.2.1	Bedeutung der frühen Reflektionen (Early Reflections) .....	21
2.2.2	Simulation der Early Reflections.....	22
2.3	Einfluss des diffusen Nachhalls .....	24
2.3.1	Nachhall nach Schroeder .....	24
2.3.2	Nachhall nach dem Prinzip des Feedback Delay Network FDN.....	29

3	Auswahl der zu übenden räumlichen Attribute .....	30
3.1	Distanz zur Quelle .....	31
3.2	Raumgröße .....	32
3.3	Raumklang .....	33
4	Versuchsaufbau .....	33
4.1	Abhörposition .....	33
4.2	Aufbau des Hallmoduls in PD .....	34
4.2.1	Simulation des direkten Schallanteils in PD .....	34
4.2.2	Simulation von Early Reflections in PD .....	35
4.2.3	Einstellungen des Feedback Delay Network in PD.....	35
4.2.4	Steuerung des Hallmoduls .....	35
5	Ablauf des Trainingsprogramms und erste Ergebnisse .....	37
5.1	Einstellungen des Trainingsprogramms.....	38
5.2	Ergebnisse und Erfahrungsberichte .....	41
5.3	Mögliche Verbesserungen des Trainingsprogramms .....	42
6	Literaturverzeichnis .....	43

Im folgenden Text wird ausschließlich das generische Maskulin verwendet, der Text richtet sich gleichermaßen an Frauen und Männer.

## 0 Einleitung

Unsere Umgebung erfassen wir mit unseren Sinnen und nehmen sie nicht nur visuell, sondern auch *auditiv* wahr. Egal ob wir uns in einem Konzertsaal befinden, eine Wanderung in den Bergen unternehmen oder uns im Badezimmer für den Tag vorbereiten, jede dieser Umgebungen besitzt einen charakteristischen Klang.

Die Informationen, welche wir durch die auditive Wahrnehmung erhalten sind nur ein Teil des kognitiven Prozesses. Im Unterbewusstsein wird die Summe aller Sinneswahrnehmungen verarbeitet und dadurch entsteht ein Abbild der Realität. Nur selten sind wir ausschließlich auf die auditive Wahrnehmung angewiesen, welche daher mit gezieltem Training noch erweitert werden kann. Eine unbekannte Umgebung nur durch *Hören* besser beschreiben zu können, ist das Ziel dieser Arbeit.

Das auditive räumliche Bild, welches wir wahrnehmen, ist hauptsächlich über die Architektur und die darin enthaltenen Materialien des Raumes gekennzeichnet, und kann aufgrund ihrer Komplexität nur näherungsweise simuliert werden. Seit vielen Jahren wird versucht, den Klang in einem Raum, nur durch Kenntnis der Eigenschaften des Raumes, nachzubilden. Auch für dieses Trainingsprogramm ist die Wahl des „Raum-Simulators“ ein wesentliches Kriterium, da für eine tatsächliche Erweiterung der auditiven Wahrnehmung ein reales Abbild von natürlich vorkommenden Klang-Kulissen unabdingbar ist. Wenn das Trainingsprogramm mit unnatürlichen, in der Realität nicht vorkommenden Klang-Bildern arbeitet, würde die Testperson zwar persönliche Verbesserungen ihrer auditiven Wahrnehmung bemerken, könnte diese aber in der Praxis nicht anwenden. Daher wurde ein besonderer Fokus auf die Auswahl des Simulationsmodells gelegt.

Die Struktur dieser Arbeit ist wie folgt aufgebaut: Zunächst werden Gütemaße der Raumakustik vorgestellt, mit einer Unterteilung in subjektive Begriffe und objektive Messgrößen. Darauf folgend werden die Möglichkeiten einen Raumklang zu simulieren, bzw. jene die für dieses Trainingsprogramms gewählt wurden, vorgestellt. Weiters wird eine Übersicht über bestehende Trainingsprogramme geboten und welche Möglichkeiten das beschriebene Programm geben soll. Schlussendlich werden die Einstellungen der akustischen Raumsimulation beschrieben, welche nötig sind um die gestellten Anforderungen zu erfüllen und der genaue Testaufbau mit ersten Ergebnissen und Rückmeldungen geschildert.

# 1 Kurzer Exkurs in die Raumakustik

## 1.1 Subjektive Begriffe

Wenn sich Menschen verschiedenster (beruflicher) Herkunft untereinander verständigen wollen, müssen zunächst gleiche / ähnliche Begriffe gefunden werden, die gleiches / ähnliches beschreiben. So auch für den Klang von Räumen, da Musiker und Techniker aufgrund ihrer verschiedenen Erfahrungen und Ausbildungen sich untereinander oft nur schwer verständigen können. Nachfolgend sind nur einige der oft verwendeten Begriffe aufgeführt. Natürlich gibt es noch viele weitere spezifischere Begriffe, um einen Raumklang (welcher nicht nur Innen-, sondern auch Außenräume, wie Naturtheater, einschließt) zu beschreiben. Hier sei auf *Beranek (1996)* und *Wilkins und Plenge* (vgl. Ahnert und Reichardt 1975) verwiesen, welche viele dieser Begriffe im Laufe einer Konzertreise mit den Berliner Philharmonikern prägten.

### 1.1.1 Halligkeit (Liveness)

Im Inneren von Gewölbekellern, leeren Zimmern mit nackten Wänden oder Badezimmern klingt es „halliger“, „lebendiger“ als z.B. in Holzhütten oder Wohnzimmern, welche eher „trocken“ oder im Extremfall „tot“ klingen. Dieser für jeden einfach zu bemerkende Unterschied wird als Halligkeit bezeichnet. Sie stimmt mit der objektiven Messgröße Nachhallzeit im sensiblen Hörbereich (350- 1400Hz) sehr gut überein. (vgl. Beranek 1996).

### 1.1.2 Wärme

Dieser Parameter steht in Verbindung mit dem Bassanteil einer musikalischen Darbietung. Mit zunehmenden Bässen (75 – 350Hz) im Vergleich zu den mittleren Frequenzen (350 – 1400Hz) steigt die empfundene „Wärme“ (vgl. Beranek 1996). Da viele Materialien Tiefpasscharakter haben, und dadurch die höheren Frequenzen dämpfen, ist in fast allen Räumen die Nachhallzeit im Bassbereich länger, was zu einer stärker empfundenen „Wärme“ führt. Falls sie zu dominant wird, beschreiben Musiker einen Raum auch als *dunkel*.

### 1.1.3 Brillianz

Mit „Brillanz“ eines Raumes werden die Attribute *klarer* und *heller* Klang in

Verbindung gebracht. Bei einem „brillianten“ Klang sind die höheren Frequenzen ( $>1400\text{Hz}$ ) dominant und werden vom Raum nur wenig bedämpft (vgl. Beranek 1996). Für Text- und Sprachwiedergabe ist eine höhere „Brillianz“ vorteilhaft, da einzelne kurze Laute mit weniger Energie (wie z.B. Konsonanten) im Raum besser hörbar sind (vgl. Ahnert und Reichardt 1981).

Im Allgemeinen ist sie aber schwerer zu erzielen, da (wie in Kap.: 1.1.2 erwähnt) viele Materialien hohe Frequenzen dämpfen, anstatt sie zu verstärken. Eine Ausnahme bilden Räume mit vielen Glaswänden, welche als eine der wenigen Materialien vorwiegend im Bassbereich dämpfen.

#### 1.1.4 Einfallswerte ASW (Spaciousness: Apparent Source Width)

Eine hohe „Spaciousness“ bedeutet, dass die Klangquelle „breiter“ wirkt. Dieser Effekt tritt auf, wenn vor allem in den ersten Sekundenbruchteilen viel Schallenergie vom Raum reflektiert wird, und dann seitlich am Hörer eintrifft.

#### 1.1.5 Einhüllung LEV (Spaciousness: Listener Envelopment)

Wie nachhallend ein Klang ist, und von welchen Richtungen er kommt, beschreibt die „Einhüllung“. Ein gleichmäßiger Einfall der Schallenergien nach der Anregung aus allen Richtungen ergibt ein größtmögliches Gefühl der „Einhüllung“. Dieses möglichst diffuse Schallfeld wird erst durch spät einfallende Schallenergien ( $t > 100\text{ms}$ ) erzeugt (vgl. Beranek 1996) und ist für eine gute Akustik wünschenswert.

#### 1.1.6 Klarheit, Durchsichtigkeit

Ist es einfach, verschiedenen Klänge (unabhängig davon, ob sie gleichzeitig oder zeitlich versetzt auftreten) zu unterscheiden, spricht man von guter Klarheit. Mit steigender Halligkeit sinkt die Klarheit meist (vgl. Ahnert und Reichardt 1981). Ihr ist auch der wichtige objektive Parameter Klarheitsmaß  $C_{80}$  zugeordnet.

#### 1.1.7 Präsenz, Intimität

Laut *Beranek (1996)* wirken architektonisch kleine Räume *intimer*. Ruft ein Klang den Eindruck von einem kleinen Raum hervor, spricht man von „akustischer Intimität“ oder auch „Präsenz“.

## 1.2 Objektive Messgrößen

Die genannten subjektiven Begriffe sind in der Praxis schwer zu handhaben, da selbst einheitliche Definitionen wie „Wärme“ oder „Klarheit“ von unterschiedlichen Personen nur bis zu einem gewissen Grad einheitlich beurteilt werden. Ziel sind also Beschreibungen, welche direkt von reproduzierbaren physikalischen Messmethoden abhängig sind und den subjektiven Empfindungen bestmöglich entsprechen. Dieses Kapitel gibt einen kurzen Überblick darüber, beschränkt sich aber auf Größen, welche im weiteren Verlauf dieser Arbeit relevant sein werden.

### 1.2.1 Nachhallzeit

Diese mit Abstand bekannteste akustische Eigenschaft eines Raumes wurde erstmals von *Sabine* beschrieben. Durch Versuche mit Orgelpfeifen (vgl. Cremer und Müller 1976) erkannte er, dass die erzeugte Schallenergie immer in der gleichen Zeit auf einen gewissen Bruchteil der Anfangsenergie zurückfällt, d.h. der zeitbezogene Energieabfall unabhängig vom Anfangspegel ist:

$$-\frac{dE(t)}{dt} = \frac{E(t)}{\tau} \quad (1)$$

$E(t)$ ...Schallenergie in Abhängigkeit der Zeit  $t$

$\tau$ ...Abklingzeit

Durch Lösen der Differenzialgleichung nach  $t$  erhält man:

$$t = \tau \ln \frac{E_0}{E(t)} \quad (2)$$

$E_0$ ...Schallenergie zu Beginn ( $t=0$ )

Sabine definierte die Nachhallzeit  $T$  als jene Zeit, innerhalb welcher der mittlere Schalldruck nach seiner Anregung auf den tausendsten Teil abfällt.

$$E_0 = 10^6 E(t) \Big|_{t=T} \quad (3)$$

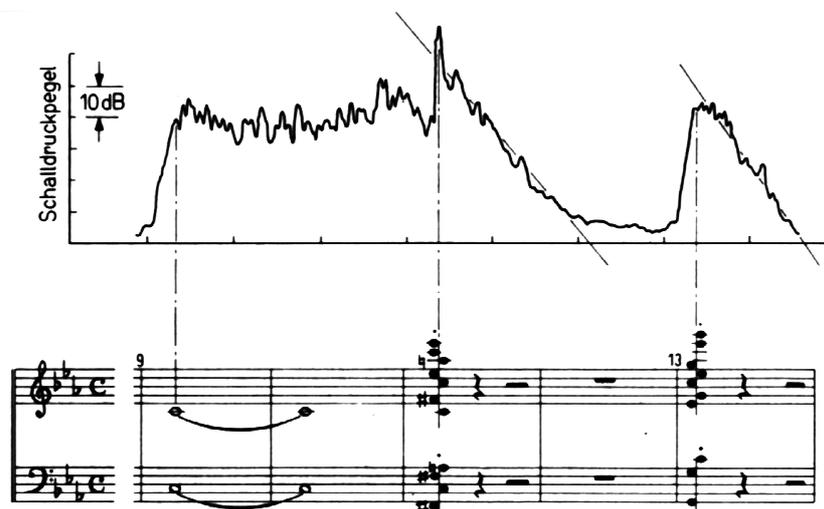
$$T = \tau \ln \frac{E_0}{E(t)} = \tau \ln 10^6 \quad (4)$$

$$E(t) = E_0 e^{-\frac{6t \cdot \ln(10)}{T}} = E_0 \cdot 10^{-\frac{6t}{T}} \quad (5)$$

Angegeben wird die Nachhallzeit meist mit der Abkürzung „RT<sub>60</sub>“ anstatt „T“, und bedeutet „Reverberation Time“ für einen Pegelabfall um 60dB.

Nach Formel (5) fällt die Energie des Nachhalls exponentiell ab, und kann daher in logarithmischer Darstellung als Gerade gezeichnet werden. Abbildung 1 zeigt eine Nachhallaufzeichnung eines Konzertes. Nach einer Schallanregung in Takt 11 und Takt 13 folgt eine kurze Pause, in der der Nachhall annähernd der eingezeichneten Gerade abklingt.

Durch die verschiedenen Absorbereigenschaften der verwendeten Materialien ist die Nachhallzeit frequenzabhängig, deshalb wird sie für mehrere Frequenzbänder getrennt angegeben. Als „Einzahlwert“ (vgl. Ahnert und Tennhardt 2008) wird die Nachhallzeit als Mittelwert der Nachhallzeiten bei 500Hz und 1000Hz Oktavmittenfrequenz berechnet.



**Abbildung 1: Nachhallaufzeichnung während eines Konzertes  
(Beethoven Op. 62, Coriolan-Ouverture, Takt 9...13  
Quelle: Cremer und Müller (1976)**

Einige Nachhallzeiten bekannter Aufführungshäuser (Beranek 1996)

Boston Symphony Hall	$RT_{60} = 1,85s$
Wien, Großer Musikvereinssaal	$RT_{60} = 2,0s$
New York, Carnige Hall	$RT_{60} = 1,8s$

Die Nachhallzeit ist wohl das am öftesten genannte Kriterium für die Beschreibung eines Raumklanges. Laien und Profis sprechen gerne von „mehr“ und „weniger Hall“ und meinen meist damit die Nachhallzeit, obwohl bewiesen ist, dass es viele (und wichtigere) Parameter gibt (vgl. Ahnert und Reichardt 1981). Da sie die verstrichene Zeit bei einem Pegelabfall von 60dB wiedergibt (unabhängig der Art des Abklingens), und diese extremen Dynamiken im Alltag selten vorkommen, kann man ihr nur in Verbindung mit anderen raumakustisch relevanten Parametern (welche in folgenden Kapiteln näher beschrieben werden) Bedeutung zumessen.

### 1.2.2 Anfangsnachhallzeit (Early Decay Time EDT)

Obwohl die Nachhallzeit an fast allen Plätzen innerhalb eines Raumes die selbe ist, schwankt die subjektiv empfundene Nachhallzeit. Durch genaue Betrachtung des Nachhall-Verlaufes kann dieser Effekt erklärt werden. Der Abfall der Energiedichte ist logarithmisch dargestellt nur annähernd eine Gerade. Fällt zu Beginn die Energie übermäßig schneller ab als im Mittel („durchhängender Verlauf“), so erscheint die Nachhallzeit auch kürzer. Offensichtlich hat genau dieser frühe Bereich einen großen Einfluss auf die empfundene Nachhallzeit. Jordan berechnete aus der Zeit, welche zwischen einem Pegelabfall von 0 auf -10dB verstreicht, die „Early Decay Time EDT“. Da bei Sprache oder musikalischen Darbietungen der Nachhall oft bei einem Abfall um wenige Dezibel schon von neuen darauffolgenden Quellsignalen verdeckt wird (vgl. Cremer und Müller 1976), hat die EDT für diese Anregungssignale mehr Aussagekraft, als die Nachhallzeit  $RT_{60}$ .

Zur Vergleichbarkeit mit der Nachhallzeit werden die gemessenen Werte der EDT mit sechs multipliziert, was einer Extrapolation auf 60dB entspricht.

Mit Anfangsnachhallzeit wurden früher auch andere Berechnungen in Verbindung gebracht, welche meist andere Integrationsgrenzen als die der EDT vorschreiben, wie z.B. der Abfall des Nachhalls um die ersten -15dB oder -20dB (vgl. Kürer und Kurze 1967 zitiert nach Cremer und Müller 1976).

### 1.2.3 Bassverhältnis BR

Diese raumakustische Eigenschaft wird eng in Verbindung gebracht mit dem subjektiven Begriff „Wärme“ Kap. 1.1.2. Das Bassverhältnis BR (vgl. Beranek 1962 zitiert von Ahnert und Tennhardt 2008) errechnet sich als Verhältnis der Nachhallzeiten bei Oktavmittenfrequenzen von 125 Hz und 250 Hz zu Oktavmittenfrequenzen von 500 Hz und 1000 Hz.

$$BR = \frac{T_{125\text{Hz}} + T_{250\text{Hz}}}{T_{500\text{Hz}} + T_{1000\text{Hz}}} \quad (6)$$

Für Musik wird ein Bassverhältnis von  $BR = 1,0 \dots 1,3$  angestrebt (Festspielhaus Bayreuth:  $BR = 1,11$ ; Festspielhaus Salzburg  $BR = 1,10$ ; Staatsoper Wien  $BR = 1,10$  (vgl. Beranek 1996)), für Sprache dagegen sollte das Bassverhältnis höchstens einen Wert von  $BR = 0,9 \dots 1,0$  aufweisen (vgl. Ahnert und Tennhardt 2008).

### 1.2.4 Dissipation

Bei der Ausbreitung der Schallwellen wird ein Teil der Schallenergie in Wärme umgewandelt (Dissipation). Dies führt zu einer Pegelminderung, welche von Temperatur, Luftfeuchtigkeit, Frequenz des Schalls und Distanz zur Quelle abhängt (für Luft als Medium). Die Dissipation hat Tiefpassverhalten und dämpft mehr mit steigender Frequenz und sinkender Luftfeuchtigkeit.

$$E(r, f, \varphi) = E_0 e^{-\mu r} \quad (7)$$

$$\text{mit dem Korrekturfaktor } \mu = \frac{85}{\varphi} \left( \frac{f}{1000} \right)^2 10^{-4} \quad (8)$$

mit  $r$ ...Distanz Quelle – Senke in m  
 $f$ ...Frequenz der Schallenergie in Hz  
 $\varphi$ ...relative Luftfeuchtigkeit in %  
 $E(r, f, \varphi)$ ...Schallenergie an der Senke  
 $E_0$ ...Schallenergie der Quelle

Formel (8) beruht auf einer Näherungsformel (vgl. Evans und Bazley 1956 zitiert von

Cremer und Müller 1976), die für relative Luftfeuchtigkeiten von 30% - 80%, einer Raumtemperatur von 20°C und im Frequenzbereich  $f < 10\text{kHz}$  gültig ist.

In der Realität spielt sie eine untergeordnete Rolle, da sie nur bei großen Distanzen einen Einfluss hat (nach Formel (7) wird Schallenergie mit 300Hz auf einer Länge von 100m um 1,64% ( $\varphi = 60\%$ ) abgesenkt).

### 1.2.5 Hallradius $r_H$

Für die vereinfachte Modellierung von Kugelstrahlern beschreibt der Hallradius  $r_H$  die Beziehung der Energien des Direktschalls und des Diffusschallfelds. Die Energie des Direktschalls kann über die Kugelwellenausbreitung beschrieben werden.

$$E_{dir} = \frac{P_{dir}}{4\pi r^2 c} \quad (9)$$

mit  $E_{dir}$ ...Energie des Direktschalls  
 $P_{dir}$ ...Schalleistung  
 $c$ ...Schallausbreitungsgeschwindigkeit  
 $r$ ...Distanz Quelle – Senke

Schaltet man eine Schallquelle in einem nicht schalltoten Raum ein, baut sich das Schallfeld allmählich auf, bis die Energiedichte einen stationären Wert  $E_{st}$  erreicht, bei dem die absorbierte Leistung des Raumes  $P_{abs}$  der von der Schallquelle abgestrahlten Leistung  $P_{dir}$  die Waage hält (vgl. Cremer und Müller 1976).

$$P_{dir} = P_{abs} = \frac{1}{4} E_{st} \cdot c \cdot A_{ges} \quad (10)$$

daraus folgt:

$$E_{st} = \frac{4P_{dir}}{c \cdot A_{ges}} \quad (11)$$

$E_{st}$ ...stationäre Energiedichte  
 $A_{ges}$ ...Absorptionsfläche mit Absorptionsgrad  $\alpha=1$   
 $c$ ...Schallausbreitungsgeschwindigkeit

Der Hallradius ist nun jener Abstand, bei dem Direktschallenergie  $E_{dir}$  und Diffusschallenergie  $E_{st}$  (=stationäre Energiedichte) gleich groß sind.

$$E_{st} = E_{dir} \quad (12)$$

Durch Umformen auf den Radius  $r = r_H$  erhält man

$$r_H = \frac{\sqrt{A_{ges}}}{4\sqrt{\pi}} \quad \text{mit } A_{ges} \approx 0,163 \frac{V}{T} \quad (13)$$

$$r_H = 0,057 \sqrt{\frac{V}{T}} \quad (14)$$

V...Volumen des Raumes

T...Nachhallzeit des Raumes, auch  $RT_{60}$

Die Raumwände halten also (bei üblichen Nachhallzeiten) die Energie im Raum, so dass auch bei größeren Abständen zwischen Sprecher und Hörer das Nutzsiegel genügend groß bleibt (vgl. Cremer und Müller 1976). In Abbildung 2 ist das Diffusschallfeld als konstanter Pegel unabhängig von der Entfernung zur Quelle erkennbar, das Direktschallfeld nimmt pro Verdopplung der Entfernung um -6dB ab (vgl. Ahnert / Reichardt 1981). Am Punkt  $r/r_H = 1$  ist die Entfernung zur Quelle gleich dem Hallradius, beide Schallfelder addieren sich zu einem um +3dB stärkeren Gesamtschallfeld.

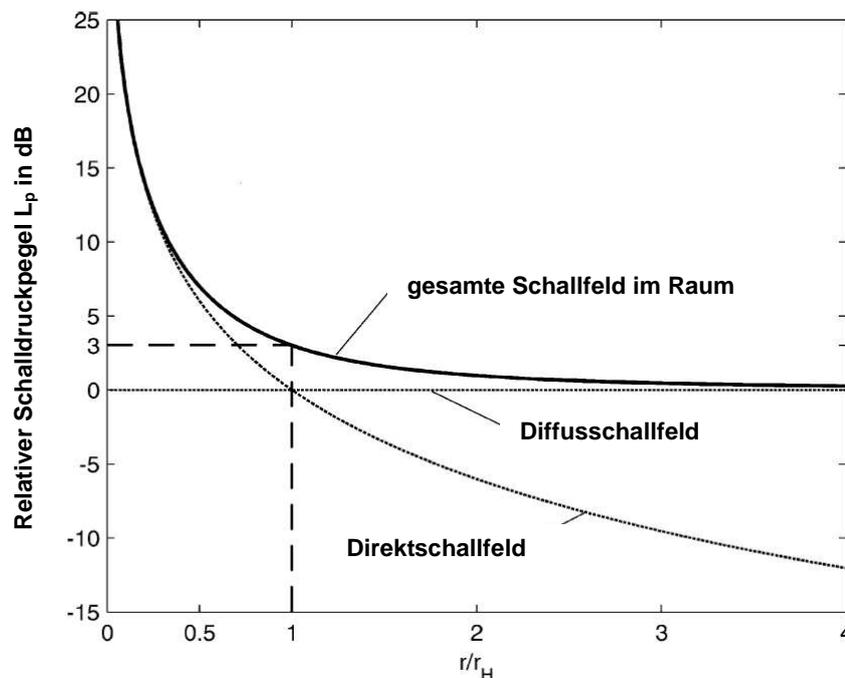


Abbildung 2: Verhältnis Direkt zu Diffus-Schallanteil im Raum

Quelle: Ahnert und Tennhardt 2008

Es ist zu beachten, dass der Hallradius nur für Kugelwellen als Quelle ein geeignetes Maß ist. Die meisten Schallquellen haben keine gleichmäßige Energieabstrahlung in alle Richtungen, d.h. sie sind gerichtet, wie z.B. Sprache. Dies berücksichtigt die sogenannte Richtentfernung.

### 1.2.6 Richtentfernung $r_R$

Die Richtentfernung  $r_R$  gibt, ähnlich wie der Hallradius  $r_H$ , die Distanz zwischen Quelle und dem Punkt im Raum an, bei dem die Energie des Direktschalls der Energie des Diffusschalls entspricht, berücksichtigt aber auch die Richtwirkung der Quelle.

$$r_R = \sqrt{\gamma} \cdot r_H \quad (15)$$

mit  $r_R$ ...Richtentfernung  
 $\gamma$ ...Bündelungsgrad  
 $r_H$ ...Hallradius

Der Bündelungsgrad  $\gamma$  ist angegeben als das Verhältnis der Schallintensität in Hauptabstrahlrichtung zum Mittelwert über alle Richtungen.

$$\gamma = \frac{I_{\text{Hauptabstrahlachse}}}{I_{\text{Mittel}}} \quad (16)$$

Bei gerichteten Lautsprechern oder z.B. Trompeten im fortissimo kann der Bündelungsgrad  $\gamma$  Werte von bis zu 100 annehmen, das gleichbedeutend mit einer Vergrößerung des „effektiven Hallradius“ um den Faktor 10 ist (vgl. Cremer und Müller 1976).

### 1.2.7 Deutlichkeitsmaß $C_{50}$

Für gute Text- und Sprachverständlichkeit ist es wichtig, dass der Nachhall im Vergleich zum Direktschall nicht zu dominant wird. Ansonsten „verschmiert“ das Signal stärker, und Silben werden dadurch schwerer unterscheidbar. Das Deutlichkeitsmaß  $C_{50}$  gibt darüber eine Aussage, indem es die verzögerten Reflexionen in zeitliche Bereiche teilt, welche für die Verständlichkeit unterstützend sind. Diejenigen, welche innerhalb der ersten 50ms nach Eintreffen des Direktschalls ( $W_{0...50ms}$ ) auftreten, unterstützen die

Verständlichkeit und Deutlichkeit (vgl. Cremer und Müller 1976), die später einfallenden Reflexionen ( $W_{50ms...∞}$ ) wirken eher störend. Dadurch ergibt sich das Deutlichkeitsmaß  $C_{50}$  als:

$$C_{50} = 10 \cdot \log \frac{W_{0...50ms}}{W_{50ms...∞}} \quad (17)$$

*Ahnert und Reichardt (1981)* begründen die Integrationsgrenze von 50ms als statistisch ermittelte Durchschnittsgröße und geben für gute Verständlichkeit ein Deutlichkeitsmaß  $C_{50} > -2\text{dB}$  an.

Parallel zum Deutlichkeitsmaß existiert der Deutlichkeitsgrad  $D$  oder  $D_{50}$  (vgl. Thiele 1953 zitiert nach Ahnert und Tennhardt 2008), welcher in Prozent angegeben wird und den Energieanteil der ersten 50ms auf die Gesamtenergie bezieht. Beide Messgrößen können laut Formel (18) ineinander umgerechnet werden, laut *Ahnert und Tennhardt (2008)* verliert der Deutlichkeitsgrad deshalb zunehmend an Bedeutung.

$$C_{50} = 10 \cdot \log \frac{D_{50}}{1 - D_{50}} \quad (18)$$

### 1.2.8 Schwerpunktszeit $t_s$

Das Deutlichkeitsmaß  $C_{50}$  setzt eine Grenze, ab der gute Verständlichkeit gegeben ist, im Gegensatz zur Schwerpunktszeit  $t_s$ , welche „ein Maß für die zeitliche Lage hoher Leistungen“ ist (vgl. Graber 2008). Sie ist definiert als (vgl. Kürer 1971 zitiert von Ahnert und Tennhardt 2008):

$$t_s = \frac{\int_{t=0}^{\infty} t \cdot p^2(t) dt}{\int_{t=0}^{\infty} p^2(t) dt} \quad (19)$$

mit  $t_s$ ...Schwerpunktszeit  
 $p$ ...Schalldruck am Hörerplatz

Mit steigender Schwerpunktszeit steigt auch der Eindruck der empfundenen „Halligkeit“ (siehe Kap.: 1.1.1.) Für Sprache gilt eine Schwerpunktszeit von  $t_s = 60\text{ms}$  bis  $80\text{ms}$  als gut, für Musik wird der Bereich auf  $t_s = 70\text{ms} - 150\text{ms}$  gesetzt. (Der größere empfohlene Bereich bei Musik ist dadurch begründbar, dass je nach Art der

musikalischen Darbietung mehr oder weniger „Halligkeit“ oder „Räumlichkeit“ empfehlenswert ist).

### 1.2.9 Hallmaß H

Das Hallmaß H beschreibt die Halligkeit eines Raumes und ist umso größer, je größer der Anteil des Nachhalls ( $t > 50\text{ms}$ ) gegenüber den frühen Reflexionen ist (vgl. Cremer und Müller 1976). Das Hallmaß setzt diese Energien in ein Verhältnis:

$$H = 10 \cdot \log \frac{W_{50\text{ms}-\infty}}{W_{0...50\text{ms}}} \quad (20)$$

Als optimal für Konzertsäle haben sich Werte zwischen 0dB und 4dB und für Musiktheater mit Konzernutzung Werte zwischen -2dB und 4dB erwiesen.

### 1.2.10 Seitenschallgrad LE, LF

Die seitlich einfallenden Reflexionen verursachen einen Effekt der Räumlichkeit und tragen mehr als die Halligkeit zum Raumeindruck bei (vgl. Cremer und Müller 1976). Sie erzeugen eine scheinbare Ausdehnung der Schallquelle (Apparent Source Width ASW siehe Kap.: 1.1.4), die nicht mehr als punktförmig wahrgenommen wird, obwohl sie annähernd eine Punktquelle ist (vgl. Ahnert und Tennhardt 2008). Wesentlich für diesen Effekt sind die frühen Reflexionen aus seitlichen Einfallsrichtungen. In schmalen hohen Konzertsälen treten sie besonders stark auf, da die seitlichen Reflexionen den Zuhörer zeitlich früher erreichen, als die Deckenreflexionen.

Der Seitenschallgrad Lateral Efficiency (LE) (vgl. Jordan 1980 zitiert von Ahnert und Tennhardt 2008) unterscheidet sich vom Seitenschallgrad Lateral Fraction (LF) (vgl. Barron 2003 zitiert von Ahnert und Tennhardt 2008) dadurch, dass *Barron* die Reflexionen ab 5ms beachtet, *Jordan* erst ab 25ms. Durch diese unterschiedliche Bewertung ergeben sich für die LF nach Barron minimal höhere Werte. Beiden gemeinsam ist die Gewichtung des Seitenschallanteils mit  $\cos^2\alpha$ , ( $\alpha$ ...Einfallswinkel), was durch Verwenden eines Mikrofons mit Achtercharakteristik erfüllt wird.

Seitenschallgrad Lateral Efficiency (LE) nach *Jordan*:

$$LE = \frac{(W_{25...80ms})_{seitlich}}{(W_{0...80ms})_{gesamt}} \quad (21)$$

mit  $W_{seitlich}$ ...Energie der einfallenden Reflektionen gewichtet mit  $\cos^2\alpha$

$W_{gesamt}$ ...Energie der von allen Richtungen einfallenden Reflektionen

Seitenschallgrad Lateral Fraction (LF) nach *Barron*:

$$LF = \frac{(W_{5...80ms})_{seitlich}}{(W_{0...80ms})_{gesamt}} \quad (22)$$

Je größer der Seitenschallgrad ist, desto „akustisch breiter“ wirkt die Quelle. Für Konzerthäuser haben sich Werte von 25% - 40% für den Seitenschallgrad (LF) als optimal erwiesen (vgl. Cremer und Müller 1976).

### 1.2.11 Raumeindrucksmaß R

Mit diesem Maß kann das Hallmaß und das Seitenschallmaß zusammengefasst werden:

$$R = 10 \cdot \log \frac{(W_{25...80ms})_{gesamt} - (W_{25...80ms})_{vorne}}{(W_{25...80ms})_{gesamt} + (W_{25...80ms})_{vorne}} \quad (23)$$

$(W_{25...80ms})_{vorne}$ ...direkt einfallende Energie<sup>1</sup> zwischen 25ms und 80ms nach Auftreten eines Impulses

Für einen Raumeindruck muss zwischen Direktschall und Raumreflexionen unterschieden werden. Nur sehr frühe Raumreflexionen sind vom Direktschall nicht unterscheidbar. Das Raumeindrucksmaß teilt die Schallenergie nach einfallender Richtung ein, und bewertet die seitlich einfallenden Reflektionen höher, als die von Richtung des Direktschalls, da die seitlich einfallenden Reflektionen mehr zum Raumeindruck beitragen (vgl. Ahnert und Reichardt 1981). Raumeindrucksmaße von  $R > 0\text{dB}$  sind einfach zu erreichen, als optimaler Wert wird  $R=(4\pm 2)\text{dB}$  angegeben (vgl. Lehmann 1974 zitiert von Ahnert und Reichardt 1981).

<sup>1</sup> gemessen mit Richtmikrofon (Öffnungswinkel  $\pm 40^\circ$ ) bei 500-1000Hz direkt auf Schallquelle gerichtet (vgl. Ahnert und Tennhardt 2008)

### 1.2.12 Klarheitsmaß $C_{80}$

Musikalische Darbietungen erfordern im Allgemeinen mehr Raumanteil, als ein gesprochener Text. Dennoch soll eine gewisse Durchsichtigkeit vorhanden sein, um z.B. feine Phrasierungen klar hören zu können. Das Klarheitsmaß  $C_{80}$  beschreibt dies, indem es früher eintreffende Schallenergie (wichtig für Klarheit) in Bezug zu später eintreffender Energie (Räumlichkeit, Gefühl des Einhüllens) stellt (vgl. Ahnert und Tennhardt 2008).

$$C_{80} = 10 \cdot \log \frac{W_{0..80ms}}{W_{80ms..\infty}} \quad (24)$$

Das Klarheitsmaß  $C_{80}$  ähnelt der Formel nach (mit verschiedenen Integrationsgrenzen) sehr dem Deutlichkeitsmaß  $C_{50}$ , welche die Verständlichkeit für Sprache beschreibt. Mit steigender Klarheit sinkt die Räumlichkeit (vgl. Ahnert und Reichardt 1981), deshalb wird für das Klarheitsmaß nicht nur ein Mindest- sondern auch ein Maximalwert empfohlen, um nicht zuviel Räumlichkeit zu verlieren. Man gibt also optimale Bereiche an, in denen sich diese beiden Attribute ergänzen. Als gute Werte gelten  $-3\text{dB} \leq C_{80} \leq +4\text{dB}$ . Dieser Bereich ist natürlich auch abhängig von der Art der gespielten Musik. So wird für die Wiedergabe von klassischer Musik (Betonung eher auf Melodien und Stimmen) eher mehr Klarheit gewünscht, was einen deutlich höheren unteren Grenzwert von  $C_{80} \geq -1.6 \text{ dB}$  bedeutet (Abdel Alim, Reichardt zitiert von Ahnert und Tennhardt 2008). Für Romantik (eher mehr Einhüllen der Musik, Melodien werden mehr durch Akkorde unterstrichen als bei Klassik) wird mehr Raumanteil empfohlen, der untere Grenzwert sinkt auf  $C_{80} \geq -4.6 \text{ dB}$ .

## 1.3 Zusammenfassung

Die hier aufgeführte Liste an Begriffen ist nicht vollständig, beinhaltet aber die wesentlichen Begriffe, welche für die weiteren Kapitel benötigt werden.

In der folgenden Tabelle findet sich eine Zusammenführung der genannten subjektiven und objektiven Begriffe:

Subjektive Begriffe	Merkmalpaar	Objektive Begriffe
Halligkeit	hallig - trocken	Nachhallzeit $RT_{60}$
Präsenz		Anfangsnachhallzeit EDT
Durchsichtigkeit	deutlich – undeutlich	Deutlichkeitsmaß/grad $C_{50}/D_{50}$
	klar – verschwommen	Klarheitsmaß $C_{80}$
		Schwerpunktszeit $t_s$
Einhüllung LEV	konzentriert - diffus	Raumeindrucksmaß R
Einfallswerte ASW		Seitenschallgrad LE, LF
	hell – dunkel	Bassratio BR
Brillanz	aufdringlich - zurückhaltend	
Wärme	scharf – stumpf	Frequenzgang der Nachhallzeit
	brilliant - matt	$RT_{60}$

**Tabelle 1: Übersicht über subjektive / objektive Messgrößen**

Quelle: Ahnert und Reichardt (1981)

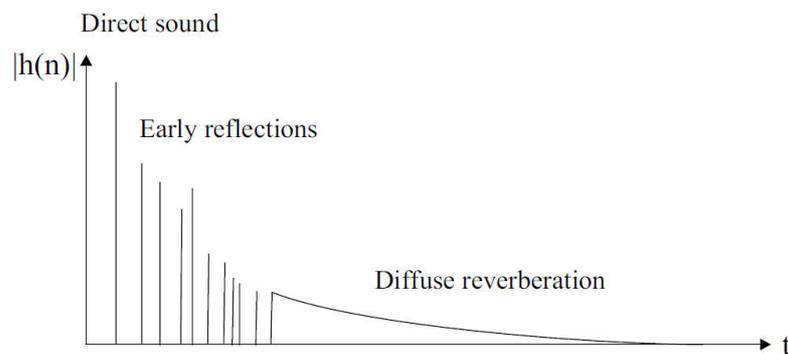
## 2 Simulation eines Raumes

Ausgehend vom Standpunkt der Signalverarbeitung kann ein Raum als Übertragungsfunktion mit einer oder mehreren Schallquellen am Eingang und einem Empfänger am Ausgang beschrieben werden. Für die Annahme, dass der Raum statisch ist (es gibt keine physischen Bewegungen), gilt er näherungsweise als LTI<sup>1</sup> System (vgl. Gardner 1998). Die Übertragungsfunktion lässt sich komplett über die Impulsantwort des Raumes beschreiben und die Signale an den Empfängern ergeben sich durch Faltung mit den Signalen der Schallquellen am Eingang.

Für den Fall, dass es den zu simulierenden Raum in der Realität nicht gibt, ist es möglich, mit genauer Kenntnis über die Beschaffenheit des Raumes, die Impulsantwort zu errechnen. Dieses Verfahren wird „auralization“ genannt (vgl. Kleinert 1993 zitiert von Gardner 1998) und bestimmt zuerst die Impulsantwort, welche dann mit dem Eingangssignal gefaltet wird. Die Nachteile dieser Methode sind, dass einerseits eine Echtzeitfaltung mit langen Impulsantworten sehr aufwendig ist und andererseits mit jeder Änderung eines Parameters die Impulsantwort neu berechnet werden muss (vgl. Gardner 1998).

<sup>1</sup> LTI-Systeme sind linear und unabhängig von zeitlichen Verschiebungen

Ein anderer Ansatz ist, nur die prägnanten Eigenschaften der Impulsantwort zu simulieren. Durch genauere Betrachtung einer aufgenommenen Raumimpulsantwort (Abbildung 3) kann man physikalische Rückschlüsse ziehen. Zuerst gelangt der Direktanteil des Schalls zum Empfänger, gefolgt von einigen Reflektionen, die aufgrund ihrer größeren Distanz zur Quelle zeitverzögert ankommen (Early Reflections). Mit steigender Zeit treffen immer mehr Reflektionen in kürzeren Zeitabständen von allen



**Abbildung 3: Schematische Darstellung der Raumimpulsantwort  $h(n)$**   
Quelle: Extra, Simmer, Fischer, Bitzer (2006)

Richtungen ein, bis man von diffusem Nachhall spricht. Das Auftrennen der Raumimpulsantwort in diese drei zeitlichen Bereiche (Direktschall, Early Reflections, diffuser Nachhall), erleichtert die Simulation, indem jeder Bereich einzeln berechnet und anschließend mit den anderen zusammengeführt wird. Im Folgenden werden die einzelnen Komponenten vorgestellt.

## 2.1 Direktanteil des Schalls

Besteht eine direkte Verbindung von Schallquelle zu Hörerplatz, so trifft der Direktanteil als erstes ein. Die Intensität nimmt mit steigender Distanz linear ab, spektrale Veränderungen sind nur über größere Distanzen bemerkbar (siehe Dissipation Kap. 1.2.5)

## 2.2 Frühe Reflektionen (Early Reflections)

### 2.2.1 Bedeutung der frühen Reflektionen (Early Reflections)

Schall legt in 100ms eine Länge von 34m, in 200ms 68m zurück. Das bedeutet, bis ein diffuser Nachhall entsteht, breiten sich einzelne Reflektionen aus, welche zeitlich

unterscheidbar sind (im Gegensatz zu den Schallenergien im diffusen Nachhall). In dieser Phase liegen die Anfangsreflexionen (Early Reflektions). Etwa 80% der einfallenden Schallenergie gelangt in dieser Phase an den Hörerort. Ihre Richtung, Betrag, Klangfärbung und zeitliche Verteilung gibt dem Hörer Informationen über den Raum, sowie seiner Beschaffenheit. Sie sind maßgeblich an der Unterscheidung zwischen „guten“ und „schlechten“ Zuhörer-Plätzen beteiligt (vgl. Reichardt 1974 zitiert von Ahnert und Reichardt 1981). Auch *Beranek (1996)* gibt diesem Zeitbereich eine große Bedeutung. Er bezeichnet diese Phase als *Early Sound* (welche aber auch den direkten Anteil einschließt, im Gegensatz zu den Early Reflections). Sie gibt dem Hörer ein Gefühl einer „breiteren“ Quelle, oder vergrößerten ASW (Apparent Source Width (siehe Kap.: 1.1.5)). *Beranek (1996)* gibt für diese Phase einen Zeitrahmen von 80ms vor, der aber nur bedingt allgemein gültig ist. So kann z.B. bei Badezimmern schon viel früher von diffusem Nachhall gesprochen werden, als bei großen Konzerthallen, da es durch die kleineren Weglängen schon viel früher zu einer Durchmischung der Schallenergien kommt. Aber die zeitliche Abgrenzung von 80ms dürfte allgemein anerkannt sein, wie man an den raumakustischen Parametern wie Klarheitsmaß  $C_{80}$ , Seitenschallgrad LE, LF und Raumeindrucksmaß R erkennt, welche sich alle an 80ms als Integrationsgrenzen orientieren.

*Griesinger (2007)* betont, dass Early Reflections die Klarheit und ASW auch negativ beeinträchtigen können. Sind sie z.B. zu laut, überlagern sie sich untereinander und mit dem Direktschall, sodass hörbare Phasenverzerrungen entstehen. Dieser Effekt ist auch in Badezimmern hörbar. Das häufig verwendete Material Fliese hat nur einen kleinen Dämpfungsfaktor, wodurch viel Energie reflektiert wird und bedingt durch die kleinen geometrischen Abmessungen in diesen Räumen haben die frühen Reflexionen eine hohe zeitliche Dichte und überlagern sich stark, sodass hörbare Kammfilter entstehen.

Sollte ein Raum zu starke Early Reflections besitzen, schlägt *Griesinger (2007)* vor, die Materialien, welche die Reflexionen lenken, mit Absorbern auszustatten, welche Frequenzen ab 700Hz dämpfen. Dadurch ist es einfacher, den Direktschall von einfallenden Reflexionen zu unterscheiden.

### 2.2.2 Simulation der Early Reflections

Der einfachste Weg zur Simulation von Reflexionen ist über das geometrische Modell eines Raumes. Trifft ein Schallstrahl auf eine Oberfläche, wird er zu einem gewissen Teil reflektiert, wie bei einem Spiegel. Wie in Abbildung 4 ersichtlich, erscheint die

Schallquelle „hinter“ der Wand, sie wird auf die andere Seite der Wand gespiegelt. Im vereinfachten Fall eines quadratischen Raumes in 2D muss die Quelle für jede Wand gespiegelt werden, das ergibt vier Spiegelungen (Abbildung 5). Es entsteht ein Raum

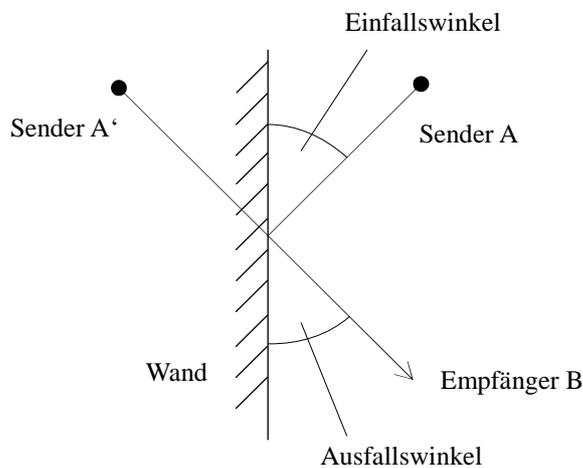


Abbildung 4: Reflektion eines Schallstrahls

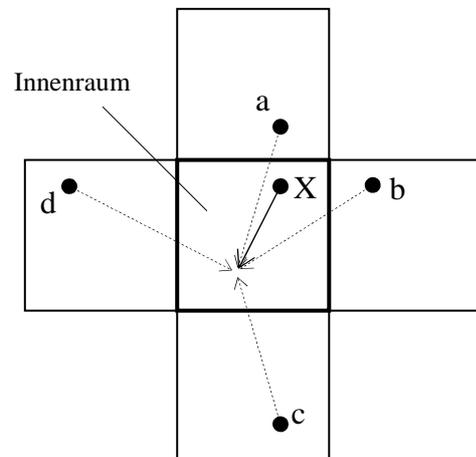


Abbildung 5: Prinzip der Spiegelquellen<sup>1</sup>

ohne Wände, dafür mit mehreren Quellen, und es ist nun einfach, die Einfallswinkel und Längen der Reflektionen zu berechnen. Diese Methode führt zu den Reflektionen erster Ordnung, in der Realität müssen aber noch jene berücksichtigt werden, welche zuerst mehrmals die Wand treffen bevor sie am Hörer eintreffen. Dafür muss das Spiegelquellenprinzip erweitert werden, indem man die Quellen nochmals spiegelt. Die Anzahl der Reflektionen ist nach der Formel  $N^k$  abhängig von der Anzahl der Wände  $N$  und der Anzahl  $k$ , wie oft die Schallstrahlen bereits eine Wand berührt haben. Da sich die Anzahl der Spiegelquellen exponentiell erhöht, ist diese Methode nur für kleinere Ordnungen von  $k$  sinnvoll (vgl. Gardner 1998). Meist reicht es aus, nur die ersten beiden Ordnungen zu berechnen um dann bereits in den Nachhall überzugehen (vgl. Musil et al. 2007) für den es effizientere Berechnungsmethoden gibt (siehe Kap. 2.3.2). In einer Simulation für die Early Reflections berechnet man anhand der Spiegelquellen die Einfallrichtung und die Länge, über die man anhand der Schallgeschwindigkeit die Zeitdifferenz bis zum Eintreffen am Empfänger berechnen kann. Um die

<sup>1</sup> Der Innenraum samt Quelle X wird an jeder Wandseite gespiegelt, dadurch ergeben sich vier Spiegelquellen a,b,c,d. Länge und Richtung dieser vier Reflektionen sind vorgegeben durch den jeweils kürzesten Weg zum Empfänger

Wandabsorption und die Dämpfung durch das Ausbreitungsmedium (Dissipation siehe Kap. 1.2.5) zu berücksichtigen, muss jede Reflektion entsprechend gefiltert werden (vgl. Lehnert und Blauert 1992 zitiert von Gardner 1998). Die Parameter für diese Filter werden einerseits über die frequenzabhängigen Absorptionskoeffizienten der verwendeten Materialien, andererseits über die Länge des Schallstrahls berechnet.

## 2.3 Einfluss des diffusen Nachhalls

Der diffuse Nachhall zeichnet sich dadurch aus, dass die Schallenergie von sehr vielen Richtungen mit einer hohen Echodichte am Hörer auftritt, sodass das Ohr einzelne Klänge nicht mehr geometrisch im Raum zuordnen kann. Die verwendete Methode für das Testprogramm richtet sich am Nachhall nach *Jot und Chaigne (1991)*, welche mit dem FDN (Feedback Delay Network) das Basisprinzip von *Schroeder (1962)* weiterentwickelten.

### 2.3.1 Nachhall nach Schroeder

Der „Late Reverberator“ von *Schroeder (1962)* ist der Ausgangspunkt für das Nachhallprinzip von FDN (Feedback Delay Network). Wesentliche Elemente sind eine Kombination von Kammfilter (Abb. 6) und Allpassfilter (Abb. 7).

Kammfilter

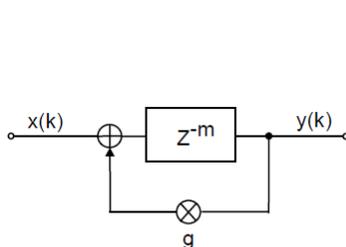


Abbildung 6: Blockschaltbild Kammfilter

Allpassfilter

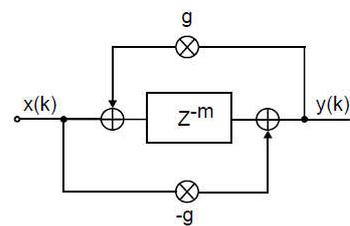
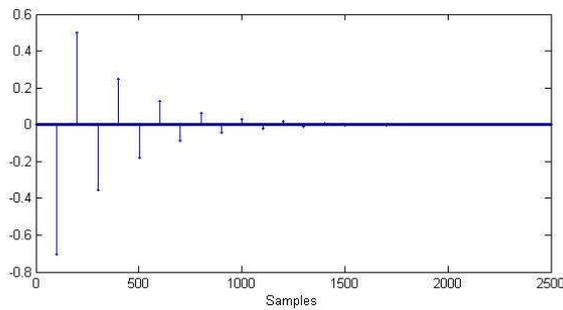


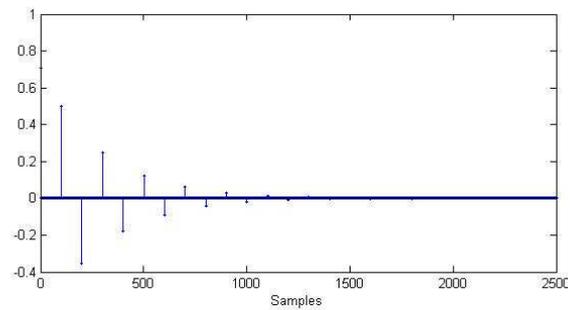
Abbildung 7: Blockschaltbild Allpassfilter

Kammfilter

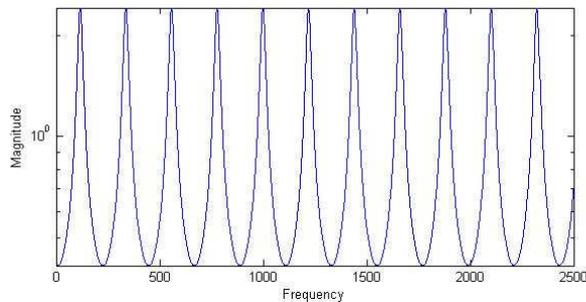


**Abbildung 8: Impulsantwort eines Kammfilters mit  $m=100$  und  $g=-0.707$**

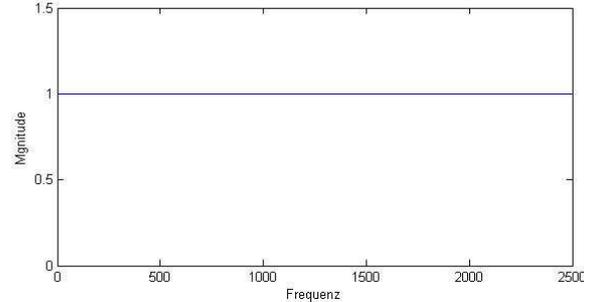
Allpassfilter



**Abbildung 9: Impulsantwort eines Allpassfilters mit  $m=100$  und  $g=-0.707$**



**Abbildung 10: Spektrum eines Kammfilters mit  $m=100$  und  $g=-0.707$**



**Abbildung 11: Spektrum eines Allpassfilters mit  $m=100$  und  $g=-0.707$**

Übertragungsfunktion Kammfilter

$$H(z) = \frac{g}{z^{-m} - g} \quad (25)$$

Übertragungsfunktion Allpassfilter

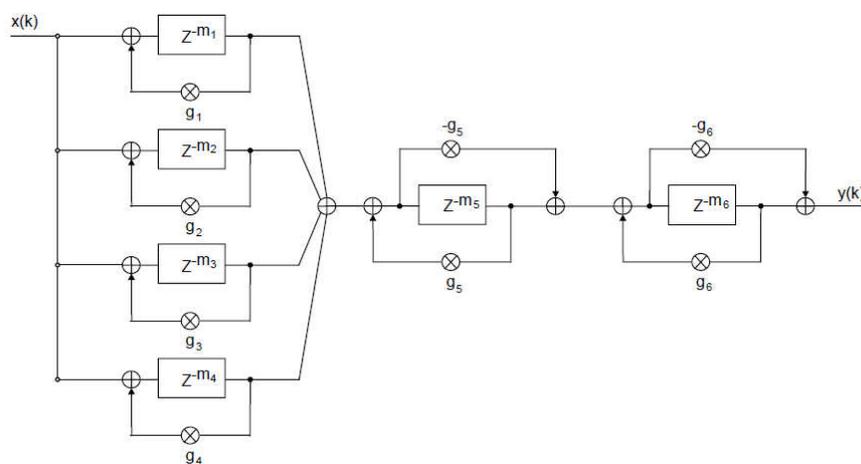
$$H(z) = \frac{z^{-m} - g}{1 - g \cdot z^{-m}} \quad (26)$$

In der Zeitebene ist sowohl beim Kammfilter (Abb. 8), als auch beim Allpassfilter (Abb. 9) eine abklingende Impulsfolge erkennbar, die Impulse treten mit einem Zeitunterschied von  $m$  auf. Durchläuft ein Signal dieses Filter, wird es bei jedem Impuls verzögert und mit der Rückkopplungsverstärkung  $g$  gewichtet, was einem abklingenden Echo nahekommt.

Allpassfilter haben ein konstantes Betragsspektrum (Abb. 11), da die Null und Polstellen paarweise gespiegelt am Einheitskreis auftreten. Sie eignen sich daher, Signale zu verzögern, ohne die Klangfarbe zu verändern.

Schroeder kombinierte beide Filter um einerseits eine hohe Echodichte (Anzahl der

Reflexionen pro Zeit) mittels Serienschaltung von mehreren Allpässen und andererseits Raummoden mittels Parallelschaltung von Kammfiltern zu erzeugen (Abbildung 12). Mit den Allpässen kann eine hohe zeitliche Dichte erreicht werden, welche realen Räumen sehr nahe kommt. Laut *Schroeder* sind mindestens 1000 Reflexionen pro Sekunde nötig, um nicht einzelne Echos zu hören, laut *Cremer und Müller (1976)* 2000 Reflexionen, laut *Beranek (1996)* und *Jot und Chaigne (1992)* 10000 Reflexionen, sodass auch schnelle Transienten, z.B. hervorgerufen von kleinen Trommeln, realistisch klingen. Die Parallelschaltung der Kammfilter erzeugt in Summe eine hohe Anzahl an Spitzen im Frequenzspektrum. Die auftretende Klangverfärbung wird über die *Frequenzdichte* beschrieben.



**Abbildung 12: Nachhall nach Schroeder**

**Quelle: Extra et al. (2006)**

## Frequenzdichte

Die Kammfilter erzeugen starke spektrale Verfärbungen des Klages, was durch hervorstechen einzelner Töne im Gesamtklang hörbar ist. Deshalb müssen die Kammfilter gewichtet werden um eine konstante *Modale Dichte* oder *Frequenzdichte*, also gleich viele Moden pro Herz, sicherzustellen. Laut *Schroeder* ergibt eine unterschiedliche Gewichtung der Rückkopplungs-Verstärkung  $g_p$  der Kammfilter mit der jeweiligen Verzögerungszeit  $m_p$  einen konstantes Abklingen der Nachhallzeit bei allen Frequenzen.

$$\gamma = g_p \frac{1}{m_p} = const. \quad (27)$$

Für ein konstantes  $\gamma$  für alle Kammfilter tritt minimale Klangverfärbungen auf.

### Zeitliche Dichte

Durch geeignete Einstellung der Kamm- und Allpassfilter kann ein perfekt exponential abfallender Nachhallverlauf erreicht werden. Nach *Schroeder* besteht ein Zusammenhang von Frequenzdichte und zeitlicher Dichte:

$$\text{Frequenzdichte:} \quad D_f = \sum_{p=0}^{P-1} \tau_p \approx P \tau \quad (28)$$

$$\text{Zeitliche Dichte:} \quad D_t = \sum_{p=0}^{P-1} \frac{1}{\tau_p} \approx P / \tau \quad (29)$$

$$\text{daraus resultierend:} \quad P \approx \sqrt{D_f D_t} \quad (30)$$

$$\tau \approx \sqrt{D_f / D_t} \quad (31)$$

mit  $D_f$ ...Frequenz Dichte

$D_t$ ...Zeitliche Dichte

$P$ ...Anzahl der Kammfilter

$\tau_p$ ...Verzögerung in Sekunden des Kammfilters  $P$

$\tau$ ... Durchschnittliche Verzögerung aller Kammfilter

Für 1000 Echos pro Sekunde und 0.15 Raummoden pro Hertz benötigt man laut *Schroeder* 12 Kammfilter mit einer durchschnittlichen Verzögerung von 12ms. Laut *Jot und Chaigne (1991)* sind 1000 Echos pro Sekunde unzureichend, für einen realen Nachhall sind ungefähr 10000 Echos pro Sekunde nötig, was laut Formel (30) 40 Kammfilter erfordern würde. Bei Verwendung von weniger Filtern  $P$  verkürzen sich laut Formel (29) bei unveränderter zeitlicher Dichte  $D_t$  auch die durchschnittlichen Verzögerungszeiten  $\tau$  der Kammfilter und dadurch werden Klangverfärbungen hörbar, weil sich laut Formel (28) die Frequenzdichte  $D_f$  verkleinert.

### Einstellen der Nachhallzeit

Durch Ausdrücken der Rückkopplungsverstärkung  $g_p$  und  $\gamma$  in logarithmischen Werten kann die Nachhallzeit  $RT_{60}$  berechnet werden. Gleichung (27) in logarithmischer Darstellung mit  $\Gamma=20\log(\gamma)$  und  $G_p=20\log(g_p)$  ergibt:

$$\Gamma = G_p / m_p \quad (32)$$

Laut *Schröder* haben alle Eigenmoden die selbe Abklingzeit,  $\Gamma$  gibt die Rückkopplungsverstärkung an. Dadurch kann die Nachhallzeit  $RT_{60}$  berechnet werden als

$$RT_{60} = -60 \frac{\text{Samplingrate}}{\Gamma} = -60 \frac{\tau_p}{G_p} \quad (33)$$

Die Konstante „-60“ bezieht sich auf die Festlegung der Nachhallzeit  $RT_{60}$  als verstrichene Zeit bei einem Abklingen der Raumantwort um 60dB

Um die Nachhallzeit  $RT_{60}$  zu verlängern, müsste man die Verzögerungszeiten  $\tau_p$  erhöhen oder die Rückkopplungsverstärkung  $G_w$  verkleinern.

### Frequenzabhängige Nachhallzeit

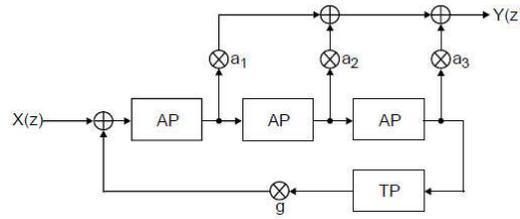
Der oben genannte Nachhall erfüllt nun die Vorgabe, dass er über das gesamte Spektrum eine konstante Abklingzeit besitzt. In der Realität ist aber die Nachhallzeit von Materialien abhängig, welche keineswegs konstante Absorption in Abhängigkeit der Frequenz haben. Meist werden tiefere Frequenzen weniger bedämpft, was zu einer längeren Nachhallzeit für tiefere Frequenzen führt.

Durch Ersetzen der Rückkopplungsverstärkung  $g_p$  der Kammfilter durch ein frequenzabhängige Funktion  $g_p(\omega) = |h_p(e^{j\omega})|$  in Gleichung (27), kann dieser Effekt erreicht werden:

$$\gamma(\omega) = |h_p(e^{j\omega})|^{\frac{1}{m_p}} \quad (34)$$

Um die Luftabsorption zu simulieren, schaltete (vgl. *Moorer 1979 zitiert von Extra et al. 2006*) einen Tiefpass erster Ordnung in die Rückkopplung. Dessen Resonanzfrequenz ist von der Distanz zwischen Schallquelle und Hörer abhängig, und wird über die Verzögerung der Kammfilter berechnet, welche für größere Räume, also für größere Distanzen wächst.

Um eine höhere zeitliche Dichte zu erreichen, kann man die Allpässe anders verschalten (vgl. *Gardner 1992 zitiert von Extra et al. 2006*), wie in Abbildung 13 zu sehen ist.



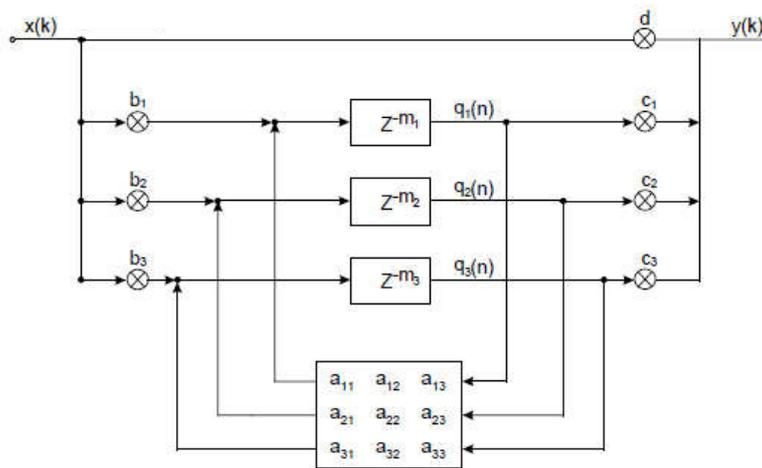
**Abbildung 13: Serienschaltung von Allpässen nach Gardner**

Quelle: Extra et al. 2006

Die zusätzliche Rückkopplung erhöht die zeitliche Dichte mit zunehmender Zeit in Abhängigkeit des Faktors  $g$  und dem Tiefpass TP. Dieser Tiefpass simuliert die Luftdissipation (siehe Kap.: 1.2.4) und mit steigender Raumgröße werden tiefere Grenzfrequenzen eingestellt. Mit wenig zusätzlichem Rechenaufwand erhöht sich die zeitliche Dichte des Nachhalls. Den Ausgang bildet eine Addition der einzelnen gewichteten Ausgänge der Allpässe, wodurch die zeitliche Dichte nochmals erhöht wird.

### 2.3.2 Nachhall nach dem Prinzip des Feedback Delay Network FDN

*Jot und Chaigne (1991)* entwickelten die erwähnten Modelle weiter, indem sie nicht mehr die Wege einzelne Rückkopplungen der Kammfilterparallelschaltung verfolgen, sondern versuchten, alle Rückkopplungen miteinander zu „vermischen“. Dieser Ansatz wird auch FDN (Feedback Delay Network) genannt, Abbildung 14 zeigt einen prinzipiellen Aufbau.



**Abbildung 14: FDN nach Jot-Chaigne**

Quelle: Extra et al. 2006

Die erwähnte Vermischung aller Rückwege findet in der Matrix A im Rückkopplungszweig statt. Nun kann zur Rückkopplung des ersten Kammfilters auch ein Teil der Schallenergie von den anderen Filtern addiert werden. Für ein Netzwerk mit vier Rückwegen geben *Jot und Chaigne (1991)* folgende Matrix als Beispiel an:

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix} \quad (35)$$

Damit das System stabil ist müssen die Faktoren in den Rückkopplungen kleiner eins sein. Dies trifft für den Fall zu, wenn die Matrix A aus dem Produkt einer unitären Matrix ( $U^T = U^{-1}$  bzw.  $\|U \cdot x\| = \|x\|$ , wenn x ein beliebiger Vektor und U die unitäre Matrix ist) und einer Diagonal-Matrix mit Werten kleiner eins besteht. Gilt der Spezialfall, dass die Matrix A selbst auch unitär ist, sind alle Pole des Systems auf dem Einheitskreis, es findet keine Abschwächung statt, die gesamte Energie bleibt im System, die Eigenmoden klingen nicht mehr ab. Für die genannte Matrix A aus Formel (34) ist dies der Fall.

Maximales Vermischen findet statt, wenn alle Werte der Matrix A ungleich Null sind. Besteht A aus einer Diagonalmatrix, hängt jeder Ausgang der Matrix nur von dem jeweiligen gewichteten Eingang ab, was der Parallelschaltung der Kammfilter nach *Schroeder* entspricht.

Betrachtet man die Serienschaltung der Allpässe von *Schroeder* näher, kann man es auch als ein Netzwerk von Allpässen mit einer triangularen Matrix ( $a_{ij} = 0$  für  $i > j$ ) im Rückkopplungszweig auffassen. Dadurch kann man dieses Netzwerk ebenfalls als Quelle von Eigenmoden mit Resonanzfrequenzen und Abklingkurven bezeichnen, wofür eigentlich die Kammfilter gedacht sind.

Die eben erwähnten Modifikationen erhöhen die zeitliche Dichte immens, um dem Raum noch eine Klangfärbung zu geben, werden zusätzlich in den Rückführungszweig Filter geschaltet. Für möglichst reale Nachbildung werden Filter verwendet, welche eine unabhängige Einstellmöglichkeit für Bässe, Mitten und Höhen geben.

### 3 Auswahl der zu übenden räumlichen Attribute

In Kapitel 1 sind einige subjektive Begriffe und objektive Messgrößen vorgestellt worden. Wie kann nun ein Testprogramm aussehen, welche Attribute sollen getestet

werden?

Auf „Golden Ears“ von *Moulton* (1995) ist die Nachhalldauer zu beurteilen. Als Quellmaterial wurden sehr trocken aufgenommene Musikbeispiele und ein Schlagzeug verwendet. Nach Zugabe eines künstlichen Nachhalls muss der Proband die Länge des Nachhalls und das „Predelay“<sup>1</sup> abschätzen, wobei beide Parameter jedesmal geändert werden. Der Nachhall wurde mit einem Lexicon NuVerb mit einem modifizierten Preset „Medium Hall“ erzeugt.

*Berg* (2005) erstellte ein Testprogramm namens OPAQUE, welches folgende Eigenschaften variiert:

- Raumgröße
- Distanz zur Quelle
- Präsenz (siehe Kap.: 1.1.7)
- akustische Breite der Quelle<sup>2</sup>
- Brillianz

Das Programm OPAQUE spielte das selbe Audiomaterial mit verschiedenen Einstellungen für den zu simulierenden Raum ab. Der Proband musste dann diese Beispiele ordnen, z.B. nach Distanz zur Quelle, wobei immer nur ein Parameter geändert wurde (z.B. blieb die Raumgröße konstant, wenn die Brillianz verändert wurde).

Basierend auf diesen Arbeiten wurde entschieden, folgende zu prüfende Attribute in das Trainingsprogramm einfließen zu lassen:

- Distanz zur Quelle
- Raumgröße
- Raumklang (Wärme, Brillianz)

### 3.1 Distanz zur Quelle

Die Wahrnehmung der Distanz zwischen Schallquelle und Hörer kann über mehrere physikalische Parameter beschrieben werden, die bedeutendsten sind (Nielsen 1991 zitiert von Neher, Rumsey, Brooks 2002):

---

<sup>1</sup> Verstrichene Zeit zwischen Direktschall und der ersten Reflektionen am Hörerort

<sup>2</sup> Mit steigender Abmessung der Quelle, steigt auch die akustische Breite, eine Punktquelle hat die kleinste akustische Breite.

- Lautstärke
- Verhältnis von direkter zu diffus einfallender Schallenergie
- Struktur der Frühen Reflektionen
- Klangfärbung

### Lautstärke

Mit steigender Distanz zur Schallquelle nimmt in schalltoten<sup>1</sup> Räumen der Schallpegel um 6dB pro Verdopplung der Distanz ab,

### Verhältnis von direkter zu diffus einfallender Schallenergie

Wenn der Raum Reflektionen zulässt (praktisch immer der Fall), addiert sich die reflektierte zu der direkt empfangenen Schallenergie (siehe Abbildung 2). Das Verhältnis dieser Energie zueinander sinkt mit steigender Distanz zur Quelle.

### Struktur der frühen Reflektionen (Early Reflections)

Der Einfluss der Early Reflections (siehe Kap.: 2.2) auf die Distanzwahrnehmung ist noch umstritten (Neher et al. 2002). Einerseits gibt es Hinweise, dass die „exakte räumliche und zeitliche Abbildung“ dieser Reflektionen die Distanzwahrnehmung dominiert (Kendall et al. 1984 zitiert von Neher et al. 2002), andererseits ist ihr Einfluss laut *Neher (2002)* eher gering, da auch auf Mono-Aufnahmen eine gute Abschätzung der Distanz möglich ist.

### Klangfärbung

Für große Distanzen ( $> 15\text{m}$ ) tritt für das Direktsignal eine Abschwächung der Höhen ein (siehe Kap.: 1.2.4), für kleinere Distanzen ( $< 25\text{cm}$ ) ist ein Anstieg der Bässe bemerkbar (Blauert 1997 zitiert von Neher et al. 2002).

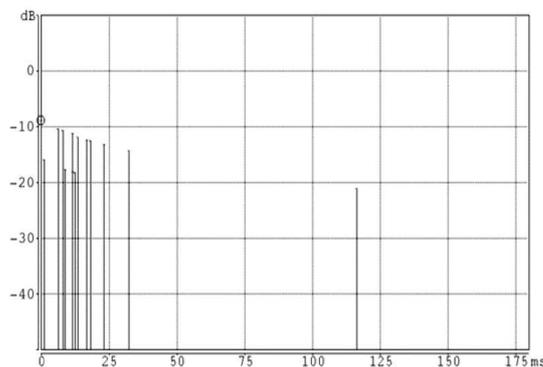
## 3.2 Raumgröße

Die Raumgröße wird vom Ohr zum großen Teil von der zeitlichen Auflösung der Early Reflections bestimmt (siehe auch Kap. 2.2). Umso größer der Raum ist, umso länger

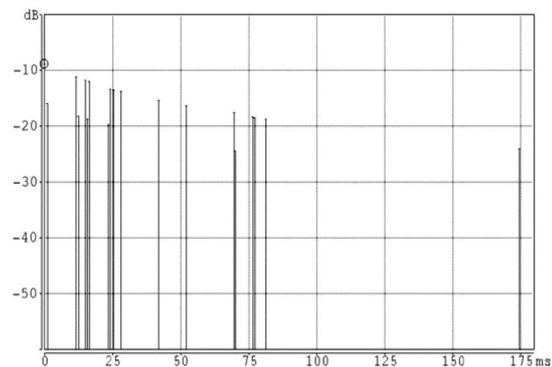
---

<sup>1</sup> Ein Raum mit Absorptionsgrad  $\alpha=1$  absorbiert die komplett einfallende Schallenergie, es kommt zu keinen Reflektionen und nur der Direktschall gelangt zum Hörer.

sind die Distanzen zu den reflektierenden Wänden und umso zeitlich verzögerter erreichen die Reflektionen den Hörer. Bedingt durch diese steigenden Verzögerungen steigt auch die Nachhallzeit, die Dissipation der Luft nimmt zu, infolgedessen wird der Raumklang minimal *wärmer* (siehe Kap. 1.1.2). Abbildung 15 und Abbildung 16 zeigen einen Vergleich von berechneten Early Reflections bis zur 2. Ordnung für rechteckigen Räume mit  $1200\text{m}^3$  (Abb. 14) und  $3780\text{m}^3$  (Abb. 15), wobei nur die Raumgröße verändert wurde. Der Direktschall ist auf dem Punkt 0s der Zeitachse eingetragen.



**Abbildung 15: zeitliche Verteilung der Reflektionen bis zur 2. Ordnung für eine Raumgröße von  $1200\text{m}^3$**   
Quelle: Neher 2004



**Abbildung 16: zeitliche Verteilung der Reflektionen bis zur 2. Ordnung für eine Raumgröße von  $3780\text{m}^3$**   
Quelle: Neher 2004

### 3.3 Raumklang

Der Raumklang kann gut über die frequenzabhängige Nachhallzeit (siehe Kap.: 1.2.1) beschrieben werden. Die subjektiven Begriffe „Wärme“ und „Brillanz“ (siehe Kap. 1.1) sind direkt abhängig von den Nachhallzeiten für tiefe (75 – 350Hz) und hohe (> 1400Hz) Frequenzen. Durch Variation dieser Parameter kann der Raumklang verändert werden.

## 4 Versuchsaufbau

### 4.1 Abhörposition

Um so flexibel wie möglich zu sein, wurde das Stereo-Format mit  $60^\circ$  Öffnungswinkel als Wiedergabesetup gewählt, wie auch von *Moulton* (1995). *Neher* (2004) wählte Stereo + Center mit  $60^\circ$  Öffnungswinkel, merkte aber an, dass ein größerer

Öffnungswinkel den Vorteil hätte, ein „breiteres“ Abbild des Raumes zu ermöglichen. Aufgrund von Kompatibilität und Flexibilität entschied er sich dennoch für die Standard – Stereo Aufstellung. *Berg (2004)* gibt über die verwendete Lautsprecheranordnung keine Auskunft, vermutlich wurde ebenfalls eine Stereo - Anordnung gewählt.

## 4.2 Aufbau des Hallmoduls in PD

Das Hallmodul ist in der Entwicklungsumgebung „Pure Data PD“ realisiert, welche von *Miller S. Puckette* entwickelt wurde. Ähnlich wie MAX/MSP ist es eine echtzeitfähige Programmierumgebung welche ursprünglich für Audio-Anwendungen optimiert wurde. Für Pure Data sind mehrere Hallmodule erhältlich, als Standard wird „Freeverb“ genannt, welches auf dem Algorithmus von *Moorer* (vgl. Puckette 2010) basiert, aber nur ungenügende Einstellmöglichkeiten bietet. Es ist nicht möglich, Parameter für Wärme des Nachhalls, Position im Raum, usw. einzustellen. Um umfassende Tests durchzuführen, ist es nötig, auch diese Parameter einzeln steuern zu können. Daher wurde auf ein Nachhallmodul zurückgegriffen, welches am IEM (vgl. Musil et al. 2007) im Zuge einer Arbeit über Binaurale Raumsimulation entwickelt wurde. Es beruht auf den Arbeiten von *Jot und Chaigne (1991)* über einen erweiterten „Feedback Delay Network (FDN) Reverberator“ (siehe Kap.: 3.3.1).

Die Funktionen des resultierenden Nachhallmoduls sind an die Anforderungen an das Trainingsprogramm angeglichen und beschränken sich auf folgenden Rahmen:

- Berechnung für zwei Dimensionen, Decken- und Bodenreflektionen bleiben unberücksichtigt
- Ordnung der Early Reflections wird auf zwei begrenzt (vgl. Musil et al. 2007)

### 4.2.1 Simulation des direkten Schallanteils in PD

Der Einfallswinkel und die Distanz zwischen Hörer und Schallquelle werden über die jeweiligen Positionen im Raum berechnet. Die ermittelte Distanz führt über die Schallgeschwindigkeit zu der zeitlichen Verzögerung, welche über ein Delay realisiert ist. Für die räumliche Abbildung wird das Signal auf die Stereokanäle mittels *Panning* (vgl. Bauer 1961 zitiert von Pulkki und Karjalainen 2001) nach Formel (36) aufgeteilt:

$$\frac{\tan(\theta_s)}{\tan(\theta_0)} = \frac{g_1 - g_2}{g_1 + g_2} \quad (36)$$

$$g_1^2 + g_2^2 = 1 \quad (37)$$

mit  $\theta_s$  als abzubildende virtuelle Quelle,  $\theta_0$  als Basiswinkel<sup>1</sup> der Stereoanordnung und  $g_1$ ,  $g_2$  als Multiplikator des Quellsignals für den jeweiligen Kanal. Formel (37) sichert, dass die Energie nur aufgeteilt wird und in Summe die selbe bleibt.

#### 4.2.2 Simulation von Early Reflections in PD

Über die Position des Hörers, der Schallquelle und den Abmessungen des Raumes können die Early Reflections über das Spiegelquellenprinzip berechnet werden. Die jeweilige Verzögerung ist über Delays realisiert, bei jeder Wandberührung wird die Absorption berücksichtigt. Die Schall-Energien werden wie beim direkten Schallanteil (siehe Kap.: 4.2.1) auf beide Stereo-Kanäle aufgeteilt.

#### 4.2.3 Einstellungen des Feedback Delay Network in PD

Das verwendete Nachhall-Modul besteht aus einem FDN nach *Jot und Chaigne (1991)* mit 12 Kammfiltern. Die Verzögerungszeiten werden über die Raumgröße errechnet und dann nichtlinear<sup>2</sup> auf die Kammfilter aufgeteilt. Auf jeder der 12 Rückkopplungen der Kammfilter befindet sich ein 3-Band-Equalizer für die Simulation der Absorbereigenschaften des Raumes mit 450Hz als Mittenfrequenz für die Bässe und 3500Hz für die Höhen. Die Rückkopplungsmatrix A ist unitär (verlustlos), deshalb wird die Nachhalllänge nur durch die Absorption der Equalizer bestimmt.

#### 4.2.4 Steuerung des Hallmoduls

Die Steuerung des Moduls findet über das Programm MATLAB<sup>®</sup> von MathWorks<sup>®</sup> statt. Als Eingabeparameter sind nötig:

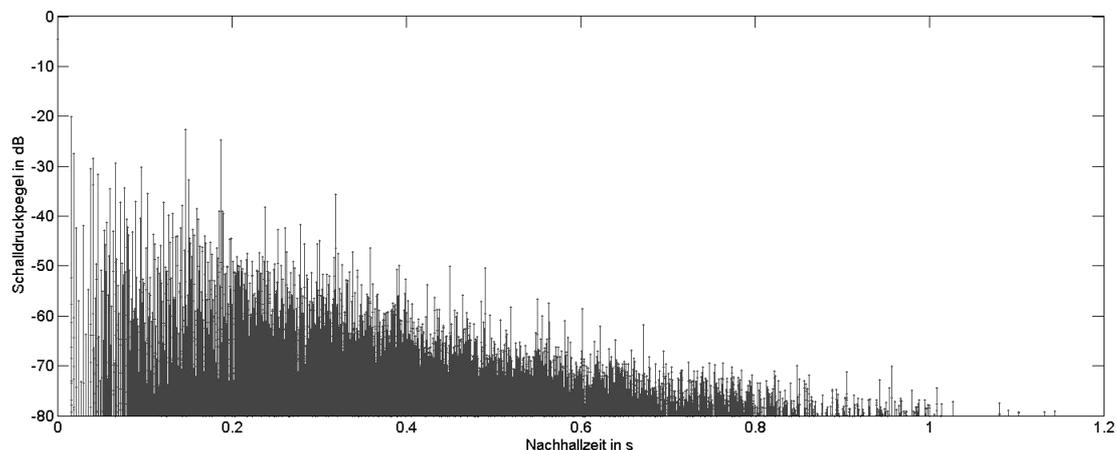
- Raumgröße (Länge x Breite)

<sup>1</sup> Einer Stereoanordnung mit 60° Öffnungswinkel hat einen Basiswinkel von 30°

<sup>2</sup> Laut Musil (2010) klingt der Nachhall bei linearer Aufteilung der Verzögerungen auf die Kammfilter unnatürlich, das Aufteilen nach einer exponentiellen Funktion löste das Problem.

- Position Schallquelle (Position X, Y)
- Position Hörer (Position X, Y)
- Raumabsorption in drei Frequenzbändern
- Schallgeschwindigkeit  $c$
- Pegel des Direktanteils
- Pegel der Early Reflections
- Pegel des diffusen Nachhalls

Über diese Parameter kann das Hallmodul alle Zwecke des Trainingsprogramms erfüllen. Zusätzlich bietet es die Möglichkeit, Raumimpulsantworten zu erstellen und zu speichern. Mithilfe eines geeigneten Auswerteprogramms (wie z.B. MATLAB<sup>®</sup>), können aus diesen Impulsantworten Informationen gewonnen werden, mit welchen man die Eigenschaften des Halls sehr gut beschreiben kann. Abbildung 17 zeigt einen Nachhallverlauf des Hallmoduls: Der zu simulierende Raum hat die Abmessungen 25 mal 25m, die Distanz zwischen Schallquelle und Empfänger beträgt 17m und ihre Achse ist nicht parallel zu den Wänden. Der Direktanteil des Schalls ist auf der x-Achse bei Null eingezeichnet. Gut erkennbar ist die steigende Anzahl der Reflektionen mit steigender Nachhallzeit.



**Abbildung 17: Nachhallverlauf des verwendeten Hallmoduls**

Aus diesem Nachhallverlauf (Abb. 17) wurden folgende objektive Messgrößen berechnet:

Nachhallzeit $RT_{60}$	~0.9 s	Kap.: 1.2.1
Bassverhältnis BR	1.3	Kap.: 1.2.3

Deutlichkeitsmaß $C_{50}$	1.38 dB	Kap.: 1.2.7
Schwerpunktszeit $t_s$	169 ms	Kap.: 1.2.8
Hallmaß H	-1.38 dB	Kap.: 1.2.9
Klarheitsmaß $C_{80}$	2.2 dB	Kap.: 1.2.12

Auffällig ist die hohe Schwerpunktszeit, die Begründung liegt in den langen Wegen der Reflektionen durch die Größe des Raumes.

## 5 Ablauf des Trainingsprogramms und erste Ergebnisse

Die Benutzerschnittstelle ist über MATLAB<sup>®</sup> von MathWorks<sup>®</sup> realisiert. Nach Aufruf des Programms kann die Testperson wählen zwischen Training der „Wahrnehmung der Distanz“, „Wahrnehmung der Raumgröße“ oder „Klang des Nachhalls“.

Die Testperson hat die Auswahl zwischen zwei Klangbeispielen, welche sich nur durch die Art des Nachhalls unterscheiden. Es wird nur ein Parameter des Nachalls geändert, z.B. wird bei „Wahrnehmung der Distanz“ nur die Distanz zwischen Hörer und Schallquelle geändert. Die Beispiele können beliebig oft abgespielt werden. Die Funktion „Antwort zeigen“ markiert die richtige Antwort, ohne dass ein Fehler gewertet wird. Diese Funktionen ist sinnvoll, wenn die Testperson noch keine Erfahrung mit diesem Trainingsprogramm hat und sich daher zu Beginn des Trainings ein Bild über die Größenordnung der klanglichen Unterschiede der Beispiele machen möchte.

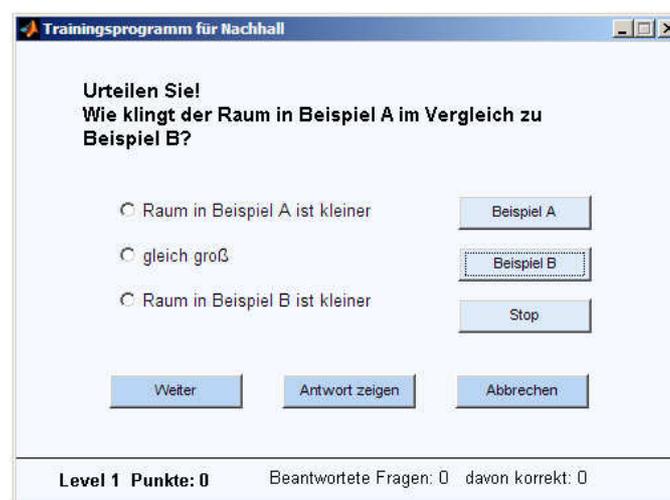


Abbildung 18: Oberfläche der Benutzerschnittstelle

Um einen Trainingseffekt zu erzielen, wird der Schwierigkeitsgrad sukzessive erhöht.

Pro richtiger Antwort bekommt man einen Punkt, pro falscher wird einer abgezogen. Sobald fünf Punkte erreicht werden, wird auf den nächsten Schwierigkeitsgrad bzw. Level erhöht und der Punktezähler zurück auf Null gestellt. Ist einmal ein Level erreicht, kann man nicht zurückfallen, ebenso sinken die Punkte für falsche Antworten nicht unter Null.

Die Änderungen, welche für einen steigenden Schwierigkeitsgrad vorgenommen werden, wurden zunächst so eingestellt, dass bis Level drei die gestellten Aufgaben (nach Meinung des Autors) mehr oder weniger bewältigbar sind, dann aber zunehmend eine Herausforderung darstellen.

## 5.1 Einstellungen des Trainingsprogramms

### Wahrnehmung der Distanz

Die Größe des Raumes beträgt 10m x 20m, Schallquelle und Hörer stehen sich in einer Flucht parallel zu den Wänden gegenüber. Abbildung 18 zeigt eine Übersicht: Der Hörer befindet sich am Punkt 5m in Richtung der Raumbreite und 2m in Richtung der Raumtiefe und ist vom Referenzklang 6m entfernt. Die Versuchsperson muss nun anhand des Klangs entscheiden, ob sich der Ort der Schallquelle und damit die Distanz geändert hat. Die Person hat einerseits ein Audiobeispiel mit der Schallquelle an der Referenzposition und andererseits das selbe Audiobeispiel, abgespielt an einer anderen Position im virtuellen Raum, zur Verfügung. Der Zufall entscheidet, ob sich die Distanz vergrößert, verkleinert, oder die selbe geblieben ist.

Für den Fall, dass sich die Distanz ändert, wird diese bei Level 1 im Vergleich zur Referenzdistanz entweder verdoppelt oder halbiert, je nach Zufall, und mit steigendem Level werden die Distanzunterschiede kleiner. In Abbildung 19 sind die Distanzen bis Level 5 als Kreuze eingezeichnet.

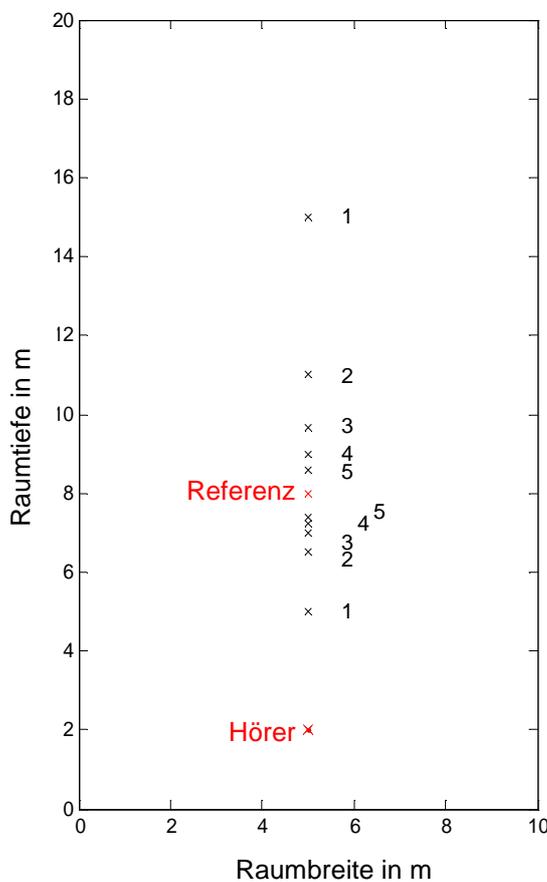
Bei diesem Test wird nur die Position der Schallquelle geändert, alle anderen Parameter bleiben konstant. Die Berechnung der neuen Position der Schallquelle wird nach Formel (38) für steigende und nach Formel (39) für sinkende Distanzen in Abhängigkeit des Level berechnet:

$$\text{Distanz}_{\text{neu, kleiner}} = \left(1 - \frac{1}{2 \cdot \text{Level}}\right) \cdot 6 \quad (38) \quad \text{Distanz}_{\text{neu, größer}} = \left(1 + \frac{1}{2 \cdot \text{Level}}\right) \cdot 6 \quad (39)$$

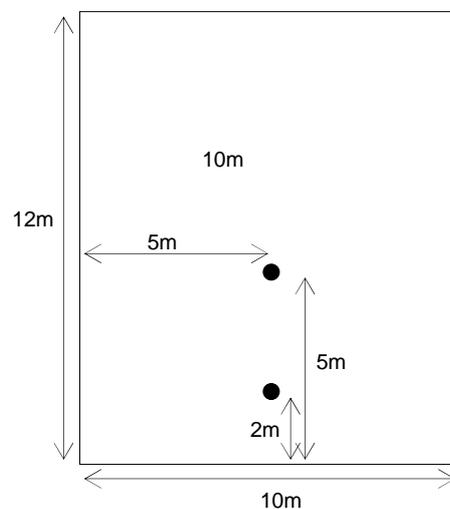
Diese Formeln sind empirisch entstanden, und wurden so angepasst, dass sich der Schwierigkeitsgrad nicht zu schnell oder zu langsam erhöht.

### Wahrnehmung der Raumgröße

Die Referenzraumgröße beträgt 10m x 12m (Breite x Tiefe), der Hörer befindet sich am Punkt 5m in Richtung Raumbreite und 2m in Richtung Raumtiefe, die Schallquelle befindet sich am Punkt 5m in Richtung Raumbreite und 5m in Richtung Raumtiefe, wie in Abbildung 20 dargestellt. Die Distanz zwischen Hörer und Schallquelle beträgt 3m und sie stehen sich direkt gegenüber.



**Abbildung 19: Schematische Darstellung der Anordnung in „Wahrnehmung der Raumgröße“**



**Abbildung 20: Schematische Darstellung der Anordnung in „Wahrnehmung der Raumgröße“**

Je nach Level wird die Raumgröße unterschiedlich stark geändert, mit steigendem Level werden die Änderungen immer kleiner. Ob der Raum kleiner oder größer wird,

entscheidet der Zufall, berechnet wird der Multiplikator für die neue Raumbreite und -tiefe nach Formel (40) für kleinere Räume und Formel (41) für größere Räume.

$$\text{Multiplikator}_{\text{Raumgröße kleiner}} = 1 - \frac{1}{\text{Level} + 1.2} \quad (40)$$

$$\text{Multiplikator}_{\text{Raumgröße größer}} = \left(1 + \frac{1}{\text{Level} + 1.2}\right)^{-1} \quad (41)$$

Zusätzlich wird pro Verdopplung bzw. Halbierung der Raumgröße der Pegel des Nachhallmoduls um 4.5dB angehoben bzw. abgesenkt. Alle anderen Parameter werden nicht verändert. Wieder gilt wie bei der Distanzwahrnehmung, dass diese Werte rein empirisch bestimmt wurden und sich noch Änderungen ergeben können.

### Wahrnehmung des Klangs des Nachhalls

Für diesen Punkt wurde ein großer Raum (22m x 30m) mit einem langen Nachhall ähnlich einer Kirche eingestellt. Schallquelle und Hörer stehen sich im Abstand von 15m in der Mitte des Raumes gegenüber. Ausgehend von den Referenzeinstellungen des Nachhalls wird die Nachhallzeit der tieferen oder der mittleren Frequenzen erhöht, was zu einem *wärmeren* oder *brillanterem* Klang führt. Die Höhen wurden nicht verändert, da sofort eine unnatürliche Überbetonung von den „Zischlauten“, wie „s“ der „sch“ bemerkbar wurde. Die Referenznachhallzeit wurde auf  $T_{\text{low}} = 2.1\text{s}$  für die tiefen Frequenzen und  $T_{\text{mid}} = 1.5\text{s}$  für die mittleren Frequenzen eingestellt. Mit steigendem Level verringern sich der Unterschied zu der Referenznachhallzeit wie in Tabelle 2 ersichtlich:

Level	$T_{\text{low}}$	$T_{\text{mid}}$
Referenz	2.1	1.5
1	5.6	3.7
2	4.6	3.2
3	4.2	2.7
4	3.6	2.4
5	3.2	2.1
6	2.8	2.0
7	2.7	1.8
8	2.4	1.6

**Tabelle 2: Eingestellte Nachhallzeit in Abhängigkeit des Levels**

## Zusammenfassung

Wie bereits erwähnt, sind die genannten Einstellungen nur eine Voreinstellung, die zunächst empirisch bestimmt wurden. Da das Hallmodul sehr vielseitig einsetzbar ist, kann das Trainingsprogramm nach möglichen Wünschen der Testpersonen mit wenig Aufwand adaptiert werden. Vor allem die Zunahme des Schwierigkeitsgrades verdient noch Nachjustierungen.

Im Anhang ist das beschriebene Programm in digitaler Form enthalten. Zum Ausführen muss zuerst die PureData-Datei namens „wurzel.pd“ und darauf folgend die MATLAB-Datei namens „TrReverberation.m“ geöffnet werden.

## 5.2 Ergebnisse und Erfahrungsberichte

Eine erste Evaluierung wurde mit zwei Testpersonen durchgeführt. Beiden Personen wurde der Ablauf des Tests erklärt.

Testperson A hat nach eigenen Angaben eine fundierte musikalische Ausbildung, sich aber noch nicht mit Akustik im technischen Sinn beschäftigt. In „Wahrnehmung der Distanzen“ waren für sie ab Level 5 keine Unterschiede mehr bemerkbar. Die Raumgrößen konnte sie bis Level 10 unterscheiden (die Referenzgröße beträgt 10m x 12m und die Vergleichsgröße bei Level 10 beträgt 9m x 11m mit einer Anhebung des Nachhalls um 1dB). Bei „Klang des Nachhalls“ erreichte sie Level 4.

Testperson B hat nach eigenen Angaben weder eine musikalische Ausbildung noch eine akustische im technischen Sinn. In „Wahrnehmung der Distanzen“ erreichte sie Level 3. Die Raumgrößen konnte sie bis Level 2 und den „Klang des Nachhall“ bis Level 3 unterscheiden. Sie benutzte die Funktion „Antwort zeigen“, wodurch sie nach einigen falschen Antworten „wieder neue Motivation fand“.

Beide Personen wünschten sich mehr Abwechslung des Rohmaterials, derzeit ist nur eine Frauenstimme zu hören. Laut Person A ist die Steigerung des Schwierigkeitsgrades angemessen, Testperson B meinte, dass die Punkteabzüge bei falschen Antworten nicht motivierend seien und mehr Rückmeldungen des Programms nötig sind. Derzeit ist nur Anhand des Punktestandes erkennbar, ob eine richtige oder falsche Antwort gegeben wurde. Außerdem waren beiden Testpersonen die Begriffe *wärmer* und *brillanter* unbekannt. Erst nach Vorspielen eines *wärmeren* Klangs im Vergleich eines *brillanteren* konnten diese Begriffe zugeordnet werden. Beide Personen konzentrierten

sich laut eigenen Angaben bei „Wahrnehmung der Distanz“ nur auf die Unterschiede der Lautstärke des Direktsignales.

### **5.3 Mögliche Verbesserungen des Trainingsprogramms**

In der Evaluierung sind bereits einige Kritikpunkte angemerkt worden. Vor allem ein Training für die Einstellung eines Nachhallprozessors wäre sinnvoll, ein Trainingsablauf könnte wie folgt lauten: vom Programm wird ein Referenzklang mit einem beliebig eingestelltem Nachhall abgespielt. Die Testperson muss nun die eingestellten Parameter des Hallmoduls ermitteln, bzw. das Hallmodul manuell so einstellen, dass zwischen Vergleichs- und Referenzklang kein Unterschied mehr hörbar ist. Dieses Training würde vor allem Tontechnikern in Live-Anwendungen nützen, wo schnelles Handeln gefordert ist, und wenig Zeit für „Trial and Error“ bleibt.

Diese Arbeit stellt nur einen Beginn für ein Trainingsprogramm da, für ein ausgereiftes Programm sind noch viele Verbesserungen möglich. Einige Möglichkeiten sind bereits bekannt, durch weitere Evaluierungen werden sicherlich noch mehr Verbesserungsvorschläge aufkommen, welche dann in das Programm einfließen werden.

## 6 Literaturverzeichnis

Ahnert Wolfgang und Reichardt Walter (1981): *Grundlagen der Beschallungstechnik*. Verlag: Hirzel Stuttgart. ISBN 3-7776-0373-2

Ahnert Wolfgang und Tennhardt Hans-Peter (2008): *Raumakustik*. enthalten in Weinzierl Stefan: *Handbuch der Audiotechnik*, Seite 188 - 266. Verlag: Springer. ISBN 978-3540343004

Beranek Leo (1996): *Concert and Opera Halls: how they sound*. Verlag: Acoustical Society of America". ISBN 1-56396-530-5

Berg Jan (2005): *OPAQUE – a tool for the elicitation and grading of audio quality attributes*. AES Convention Paper 6480.118th Convention, Barcelona, Spain, May 2005

Cremer Lothar und Müller Helmut (1978): *Die wissenschaftlichen Grundlagen der Raumakustik*. Verlag: Hirzel, Stuttgart. ISBN 3-7776-0315-5

Extra Denis, Simmer Uwe, Fischer Sven, und Joerg Bitzer (2006): *Artificial Reverberation: Comparing algorithms by using monaural analysis tools*. AES 121st Convention, San Francisco, CA, USA, 2006 Oktober 5–8

Gardner William (1998): *Reverberation Algorithms*. enthalten in Kahrs Mark, Brandenburg Karlheinz: *Applications of Digital Signal Processing to Audio and Acoustics*. Seite 85 – 130. Verlag: Kluwer Academic Publisher. ISBN 0-7923-8130-0

Griesinger David (2007): *Perception of Concert Hall Acoustics in seats where the reflected energy is stronger than the direct energy - or... Why do Concert Halls sound different - and how can we design them to sound better?*. Vortrag bei AES Convention in Wien, Österreich, Mai 2007. URL: <http://www.davidgriesinger.com/poster.doc>. aufgerufen am [15.06.2010]. zuletzt aktualisiert am [21.04.2010]

Graber Gerhard und Weselak Werner (2008): *Raumakustik*. Lehrbehelf der Lehrveranstaltung Raumakustik. Version 5.2. Institut Breitbandkommunikation Technische Universität Graz

Jot Jean Marc und Chaigne Antoine (1991): *Digital delay networks for designing artificial reverberators*. Audio Engineering Society Convention, Februar 1991.

Moulton Dave (1995): *Golden Ears User Guide*. [CD-ROM] KIQ Productions Inc. USA

Musil Thomas, Sontacchi Alois, Noisternig Markus und Höldrich Robert (2007): *Binaural-Ambisonic 4.Ordnung 3D-Raumsimulationsmodell mit ortsvarianten Quellen und Hörerin bzw. Hörer*. PD IEM Report 38/07.

URL: [http://iem.at/projekte/publications/iem\\_report/report38\\_07/report38\\_07](http://iem.at/projekte/publications/iem_report/report38_07/report38_07)  
aufgerufen am [15.06.2010].

Musil Thomas (2010): schriftliche Aufzeichnungen eines persönlichen Gesprächs am 01.06.2010 über Feedback Delay Network

Neher Tobias, Rumsey Francis, and Brookes Tim (2002): *Training of listeners for the evaluation of spatial sound reproduction*. AES Convention Paper 5584. 112th Convention 2002 May 10–13 Munich, Germany

Neher Tobias (2004): *Towards a spatial ear trainer*. Dissertation. Department of Music & Sound Recording School of Arts, University of Surrey

Puckette Miller S.: *Pd Documentation Rel-0.42-5*. URL:  
[http://crca.ucsd.edu/~msp/Pd\\_documentation/](http://crca.ucsd.edu/~msp/Pd_documentation/). aufgerufen am [19.06.2010]

Pulkki Ville und Karjalainen Matti (2001): *Localisation of Amplitude-Panned Virtual Sources: Stereophonic Panning*. AES Vol. 49 No. 9, 2001 September