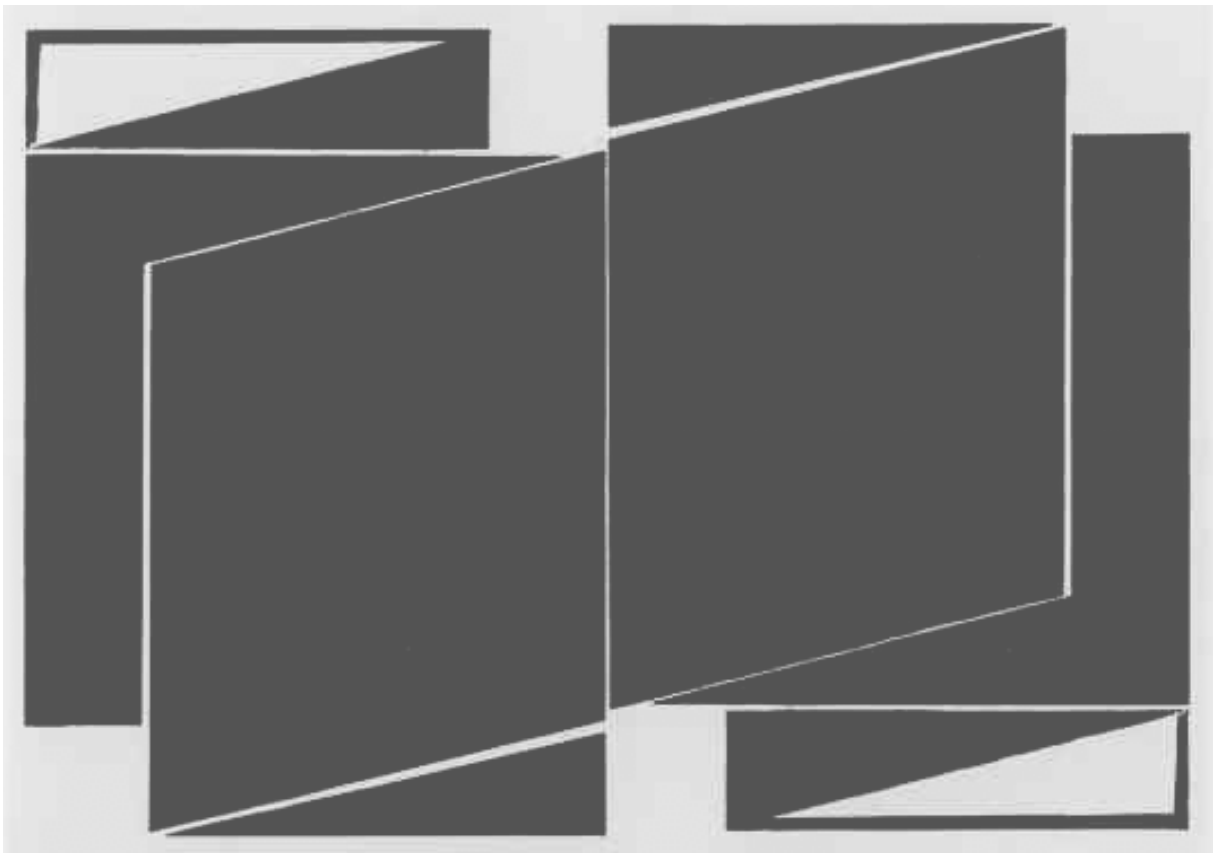

RAUMSIMULATION DURCH FALTUNG



**DIPLOMARBEIT VON ROMAN HAEFELI
FEBRUAR 2003**

AED 09/01, SAE INSTITUT, ZÜRICH

INHALT

1 EINLEITUNG	3
2 WIE FUNKTIONIERT FALTUNGSHALL	4
2.1 THEORETISCH ÜBERLEGUNGEN ZUM HALL	4
2.2 DIE RAUM-IMPULSANTWORT (RIA)	5
2.3 DIE FALTUNG	6
3 RIA-MESSUNG IN DER THEORIE	10
3.1 EINFLUSS DER MESSKETTE AUF DIE RIA	10
3.2 RIA-MESSMETHODEN	12
3.2.1 Messung mit einem Impuls	12
3.2.2 Zwei-Kanal-FFT-Analyse	14
3.2.3 Maximallängenfolgen (MLS)	15
3.2.4 FFT-Analyse mit einem Sweep	18
3.3 KORREKTUR DER ÜBERTRAGUNGSFEHLER DES MESSSYSTEMS	20
4 RIA-MESSUNG MIT VORHANDENEM EQUIPEMENT	23
4.1 WAHL UND ERSTELLEN DES ANREGUNGSSIGNALS	23
4.2 MESSUNG DES ZIMMERS	25
4.2.1 Mikrofonierung	25
4.2.2 Anregung	25
4.2.3 Wahl der f_s und Bitrate	26
4.2.4 Messung	26
4.2.5 Störpegelverringern durch Mittelung	27
4.2.6 Kompensation des Lautsprecher-Übertragungsfehlers	28
4.2.7 Faltung	30
4.2.8 Vergleich mit dem Original	30
4.3 EINFLUSS VERSCHIEDENER PARAMETER AUF DIE RIA	31
4.3.1 Anregungssignal	31
4.3.2 Mikrofonierung und LS-Position	31
4.3.3 Messung und Mittelung	32
4.3.4 Spektrum-Kompensation	32
4.3.5 Ergebnisse	32
5 TEST: KANN FALTUNG HERKÖMMLICHEN HALL KOPIEREN?	34
5.1 MESSUNG DER IMPULSANTWORT <i>TRUEVERB</i>	34
5.2 ERZEUGUNG DER VERGLEICHSDATEIEN	35
5.2.1 Stimme, verhallt durch <i>TrueVerb</i>	35
5.2.2 Stimme, verhallt durch Faltung	35
5.3 VERGLEICH	36
5.3.1 Voraussetzungen für einen Vergleich	36
5.3.2 Deutung des Ergebnisses	36
6 ABSCHLIESSENDE ÜBERLEGUNGEN ZUR PRAXIS	36
7 ANHANG	39
7.1 QUELLENVERZEICHNIS	39
7.2 CD-INHALT	40

1 EINLEITUNG

Bis vor ein paar Jahren war das Verfahren, durch Faltung einen Hall zu erzeugen, im praktischen Alltag wenig bis gar nicht bekannt. Durch die rasante Entwicklung der Prozessor-Technologie ist es nun möglich, in Echtzeit einen Faltungshall zu berechnen, was dieses Verfahren auch für den Studio-Einsatz tauglich macht. Die Erzeugung eines Faltungshalls erfordert einerseits grosse Ressourcen an Rechenkapazität, andererseits lässt dieses Verfahren keine Wünsche offen, was die technische Seite betrifft. Theoretisch gibt es keinen Hall, der sich nicht durch Faltung realisieren lässt. (Man bedenke die Tragweite dieser Aussage!) Da diese Technik einen anderen Ansatz verfolgt als die herkömmliche Hallerzeugung, müssen die Eigenheiten im Umgang mit ihr eingehend betrachtet werden, um sinnvoll mit ihr arbeiten zu können.

Was genau die Theorie ist, die dahinter steckt, soll Gegenstand dieser Arbeit sein. Mit einfachen Mitteln soll diese Theorie in die Praxis umgesetzt werden und die Problemstellungen, die sich daraus ergeben, einer näheren Betrachtung unterzogen werden. Im weiteren wird in dieser Arbeit die These aufgestellt, dass mit dem Aufkommen von Faltungshall-Geräten die Verwendung herkömmlicher Geräte aus technischer Sicht überflüssig wird. Diese Behauptung soll näher untersucht werden.

Der gegenwärtige Zeitpunkt, sich diesem Thema zuzuwenden, scheint dem Autor günstig zu sein. Seit erst ungefähr drei Jahren sind Geräte zur Faltungshallerzeugung für den Studio-Einsatz auf dem Markt erhältlich. Auch wenn sich die Preise momentan noch im Segment der gehobeneren Klasse befinden, geht der Autor davon aus, dass sie sich langsam auf dem Vormarsch in den Studio-Alltag befinden und ein zukünftiger User sich die wichtigsten Grundkenntnisse zum Thema aneignen sollte.

Um eine gute Reproduzierbarkeit der Arbeitsmethode zu gewährleisten, wurde für diese Arbeit allgemein bekannte oder gratis herunterladbare Software verwendet.

2 WIE FUNKTIONIERT FALTUNGSHALL?

2.1 THEORETISCHE ÜBERLEGUNGEN ZUM HALL

Um einen Hall zu simulieren, gab es bis jetzt schon verschiedene Ansätze. Die Erzeugung mit simplem Bandecho, Federhall, Plate-Hall, oder heute vor allem mittels digitalen Hallgeräten. Dabei wurde die Qualität des räumlichen Empfindens ständig verbessert. Bei einem Bandecho oder Delay wird die folgende Situation simuliert:

Eine punktförmige Schallquelle befindet sich in einer absolut leeren, bodenlosen Umgebung. In einiger Entfernung der Schallquelle befindet sich eine riesige, absolut schallharte Wand. In Natura kommt so eine Situation eigentlich nie vor, deshalb vermittelt ein durch Delay „verhalltes“ Signal kein richtiges Raumgefühl. Auch Federhall und Platehall simulieren eigentlich keine echten Räume. Die Feder kann nur in einer Dimension angeregt werden, der Platehall schon in immerhin zwei. Bei den heute üblichen digitalen Geräten wird die Umsetzung komplexer. Das verhallte Signal wird zusammengesetzt aus den einzelnen Komponenten, die auch beim natürlichen Hall vorhanden sind:

- Direktsignal
- frühe Reflexionen, deren Dichte kubisch ansteigt
- Hallfahne, welche theoretisch zunehmend dichter wird, aber im Pegel abnimmt.

Die Qualität eines auf diese Weise erzeugten Halls hängt von verschiedenen Faktoren ab. Einerseits spielt die Dichte (Anzahl Reflexionen pro Zeiteinheit) eine Rolle. Andererseits kommt es auch auf die Algorithmen, welche die zeitliche Staffelung und die Pegelverhältnisse der einzelnen Reflexionen bestimmen, an. Eine grössere Dichte bedeutet aber auch ein Mehraufwand an Rechenleistung. Zu berücksichtigen gilt, dass verschiedene Oberflächen unterschiedliche Reflexionseigenschaften besitzen, d.h. je nach Beschaffenheit der Oberflächen werden die hohen Frequenzen mehr oder weniger absorbiert.

In vielen Hallgeräten wird diesem Phänomen mit einem Parameter, welcher oft als „HF-damp“ bezeichnet wird, Rechnung getragen. Dieser stellt einen eingefügten Hi-cut- oder Hi-shelf-Filter in der Feedback-Schleife dar.

In der Praxis werden solche Geräte dem realen Hall natürlicher Räume bei weitem nicht gerecht. Würde man zum Beispiel ein normales Wohnzimmer bis ins Detail simulieren wollen, müsste auch die Form aller Möbel, die Fläche jedes Fensters und der Türe usw. miteinbezogen werden. Wir sehen: Auf diese Weise natürliche Räume nachzubilden, würde einen immensen, selbst mit heutigen Rechnern nicht zu bewältigenden Rechenaufwand bedeuten. Viele akustische Umgebungen lassen sich deshalb nicht, oder nur schwer mit

herkömmlichen Hallgeräten realisieren. Vor allem im Filmbereich geht es häufig darum, bestehendes Audiomaterial in ungewöhnliche Umgebungen einzubetten. Beispielsweise die Synchronisation eines Dialoges im Auto gestaltet sich sehr schwierig. Aber auch Umgebungen wie grosse Plätze, Strassenschluchten, Hinterhöfe, Wald usw. sind schwierig mit herkömmlichen Methoden nachzubilden.

Eine Theorie des Faltungshalls verfolgt einen völlig anderen Ansatz:

Die grundlegende Idee besteht darin, dass man die Akustik eines schon bestehenden Raumes als Vorlage nimmt, um vorhandenes „trockenes“ Audiomaterial zu verhallen, so als ob es in eben diesem Raum aufgenommen worden wäre. Es stellt sich deshalb zuerst die Frage, wie man die Akustik eines Raumes definieren kann, ohne sie durch Teilfaktoren wie Volumen, Form, Oberflächenbeschaffenheit usw. umschreiben zu müssen.

2.2 DIE RAUM-IMPULSANTWORT (RIA)

Ein Raum lässt sich als System betrachten. Regt man ihn in mit einem Signal an (INPUT), erklingt das Signal verändert wieder (OUTPUT). Ein Raum besitzt sogar die Eigenschaften eines LTI-Systems (LTI bedeutet *linear timeinvariant*):

- **„Linearität:** Wird das Eingangssignal in seiner Amplitude um einen Faktor λ verstärkt, so muss das Ausgangssignal des Systems ebenfalls eine um den Faktor λ verstärkte Amplitude besitzen.
- **Zeitinvarianz:** Die Übertragungseigenschaften des Systems müssen zu jedem Zeitpunkt dieselben sein und dürfen sich nicht ändern.“ [1, Seite 84]

Die Eigenschaften eines LTI-Systems werden von dessen Impulsantwort beschrieben. („Ein LTI-System ist durch seine Impulsantwort eindeutig charakterisiert.“ [2, Seite 7]) Als Impulsantwort bezeichnet man jenes Signal, welches am „Output“ ausgegeben wird, wenn man das System mit einem Impuls anregt. Dieser Impuls, der sogenannte Dirac-Stoss, soll unendlich kurz sein und eine maximale Energie besitzen. Mathematisch wird das Signal eines Dirac-Stosses so beschrieben:

Zum Zeitpunkt $t=0$ besitzt die Amplitude den Wert 1, zu allen anderen Zeitpunkten den Wert 0. Würde man nun einen Raum mit einem Dirac-Stoss anregen und den Nachhall dieses Signals aufnehmen, hätte man folglich die akustischen Eigenschaften des Raumes erfasst.

Nebenbei: Auch andere Systeme wie Filter, EQs u. ä., die linear und zeitinvariant sind, lassen sich mit einer Impulsantwort eindeutig beschreiben. Dynamik bearbeitende Effekte wie Kompressoren, Gates, DeEsser erfüllen die Bedingung der Linearität nicht; Effekte wie Pitch-Shifting, Chorus, Phaser, Flanger erfüllen die Bedingung der Zeitinvarianz nicht. Sie alle lassen sich deshalb nicht durch eine Impulsantwort charakterisieren.

2.3 DIE FALTUNG

In der digitalen Audiotechnik wird mit zeitdiskreten Signalen gearbeitet. Im Gegensatz zu zeitkontinuierlichen Systemen gibt es in zeitdiskreten Systemen eine kleinste Zeiteinheit, welche in der Audiotechnik der Länge eines Samples entspricht. Von einem CD-Track wäre diese Einheit $1 / 44'100 = 22.67\mu\text{s}$ lang. Liegt nun die Impulsantwort eines Raumes in digitalisierter Form vor, kann man sie als Folge von Messwerten betrachten, deren Anzahl der Anzahl Samples entspricht.

Nehmen wir an, wir senden ein beliebiges Signal per Lautsprecher in einen Raum und nehmen das verhallte Signal mit einem Mikrofon ab. Betrachten wir die Amplitude des Signals, welches vom Mikrofon zu einem bestimmten Zeitpunkt $t=0$ geliefert wird. Woraus setzt sie sich zusammen? Einerseits aus einem Anteil Direktsignal, welches um soviel früher (sagen wir: zum Zeitpunkt $t=-1$) gesendet wurde, wie der Schall Zeit benötigt, um die Distanz zwischen Lautsprecher und Mikrofon zurückzulegen. Ein weiterer Anteil besteht aus der Amplitude, welche das Signal hatte, kurz bevor es die Amplitude des Direktsignals erreichte (zum Zeitpunkt $t=-2$). Dieser Anteil fällt jedoch weniger ins Gewicht als der erste, da es sich um eine Reflexion handelt. Zu einem noch früheren Zeitpunkt ($t=-3$) hatte das Signal am Lautsprecher eine Amplitude, die jetzt (zum Zeitpunkt $t=0$) abgeschwächt durch zwei Reflexionen zu den ersten beiden Werten addiert werden. Handelt es sich um ein analoges Signal, so müsste man diese Rechnung für jeden beliebigen Zeitpunkt durchführen.

Konkret heisst das: Um einen präzisen Wert der Amplitude, welche am Mikrofon zum Zeitpunkt $t=0$ anliegt, zu ermitteln, müsste man eine unendliche Anzahl an Werten zueinander addieren.

In der Digitaltechnik werden die Additionen auf einen endlichen Wert beschränkt. Dieser ist abhängig von der Anzahl Samples der Raum-Impulsantwort und der Anzahl Samples des zu verhallenden Signals.

Eine Raum-Impulsantwort, welche in digitalisierter Form vorliegt, können wir folgendermassen umschreiben:

Abb. 2-1 bis 2-3: Darstellung einer Faltung

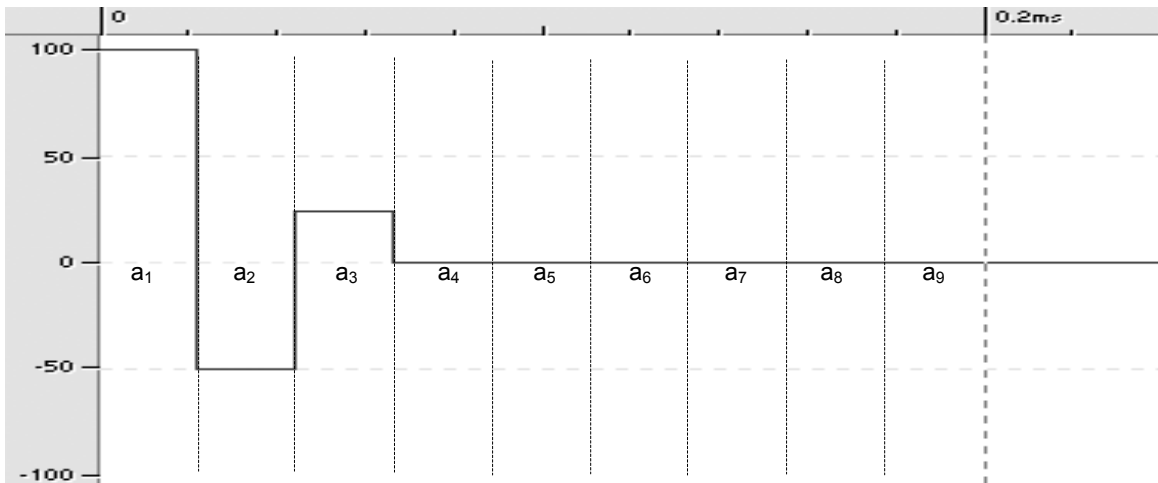


Abb. 2-1: Beispiel einer kurzen RIA (Darstellung im Zeitbereich, Amplitude in Prozent)

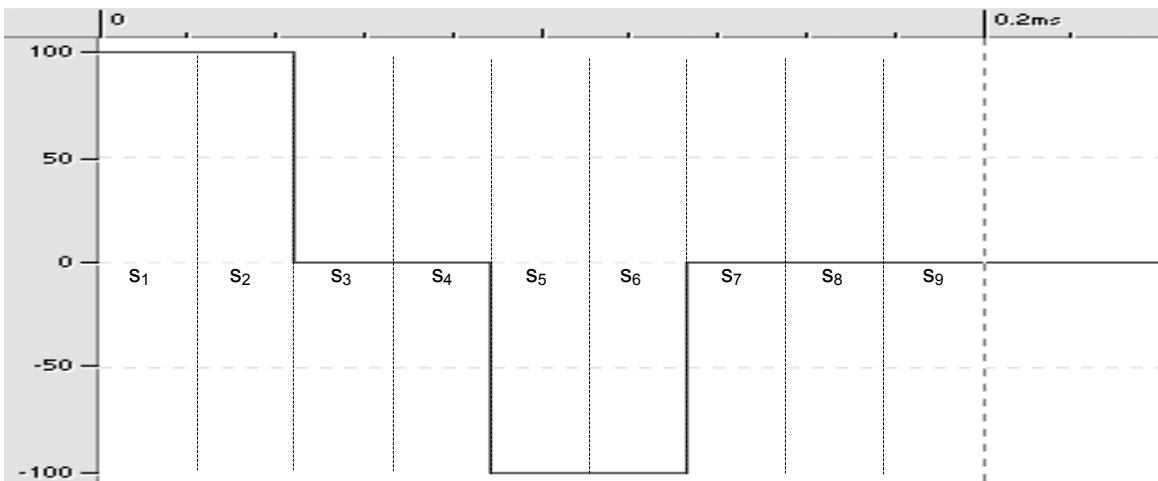


Abb. 2-2: Beispiel eines einfachen zu verhaltenden Signals

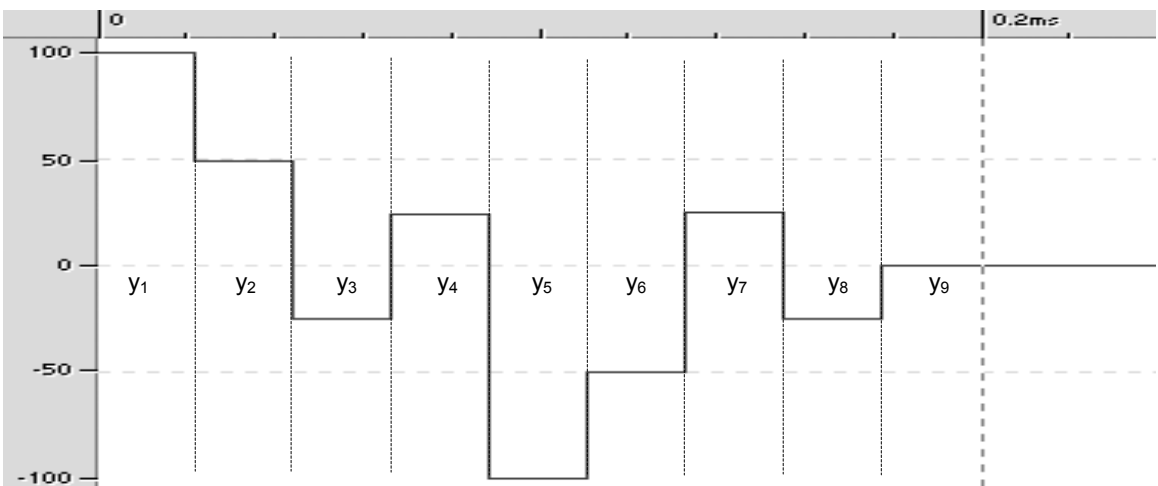


Abb. 2-3: Aus der Faltung resultierendes Signal

Jedes Sample stellt eine Reflexion dar; der Wert des Samples besagt, wie stark diese Reflexion ist (Wert=0 bedeutet keine Reflexion, Wert=1 bedeutet hundertprozentige Reflexion).

Wir nummerieren die Samples der Impulsantwort mit a_1, a_2, a_3 usw. (Abb. 2-1). Die Samples des zu verhallenden Signals werden mit s_1, s_2, s_3 usw. (Abb. 2-2) nummeriert. Da nun sowohl mit werte- als auch mit zeitdiskreten Signalen gearbeitet wird, können wir den Zeitpunkten ganze Zahlen zuordnen, d. h. auch sie werden durchnummeriert mit $t=1, t=2, t=3$ usw.

Bezeichnen wir das erste Sample der Impulsantwort als den Direktanteil, das zweite Sample als erste Reflexion, das dritte Sample als zweite Reflexion usw., dann ergibt sich als Wert für das erste Sample des verhallten Signals:

Bsp. nach Abb. 2-1 bis 2-3:

$$y(t=1) = s_1 a_1$$

$$y_1 = 1 * 1 = 1$$

Das zweite Sample Signals setzt sich zusammen aus dem Direktanteil, der sich in diesem Fall aus der Multiplikation des zweiten Samples des zu verhallenden Signals (s_2) mit dem ersten Sample der Impulsantwort (a_1) ergibt, und der ersten Reflexion, die sich ergibt aus der Multiplikation des ersten Samples des zu verhallenden Signals mit dem zweiten Sample der Impulsantwort:

$$y(t=2) = s_2 a_1 + s_1 a_2$$

$$y_2 = 1 * 1 + 1 * (-0.5) = 0.5$$

Dieses Muster lässt sich nun immer weiter fortsetzen:

$$y(t=3) = s_3 a_1 + s_2 a_2 + s_1 a_3$$

$$y_3 = 0 * 1 + 1 * (-0.5) + 1 * 0.25 = -0.25$$

$$y(t=4) = s_4 a_1 + s_3 a_2 + s_2 a_3 + s_1 a_4$$

$$y_4 = 0 * 1 + 0 * (-0.5) + 1 * 0.25 + 1 * 0 = 0.25$$

Für ein Sample des verhallten Signals zu einem beliebigen Zeitpunkt gilt folglich:

Das Sample wird mit dem „Direktanteil“ multipliziert, dazu addiert man das Produkt aus der „ersten Reflexion“ und dem vorherigen Sample, dann das Produkt aus der „zweiten Reflexion“ und dem Sample, welches zwei Zeiteinheiten zurückliegt, dann das Produkt aus der „dritten Reflexion“ und dem Sample, welches drei Zeiteinheiten zurückliegt usw.

Was hier umständlich umschrieben wird, lässt sich auch mit einer einzigen mathematischen Formel beschreiben:

$$y(t) = \sum a(i) * s(t-i) \quad \text{für } i = (-\infty \rightarrow \infty)$$

Ausgedeutet bedeutet dies, dass die Amplitude des verhallten Signals zum Zeitpunkt t aus der Summe aller Produkte von „Reflexionen“ der Impulsantwort und den

entsprechenden, immer weiter zurückliegenden Samples des zu verhallenden Signals errechnet wird.

Für die Faltung eines Signals der Länge S Samples mit einer Impulsantwort mit A Samples müssen also $S \cdot A$ Multiplikationen und eben so viele Additionen errechnet werden. Diese Tatsache macht deswegen die Faltung zu einer äusserst rechenintensiven Angelegenheit. Für eine Echtzeit-Faltung eines Signals der Samplingrate f_s mit einer Impulsantwort der Länge A Samples, müssen also pro Sekunde $2 \cdot f_s \cdot A$ Rechenoperationen durchgeführt werden [1 Seite 85].

Nehmen wir als Beispiel eine Samplingrate von 44,1kHz, wie sie bei der CD verwendet wird, und eine Impulsantwort der Länge von 6 Sekunden (das ergibt: $6 \cdot 44100 = 264600$ Samples), so muss der Rechner $2 \cdot (44100 \cdot 264600) = 23.34$ Milliarden Rechenoperationen pro Sekunde bewältigen können, dies wohlbemerkt für die Verhallung eines Monosignals. Wird eine Stereo-RIA verwendet (mono→stereo), oder ein Stereosignal mit einer Mono-RIA gefaltet (stereo→pseudostereo), oder ein Stereosignal mit einer Stereo-RIA (stereo→stereo), so verdoppelt sich diese Zahl. Deshalb ist es verständlich, dass erst seit kurzem kommerzielle Geräte zur Hallerzeugung durch Faltung hergestellt werden.

In der Rechenleistung, und nicht in der Programmierung liegt die Schwierigkeit, ein solches Gerät zu realisieren. Ob man mit einer entsprechenden Software einen Echtzeit-Faltungshall erzeugen kann, ist somit abhängig von der Leistung der Hardware, der Länge der RIA und der Samplingrate.

Das Ergebnis, also das verhallte Signal, ist um die Hallfahne der RIA ($A-1$ Samples) länger als das Ursprungssignal. Denn laut der Formel haben frühere Samples Einfluss darauf, was später geschieht.

3 RIA-MESSUNG IN DER THEORIE

Die Qualität des Faltungshalls ist abhängig von der Qualität der verwendeten RIA. Die Signal-to-Noise-Ratio einer RIA soll so hoch wie möglich sein. Denn es ist wichtig, dass die Hallfahne bis am Schluss natürlich ausklingt. Selbst kleine Fehler machen sich deutlich bemerkbar. Ein konstanter Rauschpegel in der RIA würde dazu führen, dass der Hall nach einer Weile auf einem bestimmten Pegel „stehen bleibt“. Grundsätzlich kann man sagen: Je länger die RIA ist, desto auffälliger werden Unregelmässigkeiten im Nachhall. Bei sehr schnell ausklingenden RIAs ist die Wahrscheinlichkeit gross, dass die meisten Fehler durch die vorangegangenen lautereren Signalanteile maskiert werden.

RIA-Messungen werden auch vorgenommen, um andere raumakustische Werte wie Nachhallzeit (RT60), Sprachverständlichkeit usw. zu berechnen. Für solche Anwendungen reicht eine RIA mit niedriger S/N-Ratio. Für Verhallung von Musik oder Sprache sind jedoch RIAs mit deutlich höherer Qualität erforderlich.

Da es sich bei der RIA-Messung um eine akustische Messung handelt, sind äussere Störeinflüsse, welche das Messergebnis verfälschen, eigentlich immer vorhanden. Es muss daher nach Wegen gesucht werden, diese Einflüsse so gering wie möglich zu halten.

Der wohl häufigste Störeinfluss dürfte der Umgebungslärm darstellen. Bei Messungen, die draussen stattfinden, spielt auch der Wind eine Rolle. Einerseits entstehen an den Mikrofonkapseln Rumpelgeräusche, andererseits kann bei Wind die Umgebung akustisch nicht mehr als ein LTI-System betrachtet werden; die Bedingung der Zeitinvarianz ist bei sich bewegender Luft nicht mehr gegeben. Bei bestimmten Messmethoden fallen durch Zeitinvarianz entstehende Messfehler sehr ins Gewicht.

Aber nicht nur Umwelteinflüsse beeinträchtigen das Signal. Der Frequenzgang der gesamten Übertragungsstrecke des Messsystems, vor allem des Lautsprechers, prägt die RIA.

3.1 EINFLUSS DER MESSKETTE AUF DIE RIA

Bei der Faltung eines Audiosignals mit einer RIA wird genau die Situation simuliert, die der Position von Mikrofon und Lautsprecher entspricht. Beispielsweise soll eine Sängerin virtuell in ein Studio mit besonders guter Akustik platziert werden. Dazu ist es wichtig, dass das zu verhallende Signal möglichst trocken, also ohne Raumeinfluss, aufgenommen wurde. Der verhallte Gesang schliesslich übernimmt alle Eigenschaften des bei der Messung verwendeten Systems. In unserem Beispiel hätte die Sängerin dieselbe virtuelle Abstrahlcharakteristik wie der Lautsprecher, abgesehen davon, dass ihr auch der

Frequenzgang und Phasenspektrum des Lautsprechers aufgeprägt würden. Theoretisch lässt sich mit einer einzigen RIA also nur eine spezifische Situation nachbilden. Würde man eine Violine virtuell in denselben Raum stellen wollen, müsste für die Messung der RIA ein Lautsprecher mit der gleichen Abstrahlcharakteristik, wie die der Violine, verwendet werden.

In der Praxis jedoch spielt Richtcharakteristik keine so grosse Rolle. Wichtiger für Musikanwendungen ist, dass die RIA gut klingt und den gestellten Anforderungen entspricht. Soll beispielsweise eine Solo-Trompete in eine schon bestehende Orchester-Aufnahme eingefügt werden, muss die RIA desselben oder eines ähnlich klingenden Raumes gemessen werden. Die Positionierung des Lautsprechers sollte derjenigen des virtuellen Instrumentalisten entsprechen. Es soll auf dieselbe Weise mikrofoniert werden wie bei der Aufnahme des Orchesters.

Oft werden für Messungen Lautsprecher mit Kugel-Charakteristik verwendet. Dies geschieht, damit die fertige RIA möglichst universell, und nicht nur für eine bestimmte Situation eingesetzt werden kann. Dieses Verfahren führt zu einem volleren Klang der RIA und des mit ihr gemachten Halls, weil der gesamte Raum gleichmässig mit allen Frequenzen angeregt wird, sprich es sind in der RIA mehr Reflexionen vorhanden. Für RIAs, welche die Anforderungen nach ISO 3382 [4] erfüllen, werden nur Lautsprecher eingesetzt, die in alle Richtungen das volle Spektrum abstrahlen.

Die Wahl des Lautsprechers ist entscheidend für die Art und Qualität des Ergebnisses. Die meisten klangverfärbenden Eigenschaften des Lautsprechers, die sich später auch in der RIA niederschlagen, sind unerwünscht. Am meisten ins Gewicht fällt der Frequenzgang. Praktisch jeder Lautsprecher weist für eine saubere Messung ungeeignete Unebenheiten im Frequenzgang auf, welche wiederum den Frequenzgang der gemessenen RIA verfälschen. Auch in punkto Phasentreue bereiten die Lautsprecher grosse Probleme. Da es keine im Frequenzgang absolut linearen Lautsprecher gibt, müssen Wege gefunden werden, diese Fehler zu eliminieren.

Im Grunde gilt für die Mikrofonierung dasselbe wie für die Lautsprecher. Alle Eigenschaften der Mikrofonierung (inklusive des Mikrofons) werden in die RIA übernommen. Hier gelten die gleichen Gesetzmässigkeiten, die auch sonst bei einer Mikrofonierung zu beachten sind.

Da es bei einer Verhallung um die räumliche Tiefenstaffelung geht und nicht um eine möglichst genaue Ortbarkeit, macht es eher Sinn, eine A/B-Anordnung zu verwenden. Entscheidender bei der Wahl der Mikrofonierung ist die subjektive Beurteilung des Ergebnisses. Analog zum Lautsprecher, werden häufig Mikrofone mit Kugelcharakteristik eingesetzt, um möglichst alle Reflexionen einzufangen.

Im Gegensatz zum Lautsprecher werden in der Praxis Veränderungen, die vom Mikrofon herrühren, in der RIA belassen, also nicht entfernt. Man möchte bewusst den Klang des Mikrofons in die RIA miteinbeziehen, um die Wahl des Mikrofons als Möglichkeit zur Klanggestaltung zu nutzen. Vorausgesetzt wird die Verwendung hochwertiger Mikrofone.

Alle gängigen Mikrofonierungs-Verfahren können für RIA-Messung angewandt werden. Mit entsprechender Mikrofonierung ist es möglich, auch in Stereo oder mehrkanalig (Surround, 5.1) aufzunehmen. Als praktisch erweist sich bei mehrkanaligen Messungen, dass man nicht unbedingt mehrere teure Mikrofone benötigt. Da ja immer dasselbe Anregungssignal gesendet wird, kann jede Mikrofonposition nacheinander mit nur einem Mikrofon gemessen werden. Der Zeitaufwand für eine Messung vergrössert sich, dafür werden die Aufwendungen für das Messmaterial verringert.

Auf die verschiedenen Mikrofonierungs-Verfahren soll hier nicht näher eingegangen werden. Es würde den Rahmen dieser Arbeit sprengen. Es gibt genügend Fachliteratur zu diesem Thema.

Die Wahl der Mikrofone und deren Positionierung bietet grosse Gestaltungsmöglichkeiten im Hinblick auf die fertige RIA und ist für deren Qualität entscheidend. Überlegungen über den späteren Verwendungszweck der RIA sollen miteinfließen in diese Wahl.

3.2 RIA-MESSMETHODEN

3.2.1 Messung mit einem Impuls

Die wohl simpelste Art, zu einer RIA zu gelangen, ist folgende Methode: Man sendet per Lautsprecher einen Impuls in den Raum (bzw. in die zu messende Umgebung) und nimmt die Antwort per Mikrofon auf. Fertig. Das Problem dieser Methode ist der wahrscheinlich sehr hohe Störpegel. In den seltensten Fällen ist der Aufnahmeort absolut ruhig.

Das Anregungssignal dieses Messverfahrens, ein Dirac-Stoss, ist eigentlich für messtechnische Zwecke ungeeignet. Denn je höher die Energie des Anregungssignal, desto geringer der Einfluss von äusseren Störungen auf das Messergebnis. Da aber ein Dirac-Stoss naturgemäss sehr kurz ist, wird die gesamte abgestrahlte Energie auf einen sehr kurzen Zeitraum beschränkt. Zudem hat ein Dirac-Stoss die Eigenschaft, dass er dieselbe spektrale Verteilung aufweist wie weisses Rauschen. Das bedeutet, dass die tieffrequenten Anteile relativ schwach vertreten sind. Dies wiederum hat zur Folge, dass niederfrequente äussere Störungen stark zu einer Verfälschung des Messergebnisses führen. Um dieser ungünstigen spektralen Verteilung entgegenzuwirken, könnte man auch einen rosa gefärbten

Impuls als Anregungssignal verwenden. Das Ergebnis müsste dann allerdings im Frequenzgang angepasst werden. Am einfachsten geht das mit einem FFT-Filter (FFT bedeutet Fast Fourier Transformation), damit die RIA die Färbung annimmt, als wäre sie mit einem Dirac-Stoß als Anregungssignal entstanden.

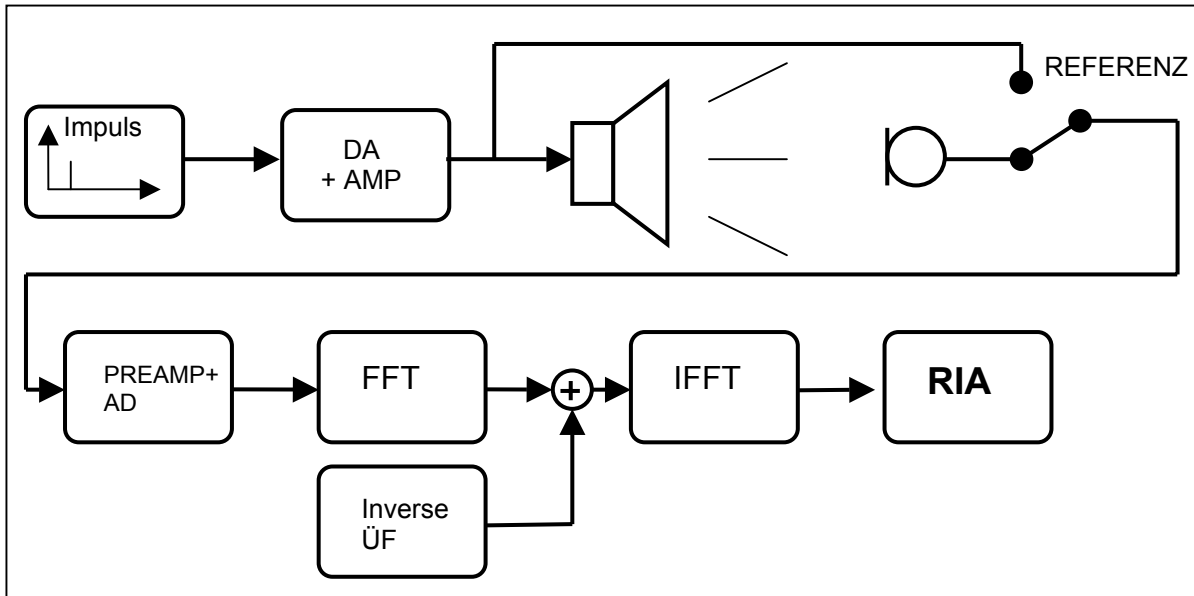


Abb. 3-1: Flussdiagramm der Messkette mit Dirac-Impuls als Anregung

Da ein Raum (oder die zu messende Umgebung) ein zeitinvariantes System darstellt, müssen mehrere Messungen ein identisches Messergebnis liefern, falls:

- die Position und Abstrahlrichtung des Anregungssignal-Senders (sprich: Lautsprecher) nicht verändert wurde,
- die Position des Mikrofons dieselbe bleibt,
- keine Luftbewegung (Wind) stattfindet.

Um äussere Störeinflüsse nachträglich zu verringern, nutzt man folgende Tatsache: Mischt man zwei komplexe, nicht phasengleiche Signale zusammen, entsteht ein Signal mit einem Pegel, der im Durchschnitt 3dB höher liegt als der Pegel des einzelnen Signals. Mischt man aber zwei identische, phasengleiche Signale zusammen, so ist der Pegel des Ausgangssignals um 6dB höher als das einzelne Eingangssignal, denn die Amplitude verdoppelt sich.

Werden nun zwei Ergebnisse der Messungen phasengleich zu einander addiert (=gemischt), so erhöht sich der Pegel des Nutzsignals um 6dB, derjenige der äusseren Störeinflüsse nur um 3dB, da es sich hierbei nicht um bei jeder Messung identische Signale handelt. Es resultiert also eine Verbesserung der S/N-Ratio um 3dB. Um nochmals 3dB zu gewinnen, müssen nun zwei weitere Messergebnisse dazugemischt werden (also 4 insgesamt).

Pro Verdopplung der Anzahl Messungen wird eine Erhöhung des Nutzsignals um 3dB gegenüber des Störanteils bewirkt. Um den Störpegel um x dB zu reduzieren, müssen also $n=2^{(x/3)}$ Messergebnisse miteinander addiert werden bzw. gemittelt werden.

Diese Methode zur Reduzierung des Störgeräuschepegels kann sehr aufwändig werden, denn pro 3dB Verbesserung der S/N-Ratio verdoppelt sich auch die Messzeit. Es lohnt sich also, die Störeinflüsse während der Messung so gering wie möglich zu halten. Ist es die Absicht, eine RIA mit einer S/N-Ratio von 90dB zu erhalten, darf, um die Anzahl der Messungen für eine Mittelung auf 64 zu beschränken, die S/N-Ratio der einzelnen Messung nicht kleiner sein als 72dB. Es ist also extrem schwierig, mit dieser Methode gute, für Verhallungszwecke geeignete RIAs zu erhalten.

3.2.2 Zwei-Kanal-FFT-Analyse

Um diese Methode zu verstehen, muss man wissen, was FFT ist bzw. was sie macht. FFT bedeutet schnelle (engl. *fast*) Fourier-Transformation. Mit Hilfe der FFT lässt sich ein Signal vom Zeitbereich in den Frequenzbereich übertragen. „Im Zeitbereich dargestellt“ bedeutet, dass die Amplitude als Funktion der Zeit betrachtet wird. Eine Kurve im Zeitbereich zeigt an, welchen Wert die Amplitude des Signals zu einem bestimmten Zeitpunkt hat. Die Darstellung im Frequenzbereich löst die Amplitude und Phase in Abhängigkeit der Frequenz auf, d.h. die Amplitude ist eine Funktion der Frequenz. Die inverse FFT (IFFT) wiederum überträgt ein Signal vom Frequenz- in den Zeitbereich. Ein Signal lässt sich so (fast) verlustfrei hin- und herübertragen.

Wie der Name schon sagt, benötigt man zur Messung der RIA mit der Zwei-Kanal-FFT-Analyse zwei Eingangskanäle (eines A/D-Wanders). Ein beliebiges Signal (kann auch Musik sein) wird als Anregungssignal verwendet. Es wird also über einen Lautsprecher in den Raum gesendet und über ein Mikrofon einem A/D-Wandler zu geführt. Bevor es aber zum Lautsprecher gelangt, wird es abgezweigt und zu einer Verzögerungsstufe geführt. Die Verzögerungszeit muss so eingestellt werden, dass das abgezweigte Signal dieselbe Laufzeit besitzt wie das Signal, das die gesamte Messstrecke durchläuft (Lautsprecher – Luft – Mikrofon), damit beide Signale phasengleich anliegen. Auch das verzögerte Anregungssignal wird einem A/D-Wandler zugeführt.

Es werden beide Signale getrennt einer FFT-Analyse unterzogen, so erhält man zwei Spektren, einerseits vom (verzögerten) Anregungssignal, andererseits vom verhallten Signal. Nach einer Division des „Anregungsspektrums“ durch das „verhallte“ Spektrum bleibt die Übertragungsfunktion des Raumes übrig. Es muss nur noch durch eine inverse FFT in den Zeitbereich übertragen werden, damit eine hör- und verwendabre RIA vorliegt.

Der Vorteil dieser Methode ist, dass ein beliebiges Anregungssignal verwendet werden kann, da ja durch die Division (bei logarithmischer Darstellung: Subtraktion) sich der Frequenzgang des Signals auslöscht; übrig bleibt die Information des Raumes. Es liegen aber nur in denjenigen Frequenzen eine Informationen über den Raum vor, die im Anregungssignal vorhanden waren.

Und hier sind wir schon beim Nachteil, falls man den Raum mit Musik anregt, denn in Musik sind kaum alle Frequenzen vertreten. Die S/N-Ratio in den fehlenden Frequenzbereichen ist also denkbar schlecht. Weiter ist die FFT nur dann genau, wenn sich das Signal periodisch mit der Blocklänge wiederholt [3, Seite 110]. Die FFT wird jeweils in Blöcken mit der Länge von 2^n Samples vorgenommen. Weil sogar weisses Rauschen diese Forderung nicht erfüllt – betrachtet man das Spektrum eines Ausschnittes einer Blocklänge von weissem Rauschen, so sind auch hier tiefe Frequenzeinbrüche feststellbar -, müssen auch hier viele Mittelungen gemacht werden, um zu verhindern, dass fehlende Frequenzen im Anregungssignal zu willkürlichen Informationen in diesen Bereichen führen.

Fazit: Auch diese Methode erfordert langwierige Messvorgänge und das Ergebnis ist nur begrenzt frei von Störanteilen.

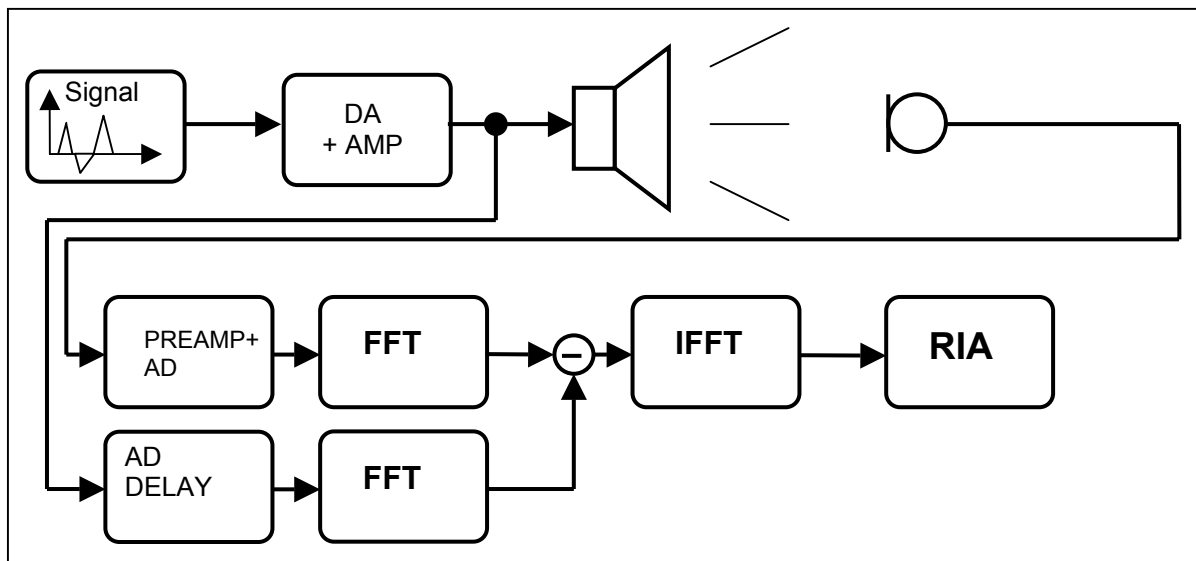


Abb. 3-2: Flussdiagramm einer 2-Kanal-FFT-Analyse

3.2.3 Maximallängenfolgen (MLS)

Bei der Maximallängenfolge (*maximal length sequence*) handelt es sich um eine ganz besonderer Art von Signal. Die Amplitude eines solchen Signals kann nur zwei Werte annehmen, 1 und 0, oder positiver und negativer Vollausschlag. Man kann es also Folge von binären Zahlenwerten betrachten. Das Signal besteht aus Folgen, die sich aus Zahlen mit n bit zusammensetzen. Aus diesen n bit lassen sich nun $2^n - 1$ (denn die Folge, bei der alle bit 0 sind, wird aus der Sequenz gekippt) verschiedene Folgen erstellen, welche in zufälliger

Reihenfolge aneinandergereiht werden. Der Vorteil der Verwendung der MLS besteht darin, dass es sich hierbei um ein deterministisches Signal, dessen Spektrum dieselbe Färbung aufweist wie weisses Rauschen. Der Unterschied ist jedoch, dass weisses Rauschen nur über einen langen Zeitraum gemessen wirklich weiss ist, bei der MLS ist der Zeitraum festgelegt auf die Länge der MLS.

Der theoretische Dynamikgewinn gegenüber einer Messung mit Einzelimpuls beträgt 3 dB mal der Ordnung n der Maximalfolgen (typisch: n = 10 bis 20). Die Länge einer Messung berechnet sich aus der Ordnung (jede Folge besitzt n bit) und der möglichen Anzahl Folgen:

$$S = n(2^n - 1)$$

Die Variable S bezeichnet die Anzahl Samples. Eine Sequenz der Ordnung 10 ist folglich $10 \cdot 1023 = 10230$ Samples lang. Sie ist zwar nur 0.23 Sekunden lang (S/f_s), dennoch ist die S/N-Ratio des Messergebnisses um 30dB grösser als bei einer einmaligen Messung mit einem Impuls. Möchte man die S/N-Ratio um 60dB verbessern, muss eine Sequenz der Ordnung 20 verwendet werden. Die gesamte Messzeit wächst jetzt schon auf knapp 6 Minuten.

„Nachdem die MLS in den Raum abgegeben und wieder aufgenommen wurde, wird sie mit dem bekannten Anregungssignal verrechnet. Dabei wird, bildlich gesprochen, über eine Kreuz-Korrelations-Funktion das Werte-Raster des Anregungssignals mit dem aufgenommenen, „verraumten“ Anregungssignal abgeglichen, um durch inverse Faltung die Raum-Impulsantwort zu generieren. Dieser Vorgang wird über eine rechenextensive Fast-Hadamard-Transformation (FHT) erledigt, die rein im Zeitbereich arbeitet.“ [1, Seite 28]

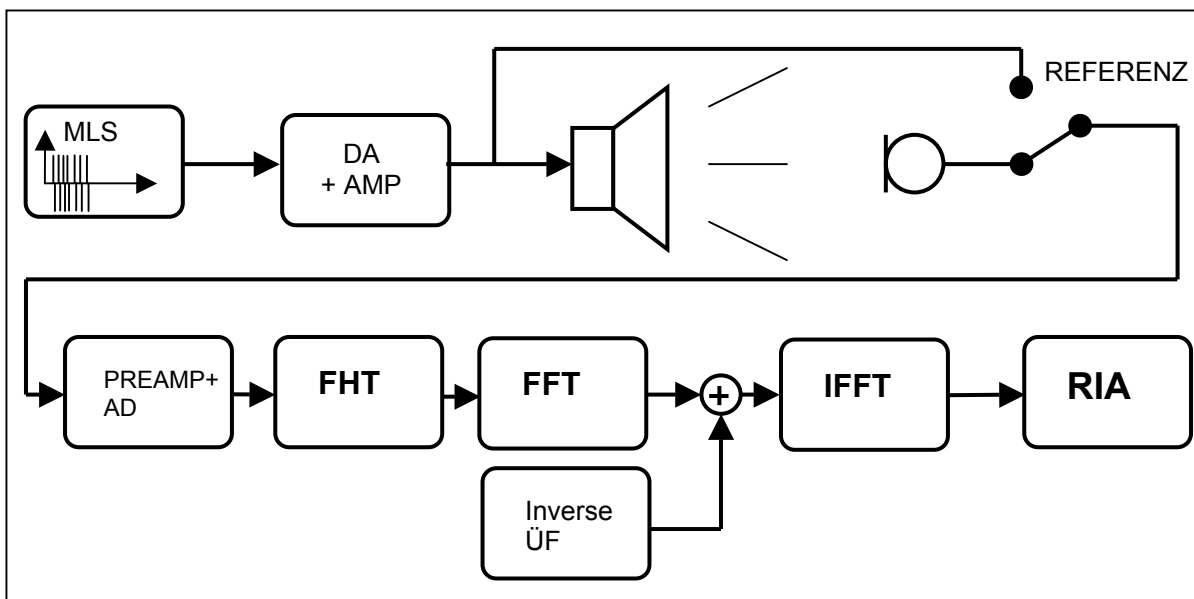


Abb. 3-3: Flussdiagramm der Messstrecke mit MLS als Anregung

Der Vorteil dieser Messmethode ist, dass man theoretisch nur einmal eine Transformation, die FHT, durchführen muss. Um die Übertragungsfehler der Messstrecke zu eliminieren, ist aber noch eine FFT mit anschliessender IFFT erforderlich.

Um ein korrektes Ergebnis zu erhalten, ist es notwendig, dass der zu messende Raum während des Messbeginns bereits eingeschwungen ist, da nur über den Zeitraum der Länge der MLS gemessen wird und sonst gewisse Anteile, die erst nach Messende im Mikrofon eintreffen würden, fehlen würden. Dies erfordert, dass zuerst ein gesamter Durchgang der MLS gesendet wird und erst bei Beginn des zweiten Durchlaufs die Messung gestartet wird.

Leider besitzen Lautsprecher die Eigenschaft, Verzerrungen zu produzieren. Dies führt zu einem erhöhten Störpegel in der resultierenden RIA. „Die Intermodulationsprodukte verteilen sich über die gesamte Länge der Impulsantwort“ [3, Seite 113], können also später nicht entfernt werden.

Es ist also nicht möglich, eine bestimmte Grösse der S/N-Ratio zu übertreffen. Selbst unter optimalen Bedingungen, wie wenig Umgebungslärm, viele Mittelungen, ist ein besserer Wert als 60-70dB [3, Seite 114] in vernünftiger Zeit nicht zu erreichen.

Will man der Tatsache Rechnung tragen, dass mit weiss gefärbten Anregungssignalen in den unteren Frequenzbereich nur eine Schwache S/N-Ratio zu erreichen ist, ist es erforderlich, das Anregungssignal mit einer Preemphasis zu versehen, welche die tieffrequenten Anteile verstärkt. Die RIA enthält dann ebenfalls diesen „verbogenen“ Frequenzgang, die Emphasis. Diese muss durch ein Filter mit zur Preemphasis invertiertem Spektrum wieder entfernt werden. Dies ist am einfachsten im Frequenzbereich zu realisieren, wofür eine FFT und anschliessend eine IFFT erforderlich ist.

Für RIA-Messungen im Freien ist diese Methode nicht geeignet, denn die Messung mit MLS ist äusserst empfindlich gegenüber Zeitinvarianzen. Wind würde verunmöglichen, dass das Sampleraster des Anregungssignals mit dem des aufgenommenen Signals übereinstimmt. Diese Verschiebung führt zu nicht vorhersehbaren Rechenartefakten in der ermittelten RIA. Diese Fehler äussern sich als häufig auftretende Peaks in der Impulsantwort (prasselndes Geräusch). Wird ein Audio-Signal mit einer solchen RIA gefaltet, machen sich diese Fehler als viel zu laute Reflexionen bemerkbar. Die empfundene Akustik hat gar nichts mehr mit dem Original zu tun.

3.2.4 FFT-Analyse mit einem Sweep

Ein Sweep entspricht einem Sinuston, dessen Frequenz mit fortlaufender Zeit stetig steigt. Handelt es sich um einen Sweep, dessen Frequenz pro Zeiteinheit eine Oktave durchläuft, also logarithmisch steigt, so spricht man von einem Sweep mit rosa Färbung. Ein Sweep, dessen Frequenz linear steigt (x Hz/s), besitzt eine weisse Färbung.

Häufig wird der lineare Sweep als Time-Stretched-Pulse bezeichnet, denn er enthält, wie ein Dirac-Impuls, alle Frequenzen in gleichmässiger Verteilung, nur dass die Laufzeiten mit der Frequenz steigen. Würde man einen Dirac-Stoss in den Frequenzbereich übertragen, dann die Gruppenlaufzeit mit steigender Frequenz erhöhen und zum Schluss das Ganze per IFFT wieder in den Zeitbereich übertragen, würde man einen linearen Sweep erhalten.

Wird ein linearer Sweep als Anregungssignal verwendet, muss die Aufnahmezeit etwas länger als der Sweep sein, da man auch den Nachhall der obersten Frequenz mitaufnehmen will. Da anschliessend eine FFT vorgenommen wird, kann die Aufnahmezeit schon im Vorfeld auf eine Länge von 2^N Samples festgelegt werden. Im Zweifelsfalls sollte sie lieber zu lang gewählt werden, um die Impulsantwort nicht abzuschneiden, bzw. es würden bei Verwendung eines ansteigenden Sweeps die oberen Frequenzen abgeschnitten. Sobald der aufgenommene Sweep nach der FFT im Frequenzbereich vorliegt, wird sein Spektrum durch das Spektrum des Anregungssweeps dividiert, sodass nur noch die Rauminformation übrig bleibt.

Das Ergebnis entspricht aber noch nicht der RIA, wie wir sie brauchen. Um zu simulieren, dass alle Frequenzen gleichzeitig abgestrahlt würden, müssen die Laufzeiten mit ansteigender Frequenz zurück (in der Zeit rückwärts) verschoben werden und zwar entsprechend den Zeiten, bei denen sie vom Sweep durchlaufen worden waren. Würde der Sweep beispielsweise 100Hz nach einer Sekunde, 300Hz nach zwei Sekunden passieren, so müsste die Laufzeit von 100Hz um eine Sekunde, die von 300Hz um 2 Sekunden verkürzt werden. So türmt sich die Summe der Amplituden aller Frequenz zum Zeitpunkt 0 zu einem Impuls auf. Nach anschliessender IFFT liegt die Raum-Impulsantwort im Zeitbereich vor.

Ein ausgesprochener Vorteil dieser Methode ist, dass Sweeps mit beliebiger spektraler Färbung verwendet werden können. Es ist somit möglich, einen grösseren Geräuschspannungs-Abstand in den Frequenzen, bei denen das menschliche Gehör am empfindlichsten ist, zu erreichen, indem sie stärker gewichtet werden. Eine stärkere Gewichtung bedeutet aber nicht, dass entsprechende Frequenzen mit einer höheren Amplitude gesendet werden, sondern dass die Anstiegsgeschwindigkeit an diesen Stellen langsamer wird, um die Energie zu erhöhen. Man kann also beliebige Sweeps als Anregung verwenden. Wichtig ist vor allem, dass die Amplitude konstant bleibt, um den niedrigen Crest-Faktor eines Sinus-Signals auszunützen.

Der Crest-Faktor beschreibt das Verhältnis zwischen der maximalen Amplitude des Signals und der im Signal enthaltenen Energie und wird in dB angegeben. Ein niedriger Crest-Faktor ist bei Messungen von Vorteil, denn es kann bei derselben Aussteuerung mehr Energie gesendet werden. Dies wiederum schlägt sich in einer erhöhten S/N-Ratio wieder.

Das beste Ergebnis liefert ein Rechtecksignal, da sich die Amplitude zu jedem Zeitpunkt auf Vollausschlag befindet. Da hier die gesamte mögliche „Energieausbeute“ ausgenutzt wird, beläuft sich der Crest-Faktor eines solchen Signals gegen 0dB. Das gilt ebenso für Puls-Signale. Weil man aber zu jedem Zeitpunkt nur eine Frequenz betrachten will, kommt für den Sweep nur ein Sinus in Frage, dessen Crest-Faktor 3dB beträgt. Dies ist ein Wert, der sich durchaus sehen lassen kann, denn würde anstatt eines logarithmischen Sweeps rosa Rauschen (welches dieselbe spektrale Färbung aufweist) verwendet werden, müsste man sich mit einem Crest-Faktor von ca. 12,5dB¹ begnügen.

Um später schon das richtige (Phasen- und Amplituden-)Spektrum zur Division zu besitzen, ist es von Vorteil, sich das Anregungssignal im Frequenzbereich zu kreieren. Dabei ist zu beachten, dass in den Frequenzbereichen, die man stärker gewichten will, die Gruppenlaufzeit steiler ansteigt, denn eine steil steigende Gruppenlaufzeit äussert sich in einem langsamen Anstieg der Frequenz im Sweep. Um die Hüllkurve im Zeitbereich konstant zu halten – man will ja den niedrigen Crest-Faktor von 3dB beibehalten –, muss die Amplitude der Frequenzbereiche mit steilem Gruppenlaufzeit-Anstieg dementsprechend niedriger sein. Ist beispielsweise ein logarithmischer Sweep erwünscht, hat das Phasenspektrum ebenfalls einen logarithmischen Verlauf. Um aber die Amplitude im Zeitbereich konstant zu halten, muss sie im Frequenzbereich um 3dB pro Oktave fallen. Allgemein gesagt: „Die Zunahme der Gruppenlaufzeit pro Frequenzschritt muss proportional zum Quadrat der gewünschten Amplitude bei dieser Frequenz sein.“ [3, Seite 115]

Werden diese Bedingungen eingehalten, kann man sich Sweeps mit beliebiger Färbung kreieren, ohne den Vorteil des geringen Crest-Faktors einzubüssen.

Die Verwendung eines Sweeps als Anregungssignal bringt einen weiteren, ganz entscheidenden Vorteil mit sich. Zu einem bestimmten Zeitpunkt wird ja nur eine Frequenz betrachtet, darum besteht der Sweep auch aus einem Sinus. Spätestens wenn das Signal vom Lautsprecher in den Raum übertragen wurde, ist das Signal kein reiner Sinus mehr, da jeder Lautsprecher, sei er noch so hochwertig, Verzerrungen produziert. Der Klirrfaktor eines Lautsprechers führt also dazu, dass das Anregungssignal verfälscht wird. Dieser Messfehler ist folglich auch in der Aufnahme vorhanden. Durch die abschliessende IFFT werden diese

¹ Dieser Wert wurde mit Hilfe der Software WAVELAB ermittelt, denn er konnte in der verwendeten Literatur nicht aufzufinden gemacht werden. Es wurde hierbei die Funktion „Globale Analyse“ auf ein rosa Rauschen der Länge von 30 Sekunden angewandt. Das rosa Rauschen wurde mit dem „Signal-Generator“ von WAVELAB erzeugt.

Fehler aber zu negativen Zeiten hin verschoben. Sie erscheinen, bevor die eigentliche Impulsantwort beginnt, darum lassen sie sich einfach von der RIA wegschneiden.

Durchläuft der Sweep nämlich gerade 100Hz (z. Bsp. nach einer Sekunde), entstehen durch die Verzerrungen Obertöne mit den Frequenzen 200Hz (k_2), 300Hz (k_3) usw. Störungen mit der Frequenz 200Hz treten auf jeden Fall auf, bevor der Sweep 200Hz (z.Bsp nach zwei Sekunden) passiert. Bevor die IFFT durchgeführt wird, werden die Gruppenlaufzeiten so verändert, als ob alle Frequenzen gleichzeitig gesendet worden wären. Bei 200Hz beträgt die Gruppenlaufzeit dementsprechend -2 Sekunden, bei 100Hz -1 Sekunden. k_2 von 100Hz, die nach einer Sekunde auftreten, erscheinen nach der IFFT also 1 Sekunde, bevor die RIA beginnt.

Die FFT-Analyse mit Sweeps hat drei Vorteile:

1. Es ist immer möglich, den niedrigen Crest-Faktor eines Sinus-Signals (3dB) voll auszuschöpfen.
2. Es ist möglich, Anregungssignale mit beliebiger Färbung zu verwenden. Man kann also frei bestimmen, wie sich die S/N-Ratio sich über das ganze Spektrum verteilt.
3. Durch die Messanlage (vor allem vom Lautsprecher) entstehende Verzerrungen entfallen und beeinträchtigen das Messergebnis nicht.

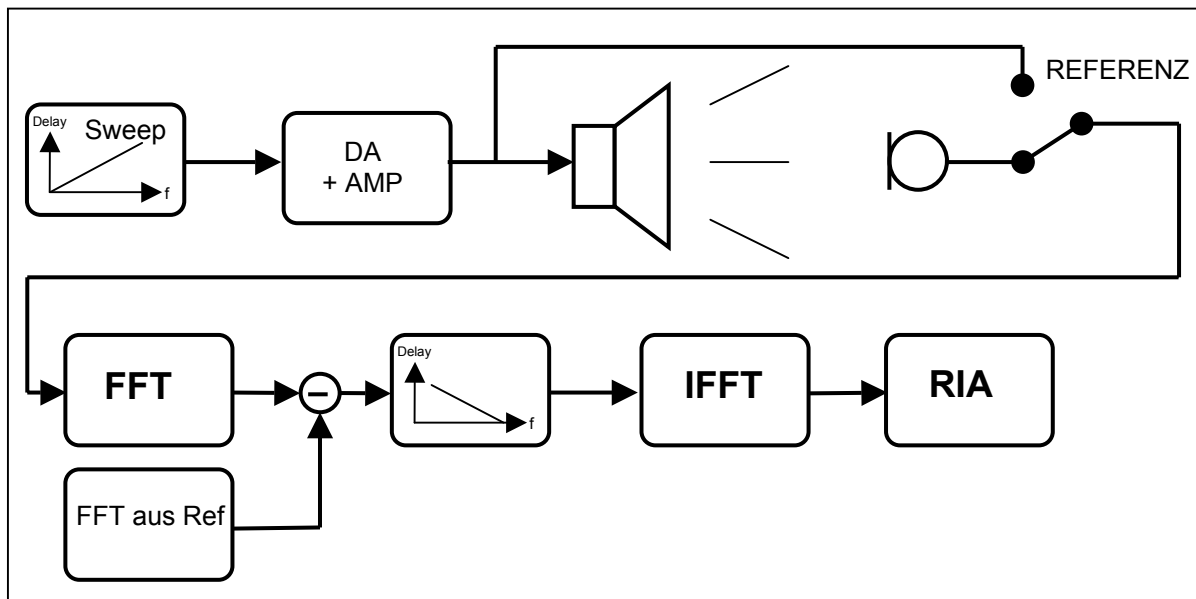


Abb. 3-4: Flussdiagramm der Messstrecke mit einem Sweep als Anregung

3.3 KORREKTUR DER ÜBERTRAGUNGSFEHLER DES MESSSYSTEMS

Die oben beschriebenen Methoden führen zu einer Verringerung des Störpegels, welcher durch äussere Einflüsse verursacht wurde. Die Übertragungsfehler der gesamten Messkette sind aber nach wie vor in der RIA vorhanden. Am meisten ins Gewicht fällt die

spektrale Verfärbung durch den Lautsprecher. Die durch das Messsystem entstehenden Fehler können entfernt werden. Hierzu muss das Messsystem selbst gemessen werden. Da, wie schon erwähnt, die Übertragungseigenschaften des Mikrofons nicht ausgeglichen werden sollen, müssen die Übertragungsfunktionen der Wandler, des Verstärkers und des Lautsprechers berücksichtigt werden.

Zur Ermittlung der ÜF (Übertragungsfunktion) der Wandler, kann der D/A-Wandler über ein Kabel mit dem A/D-Wandler verbunden werden. Anschliessend wird ein Messsignal, dessen Spektrum bekannt ist, über den D/A-Wandler ausgegeben und gleichzeitig über A/D-Wandler aufgenommen. Die Division des Spektrums des aufgenommenen Signals durch das Spektrum des Messsignals ergibt die ÜF der beiden Wandler. Dafür bietet sich ein Sweep als Messsignal an.

Zur Ermittlung der ÜF des Verstärkers kann auf die gleiche Weise verfahren werden, d.h. der Verstärker wird zwischen D/A- und A/D-Wandler eingeschleuft. Hierzu muss aber der Pegel des Verstärker-Ausgangs für den Anschluss an den A/D-Wandler entsprechend angepasst werden, sonst verabschiedet sich die Soundkarte für immer, da bei voll aufgedrehter Lautstärke gemessen werden soll (entspricht der Messsituation).

Die ÜF des Lautsprechers zu eruiieren, gestaltet sich etwas schwieriger. Da aber Lautsprecher am meisten von der Linearität abweichen, ist dieser Schritt am wichtigsten für die Korrektur. Die Frequenzgang-Messung eines Lautsprechers stellt grosse Anforderungen an die Messumgebung. Da explizit nur die Übertragungsfunktion des Lautsprechers gemessen werden soll, müssen allfällige Raumeinflüsse vermieden werden. Hierzu existieren verschiedene Ansätze.

Studios, die für solche Messungen eingerichtet sind, besitzen (fast) schalltote Räume, deren Wände meterdick mit absorbierenden Materialien ausgestattet sind. Ab einer Frequenz von etwa 200Hz sind solche Räume absolut schalltot.

Wenn ein Impuls als Anregung benutzt wird, kann auch in einem normalen Raum gemessen werden. Die Impulsantwort des Lautsprechers muss aber abgeschlossen sein, bevor die erste Reflexion beim Messmikrofon eintrifft. Dazu muss der Lautsprecher in der Mitte des Raumes positioniert werden, damit der Abstand zur nächsten Wand/Boden/Decke möglichst gross wird. Je weiter weg sich das Mikrofon vom Lautsprecher befindet, desto geringer ist die Differenz zwischen der Distanz Lautsprecher/Mikrofon und der Distanz, welche die erste Reflexion zurücklegt. Diese Methode ist ziemlich ungenau, weil das Anregungssignal relativ wenig Energie besitzt. Mit Mittelungen die Ergebnisse verbessert werden.

Ein andere Methode macht sich zu Nutze, dass ein Grenzflächen-Mikrofon von der Fläche, in der es eingebettet ist, keine Reflexion empfängt. Im quasi-freien Feld, zum Beispiel auf einer Wiese, kann man so verhindern, dass überhaupt Reflexionen zum

Mikrofon gelangen, wenn der Lautsprecher in der Höhe und das Mikrofon am Boden montiert wird. An einer absolut schallharten Fläche wird der Pegel um 6dB angehoben. Da aber selbst harte Oberflächen wie Beton für hohe Frequenzen porös sind, erhöht sich der Pegel nicht gleichmässig über alle Frequenzen, was wiederum zu einer Verfälschung des Messergebnisses führt.

Falls ein Sweep als Anregungssignal für die RIA-Messung dient, kann die Übertragungsfunktion des Lautsprechers bereits ins Anregungssignal implementiert werden, ohne Crest-Faktor-Verlust. So ist die die S/N-Ratio in jeder Frequenz maximal und eine spätere Kompensation der Lautsprecher-Übertragungsfunktion wird hinfällig.

Eine Kompensation der Übertragungsfehler soll nur dann durchgeführt werden, wenn sichergestellt ist, dass die gemessenen Übertragungsfehler genau und frei von Rauschanteilen sind, da sie sonst in die RIA miteinfließen würden.

4 RIA-MESSUNG MIT VORHANDENEM EQUIPEMENT

Es soll versucht werden, mit einfachen Mitteln, ohne teure Soft- oder Hardware zu verwenden, eine RIA zu erstellen und anhand dieser Audiosignale zu verhallen. Zur Erstellung und Verhallung kamen verschiedene Programme zum Einsatz. Für messtechnische Zwecke wurde meistens die Software WAVELAB (Version 3.0) der Firma Steinberg verwendet, welche umfangreiche Funktionen zur Signalbearbeitung, sowie zur Signalanalyse anbietet.

Da die Hardware für Faltung noch extrem teuer ist, wurde verzichtet, ein entsprechendes Gerät zu mieten, was sich auch als gar nicht nötig erwies. Die Audio-, Editing- und Mastering-Software SAMPLITUDE von SEK'D hat bereits die Funktion zur Berechnung einer Faltung implementiert¹. Der Grund für die Verwendung von zwei verschiedenen Programmen liegt nicht darin, dass das Programm SAMPLITUDE, welches Faltungshall erzeugen kann, für die RIA-Messung weniger geeignet gewesen wäre, sondern darin, dass dem Autor die Handhabung von WAVELAB vertrauter ist und ihm in einigen Punkten intelligenter schien.

Von AUDIOEASE gibt es zwar ein VST-Plugin, welches fähig ist, durch Faltung einen Hall zu erzeugen, und ermöglicht hätte, ausschliesslich in WAVELAB zu arbeiten, es wurde jedoch nur für Mac-Systeme konzipiert.

4.1 WAHL UND ERSTELLEN DES ANREGUNGSSIGNALS

Dem Autor stand bis zur Fertigstellung dieser Arbeit ein nur beschränkt geeignetes Werkzeug für die FFT und IFFT bzw. FHT zur Verfügung. In der objektorientierten Audio-Programmiersplattform PD² sind die Möglichkeiten zur Transformation in Frequenz-, bzw. Zeitbereich zwar gegeben, das Programm war dem Autor noch zu wenig vertraut, um damit sinnvoll und effizient arbeiten können.

Die Entscheidung fiel somit unweigerlich auf die am einfachsten zu durchführende Messmethode, welche als Anregung einen Dirac-Stoss verwendet. Da für die Faltung (und auch während dieser Messung) mit digitalen, sprich zeitdiskreten Signalen gearbeitet wird, ist die kürzest mögliche Zeit eines Impulses genau ein Sample lang ($22.7\mu\text{s}$ bei einer f_s von 44.1kHz).

¹ SAMPLITUDE kann als Trial-Version (entspricht Version 6.0) gratis heruntergeladen werden. Die Benutzungszeit ist in dieser Version auf 90 Tage beschränkt.

² PD (Pure Data) ist sozusagen die Gratis-Version von max/msp und läuft im Gegensatz nur auf PC-Systemen.

Es musste also ein Signal hergestellt werden, dessen Amplitude während genau eines Samples maximal (Vollausschlag) ist, sonst immer den Wert 0 besitzt. In WAVELAB kann man ab einer bestimmten Zoom-Stufe mit dem Bleistift-Werkzeug den Verlauf des Signals von Hand zeichnen. Es wurde eine Datei erstellt, welche nur Stille enthielt.

Bei genügend grosser Zoom-Stufe, werden Hilfslinien, welche das Sample-Raster darstellen, sichtbar. Mit dem Bleistift liess sich nun der Wert genau eines Sample verändern. Um sicher zu gehen, dass die Amplitude dieses Samples maximal ist, wurde dieses markiert und mit Normalisierungs-Funktion auf Vollausschlag gebracht. Ob positiver oder negativer Vollausschlag spielt für die Messung keine Rolle, sondern äussert sich nur in der Phase der RIA. Bei negativem Impuls wird die Phase der ermittelten RIA umgekehrt.

Erst später wurde herausgefunden, dass der Signalgenerator von WAVELAB bereits ein Preset besitzt, welches einen Dirac-Stoss der Länge eines Samples erzeugt¹.

Bevor aber ein Anregungssignal erzeugt wurde, musste überlegt werden, mit welcher Samplingrate man arbeiten will. Denn ein Signal kann nur mit einer RIA der selben Samplingrate gefaltet werden (Dies wird aus der Formel für die Faltung ersichtlich). Es wäre zwar möglich, die Samplingfrequenz des Signals oder der RIA anzupassen, dies sollte möglichst vermieden werden, da vor allem bei ungeraden Verhältnissen, wie 48/44,1, eine solche Wandlung unweigerlich Rechenfehler mit sich bringt. Die Wahl der Samplingrate spielt vor allem dann eine Rolle, wenn ein Dirac-Stoss als Anregung verwendet wird. Das Spektrum eines Dirac-Impulses ist dasjenige von weissem Rauschen. Konkret bedeutet dies, dass das Frequenzband zwischen 0 und 20'000Hz dieselbe Energie wie das Band zwischen 20'000 und 40'000Hz besitzt. Da aber das menschliche Gehör - grosszügig gemessen - nur Frequenzen bis 20kHz wahrnehmen kann, sind die Informationen darüber irrelevant.

Es mag Gründe geben, eine hohe Samplingrate zu verwenden; die energetische Überlegung, dass bei einer niedrigeren Samplingfrequenz mehr Energie in den hörbaren Bereich fällt, führte schliesslich zur Wahl der üblichen (CD-)Samplingfrequenz von 44,1kHz. Ausserdem ist 44,1kHz eine gängige Samplingrate.

Abschliessend wurde der Impuls mehrere Male mit Pausen dazwischen kopiert, die länger sein müssen als die vorraussichtliche RIA. So muss das Anregungs-File für mehrere Messungen nur einmal abgespielt werden. In WAVELAB ist es alternativ möglich, das Signal im Loop abzuspielen.

¹ Der Signalgenerator befindet sich unter dem Karte „Analyse“. Das Preset heisst „Single Sample (Test Early Reflections)“

4.2 MESSUNG DES ZIMMERS

Als erstes Messobjekt wurde das Zimmer des Autors gewählt. Dies war einerseits am einfachsten, da sich die gesamte Messeinrichtung (Computer mit D/A- und A/D-Wandler, Verstärker, Lautsprecher) bereits darin befinden. Andererseits war es für den Autor interessant zu erforschen, wie denn Audiosignale von dem Raum, in dem sich die vom Autor meistbenutzte Abhöre befindet, verändert werden. Es wollte nun erkundet werden, welchen Einfluss der Raum auf das Signal hat, welches schliesslich zu den Ohren gelangt.

4.2.1 Mikrofonierung

Die zu ermittelnde RIA sollte folgende Situation nachbilden: Der Hörer befindet sich in der optimalen Abhörposition zwischen den beiden Lautsprechern und schaut der Mittel-senkrechten dieses Dreiecks entlang, also genau zwischen den Lautsprechern hindurch. Der Einfachheit halber wurde ein Stereo-Mikrofon von SONY verwendet, das in die Position des virtuellen Zuhörers gebracht wurde.

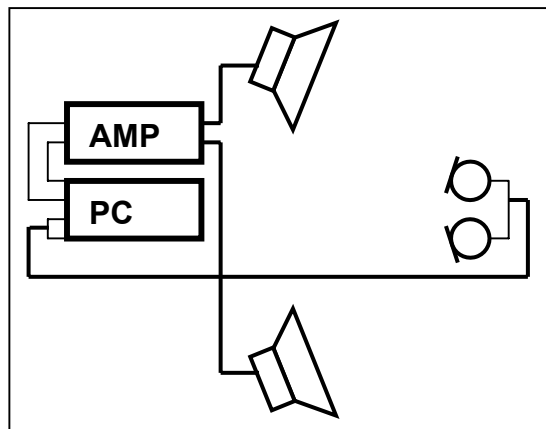


Abb. 4-1: Konstellation der Messgeräte

Das sehr kompakte Mikrofon des Typs ECM-DS70P mit Nierencharakteristik¹ wurde eingesetzt, da es nach subjektiver Beurteilung die Eigenschaft besitzt, die Umgebung räumlich sehr plastisch abzubilden. Es wurde konzipiert zur Verwendung mit den hauseigenen MD-Rekordern, DV-Cams usw. Der Geräuschspannungsabstand dieses Mikrofons ist, dem relativ niedrigen Preis von 150 SFr. entsprechend, nicht sehr hoch. Es wurde trotzdem eingesetzt, da der Umgebungslärm während der Messung das Eigenrauschen des Mikrofons deutlich übertraf.

4.2.2 Anregung

Als Anregungssignal-Sender wurde die vorhandene Abhöre „missbraucht“. Sie besteht aus einem älteren NORDMENDE-Verstärker, dessen Geräuschspannungsabstand von über 86dB keine Rolle spielt, und zwei HiFi-Lautsprecher des ebenfalls älteren Modells S-66 von JVC. Der verhältnismässig niedrige Klirrfaktor des Verstärkers von $\leq 0.03\%$ (bei 25W) kann vernachlässigt werden, da derjenige des Lautsprechers ihn bei weitem übertrifft. Bei den Lautsprechern handelt es sich um Bass-Reflex-Boxen. Daten zu den Lautsprechern, die für die Messung eventuell relevant sein könnten, wurden keine gefunden. Die Lautsprecher wurden in ihrer ursprünglichen Position belassen, weil eben die übliche Abhör-Konstellation

¹ 2x Niere. Winkel unbekannt, klingt aber sehr realistisch.

getestet werden sollte. Aus demselben Grund wurde das Anregungssignal über beide Lautsprecher gesendet. Günstiger Nebeneffekt hierbei ist, dass damit auch die Abstrahl-Energie erhöht wurde, was sich wiederum in einer besseren S/N-Ratio auswirken sollte.

Auch bei hoher Verstärkung wurde der Dirac-Impuls nicht als sehr laut empfunden, was darauf schliessen lässt, dass die Ohren empfindlicher gegenüber dem Durchschnittspegel eines Signals als dessen Peaks sind. Um eine gute S/N-Ratio zu erzielen, wurde der Volume-Regler auf Rechtsanschlag gedreht, wobei ein interessantes, dem Autor nicht schlüssig erklärbares Phänomen auftrat: Die Lautstärke nahm ab einem bestimmten Punkt beim Aufdrehen des Reglers wieder ab. Vermutlich hat es mit der enormen Flankensteilheit des Signals zu tun, welches die Lautsprecher deswegen nicht mehr sauber übertragen können. Die empfundene Lautstärke entsprach in etwa den Peaks des Signals, welches vom Mikrofon aufgenommen wurde. Der Regler wurde folglich so eingestellt, dass die maximale Lautstärke erreicht wurde.

4.2.3 Wahl der f_s und Bitrate

Sobald das Routing vom Stereo-Signal des Mikrofons zum A/D-Wandler stand, konnte die optimale Aussteuerung eingestellt werden. Da die Peaks des Input-Signals nur um etwa 0.1dB variierten, wurde ein eher kleiner Headroom von 2dB belassen.

Bevor die Messung beginnen konnte, musste zuerst das richtige Aufnahme-Format gewählt werden. Die Sampling-Rate wurde aus vorhin erwähnten Gründen auf 44.1kHz festgelegt. Wegen des eher hohen Pegels des Umgebungslärms, würde eine Auflösung von 16bit absolut ausreichen. Da aber in einem späteren Arbeitsgang in der digitalen Ebene zur Mittelung eine Senkung des Pegels der gesamten Messdatei um -30dB durchgeführt werden sollte, wurden sicherheitshalber 24bit gewählt. Die Systemdynamik mit 24bit ist um 48dB grösser als mit 16bit, das Quantisierungsrauschen nimmt somit trotz Senkung nicht zu, weil die RIA schliesslich in 16bit gespeichert wird.

Abgesehen davon, dass die S/N-Ratio der fertigen RIA zu klein sein wird, um 24bit zu rechtfertigen, kann der Faltungshall-Erzeuger von SAMPLITUDE der Version 6.0 nur RIAs einer Auflösung von 16bit verarbeiten.

4.2.4 Messung

Sobald all diese Einstellungen vorgenommen worden waren, konnte mit der Messung begonnen werden. Weil später per Mittelung der Störpegel der RIA abgesenkt werden sollte, wurden in einem Zug gleich mehrere Messungen durchgeführt.

Um eine Messung dieser Art vorzunehmen, ist eine full-duplex-fähige¹ Soundkarte notwendig, da während der Aufnahme das Anregungssignal abgespielt werden muss. Es

¹ Eine full-duplex-fähige Soundkarte kann gleichzeitig abspielen und aufnehmen.

macht Sinn, mit demselben Gerät gleichzeitig abzuspielen und aufzunehmen, denn so werden D/A- und A/D-Wandler vom selben Takt getriggert. Dies hat den Vorteil, dass sich die RIAs jeder einzelnen Messungen identisch ins Sample-Raster einfügen. Würden zur Aufnahme und Wiedergabe getrennte Geräte mit jeweils eigenem Takt verwendet werden, würden die Amplitudenverläufe der einzelnen Messungen leicht differieren, bedingt durch die leichte Verschiebung der beiden Takte. Mit der hier angewandten Methode wurde gewährleistet, dass sich die Messung nach einer bestimmten, von der Länge des Loops des Anregungssignals abhängigen Anzahl Samples stetig wiederholt.

Die gesamte Messdatei, welche mehrere Messungen enthält, konnte dann in gleich lange Teile, die gleich viele Samples lang sind wie die Anregung, geschnitten werden, um die einzelnen, für die Mittelung benötigten Messungen zu erhalten.

Abschliessend wurde einen späteren Vergleich mit denselben Einstellungen und Positionierungen von Mic/LS ein Sprachsample (die Datei „Stimme-Orig“) in den Raum gesendet und aufgenommen.

4.2.5 Störpegelverringern durch Mittelung

Ohne die Möglichkeit, mit einer speziell für Messungen entwickelten Software zu arbeiten, stellt dieser Arbeitsgang den mühseligsten Part dar. Im Lauf dieser Arbeit konnte keine bessere und effizientere Methode gefunden werden als die nachfolgend beschriebene.

Wie schon erwähnt, führt eine Mischung zweier Messergebnisse des identischen Systems zu einem Pegelanstieg von 6dB, der Anteil des Umgebungslärms steigt aber nur um 3dB. Die einzelnen Messungen wurden gut ausgesteuert, eine Mischung zwei oder mehrerer Messungen würde folglich zu einer Übersteuerung führen. Deshalb musste später der Pegel um einen Wert, der abhängig ist von der Anzahl Mittelungen, gesenkt werden. Damit Aufwand und Ergebnis in einem vernünftigen Zusammenhang zueinander stehen, wurde beschlossen, die S/N-Ratio um 15dB zu erhöhen. Das bedeutet, die gewünschte RIA muss aus $2^{(15/3)}=32$ Messungen gemittelt werden.

Damit die fertige RIA den gleichen Pegel besitzt wie der Durchschnitt der einzelnen Messungen, mussten alle Messungen bzw. die ganze aufgenommene Datei um 30dB ($15/3 * 6dB = 30dB$) gesenkt werden. Da bei jeder Verdoppelung der Mittelungen das Ergebnis doppelt so laut (entspricht 6dB) wird, entspricht der Wert der Absenkung der Anzahl Verdoppelungen multipliziert mit 6dB. Es ist deshalb anzuraten, das Dateiformat vor der Senkung auf 24Bit zu ändern, falls das noch nicht schon der Fall ist. Während dieser Messung wurde die Möglichkeit genutzt, schon mit 24Bit Auflösung aufzunehmen.

Als nächstes wurden die einzelnen Messungen getrennt und als einzelne Dateien gespeichert. Hierzu wurde die gesamte Datei durchgehört und die 32 Messungen, welche am wenigsten Umgebungsgeräusche enthielten, mit einem Marker versehen. Dieser

Vorgang geschah in der Software WAVELAB. Die 32 ausgewählten Messungen wurden anschliessend sample-genau ausgeschnitten. Bei der Mittelung ist es äusserst wichtig, dass die Messungen absolut deckungsgleich miteinander gemischt werden. Zeitlich verschoben dazugemischte Messungen würden zu einer Verwischung der gewünschten RIA führen und die Erhöhung der S/N-Ratio würde ausbleiben.

Es gibt zwei Wege, die Messungen sample-genau auszuschneiden:

1. Es wird immer im selben Abstand geschnitten. Der Abstand entspricht der Länge des Loops des Anregungssignal. Praktisch wäre die Möglichkeit, diesen (einfachen) Vorgang zu automatisieren.
2. Die hier angewandte Methode war die „optische“: Da alle Messungen fast identisch sind und der Beginn jeder einzelnen RIA gut sichtbar ist, konnte der jeweilige Schnittpunkt von Auge gefunden werden. Wichtig dabei ist, einen gut identifizierbaren Schnittpunkt zu wählen, beispielsweise 2 Samples vor dem ersten Impuls¹. Andernfalls besteht die Gefahr, dass man nicht immer am selben Punkt schneidet und die einzelnen RIAs nicht deckungsgleich gemischt werden. Hier erweist es sich als Vorteil, wenn Input und Output gleich getaktet wurden, denn somit sehen alle Messungen identisch aus und der Schnittpunkt ist einfacher zu lokalisieren. Glücklicherweise ist WAVELAB mit einem dynamischen Zoom ausgestattet. Mit einer Mausbewegung kann auf Sample-Ebene gezoomt und der Beginn des Bereichs, der ausgeschnitten werden soll, markiert werden. So ist man relativ schnell, auch 32 Bereiche sample-genau auszuschneiden.

Alle 32 Einzelmessungen mussten schliesslich miteinander vermischt werden. In WAVELAB geschah dies mit der Menu-Funktion *Einfügen spezial*→*Mischen*. Bei Anwendung dieser Funktion muss sich der Cursor an Datei-Anfang befinden, um Deckungsgleichheit der vermischten Dateien zu gewährleisten. Wenn der Peak der daraus resultierenden RIA denselben Pegel besitzt wie die Peaks der einzelnen Messungen vor der Absenkung, kann davon ausgegangen werden, dass die einzelnen Dateien deckungsgleich gemischt wurden. Der Pegel der Störgeräusche sollte nun um den angestrebten Wert leiser sein als vor der Mittelung, was bei dieser Messung auch der Fall war.

4.2.6 Kompensation des Lautsprecher-Übertragungsfehlers

Auf eine Kompensation wurde in dieser Messung verzichtet. Neben dem SONY-Mikrofon, über dessen Frequenzgang keine Daten herausgefunden wurden, stand ein Kondensator-Mikrofon von SENNHEISER mit Nierencharakteristik, es handelt sich um das K6,

¹ Auch hier trat ein interessantes Phänomen auf: Jeweils vor dem eigentlichen Ausschlag ist ein kurzes „Vorschwingen“ sichtbar, vermutlich verursacht durch die Trägheit des Lautsprechers.

zur Verfügung. Es wurde eine Messung der Lautsprecher-Übertragungsfunktion mit diesem Mikrofon durchgeführt. Als Anregung diente ebenfalls ein Impuls.

Es stellte sich aber heraus, dass die Impulsantwort des Lautsprechers noch nicht abgeschlossen war, als die erste Reflexion beim Mikrofon eintraf. Trotzdem wurden 16 Mittelungen durchgeführt, um die S/N-Ratio um 12dB zu reduzieren. Der Zeitpunkt wurde berechnet, wann die erste Reflexion eintreffen sollte, da sie nicht deutlich sichtbar war. Ab diesem Punkt wurde die Lautsprecher-Impulsantwort weggeschnitten.

Eine versuchsweise Kompensation ergab, dass die resultierende RIA sehr dumpf klang und deshalb nicht weiter verwendet werden konnte. Eine nachträgliche Betrachtung des K6-Datenblatts deckte das Problem auf; das K6 hebt die Frequenzen um 10kHz bis zu 6dB an, was den Lautsprecher für die Messung höhenlastiger erscheinen lässt, als er wirklich ist. Weil die Impulsantwort bei Eintreffen der ersten Reflexion nicht abgeschlossen war, konnte sowieso nicht von einer genauen Messung ausgegangen werden.

Abschliessend wurde die Hallfahne der RIA von dem Punkt ausgefadet, wo der Pegel nicht mehr weiter sank (Rauschen).

Die Idee dieser Messung war ein Vergleich zwischen einem durch Faltung erzeugten Signal und einem „echten“ verhallten Signal. Da beide Signale den Messfehler der gesamten Messkette enthalten, ist eine Kompensation der RIA für diesen Vergleich überflüssig. Das „echte“ Signal wurde ja über dieselbe Messstrecke in den Raum gesendet und wieder aufgenommen wie die RIA, die zur Verhallung diente.

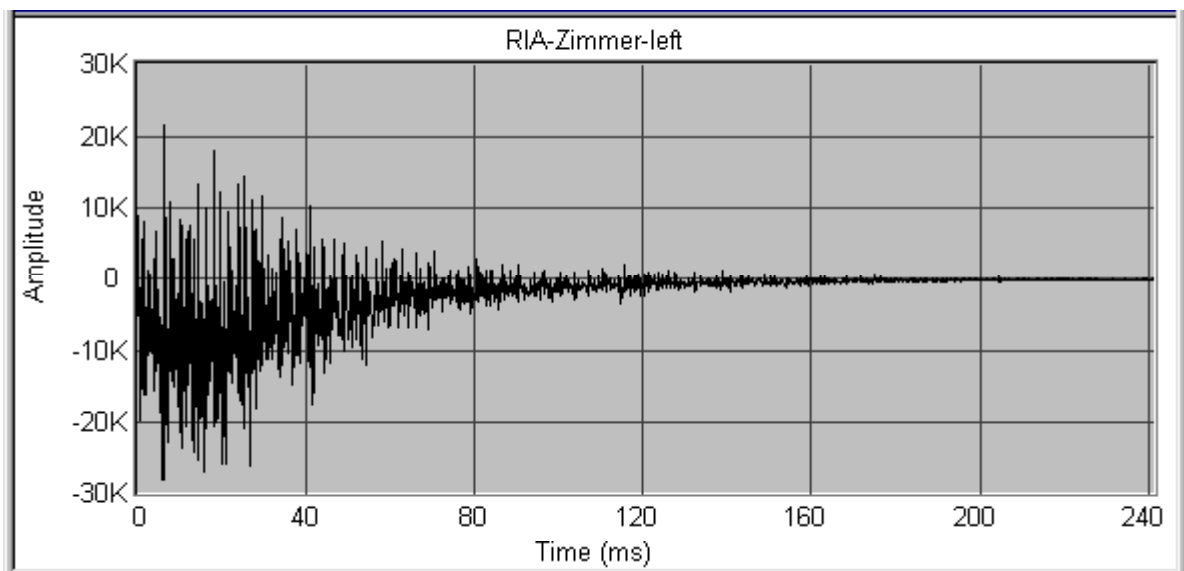


Abb. 4-2: Linker Kanal der fertigen RIA-Zimmer¹

¹ Der DC-Offset der RIA betrug nur 0,02%, entgegen der Erwartung, die aus dieser Grafik entsteht.

4.2.7 Faltung

Weil das Gehör bei Sprache sehr sensibel auf klangliche Veränderungen ist, wurde beschlossen, zur Faltung mit der RIA eine Datei mit Sprache zu verwenden (Datei „Stimme-Orig“: CD-Track 1). Wie schon erwähnt, kann in SAMPLITUDE eine Faltung durchgeführt werden. SAMPLITUDE bietet hierzu zwei Möglichkeiten unter dem Menü *Effects*¹:

Einerseits arbeitet der *Room-Simulator* mit Faltung, bietet aber noch andere Parameter zur Modifikation des resultierenden Halls an.

Weiter gibt es die Funktion *Convolution* (Faltung), mit der man zur Faltung eine geöffnete Datei wählen kann und die nur die Parameter *Original* (Originalanteil) und *Result* (Pegel des Resultats) enthält.

Für diese Arbeit wurde immer die Funktion *Convolution* gewählt, da auf weitere Parameter wie HF-Damp usw. verzichtet wurde, denn das Ergebnis sollte ausschliesslich von der RIA determiniert sein. Es wurde herausgefunden, dass die Beschriftung des *Result*-Reglers keine mathematisch stimmige Angabe darstellt. Er wurde vermutlich so konzipiert, dass bei 0dB das Ergebnis der Faltung mit einer echten (also sehr dichten) RIA nicht übersteuert. Faltet man ein Signal mit einem voll ausgesteuerten Dirac-Impuls bei 0dB, müsste das Ergebnis identisch mit dem Original und dementsprechend gleich laut sein, es ist aber wesentlich leiser. Falls mit RIAs von Effekt-Geräten, die eine viel geringere Dichte haben, gearbeitet wird, kann dieser Regler ohne weiteres auf das Maximum von +20dB eingestellt werden.

Für die Ergebnisse dieser Arbeit wurde der *Original*-Anteil immer ausgeblendet (*Original* auf 0%), denn die RIA selbst enthält den Direktschall-Anteil, sprich den ersten, beim Mikrofon eintreffenden Impuls. So bestimmt die RIA eigentlich schon das richtige Verhältnis von Direkt- und Effekt-Anteil.

Es ist nicht möglich, den Pegel des Ergebnisses vorauszusehen, da dieser abhängig von der Dichte der RIA ist. Deshalb wurden die verhallten Signale für den Vergleich abschliessend normalisiert.

4.2.8 Vergleich mit dem Original

Bei einem A/B-Vergleich zwischen dem kurzen Sprach-Track, der effektiv in das Zimmer gesendet wurde, und demselben mit der gemessenen RIA gefalteten Track, wurde der „Echte“ nach subjektiver Beurteilung des Autors als voller klingender empfunden; sie klingen aber beide sehr ähnlich. Eine FFT-Analyse beider Signale ergab, dass das „echte“ Signal in den tiefen Frequenzen etwas mehr Energie besitzt als das gefaltete. Das bedeutet, die RIA besitzt neben den Übertragungsfehlern der Messkette zusätzlich eine Ungenauigkeit im Frequenzgang. Der Vergleich kann auf der beigelegten CD angehört werden (Track-Nr. 2).

¹ Die hier angebotenen Effekte arbeiten alle destruktiv und nicht in Echtzeit.

4.3 EINFLUSS VERSCHIEDENER PARAMETER AUF DIE RIA

Im Flur wurden gleich mehrere Messungen vorgenommen. Da zwei völlig unterschiedliche Mikrofone, das ECM-DS70P von SONY und das K6 von SENNHEISER, zur Verfügung standen, wurde die selbe Messung mit beiden Mikrofonen durchgeführt. Dabei wurde Position von Mikrofon/Lautsprecher jeweils nicht verändert. Mit dem K6 wurde eine Messung mit einem „braunen“ Impuls als Anregung durchgeführt. Alle Messungen geschahen mit geschlossenen Türen. Darum wurde unter den selben Voraussetzungen die RIA des Flurs mit offenen Türen gemessen, insgesamt sechs Türen. Nach jeder Messung wurde für einen Vergleich, wiederum mit unveränderter Konfiguration, das Sprachsample „Stimme-Orig“ abgespielt und aufgenommen.

Folgende Tabelle stellt eine Liste der vorgenommenen Messungen dar, wobei jeweils ein Parameter der Grundeinstellung (RIA A) variiert wurde:

	Spektrum der Anregung	Mikrofon	Türen	Hörbeispiel CD-Track
RIA A	braun	K6	geschlossen	3
RIA B	weiss	K6	geschlossen	4
RIA C	braun	K6	offen	5
RIA D	braun	ECM-DS70P	geschlossen	6

Tab. 4-1: Parameter der verschiedenen Messungen

4.3.1 Anregungssignal

Um eine andere Energieverteilung auszuprobieren, wurde alternativ zum Dirac-Impuls ein Impuls mit brauner Färbung verwendet. Ein solcher Impuls entspricht im Prinzip der plötzlichen Änderung von 0 auf DC Vollausschlag. Da kein Lautsprecher dauerhaft Gleichspannung übersteht und sehr tiefe Frequenzen für eine RIA nicht relevant sind, kann ein entsprechend gefiltertes Signal eingesetzt werden. Zur Erzeugung wurde mit dem Signalgenerator von WAVELAB eine Rechteck-Welle mit 0.1Hz generiert, bei der die zweite Hälfte weggeschnitten wurde. Anschliessend wurden mit einem Tiefpass der Grenzfrequenz 40Hz die tieffrequenten Anteile weggefiltert.

4.3.2 Mikrofonierung und LS-Position

Weil nur ein K6 zur Verfügung stand, mussten linker und rechter Kanal getrennt gemessen werden. Da der Autor nach eigener Beurteilung davon ausgeht, dass die Nieren des ECM-DS70P nicht geradeaus nach vorne schauen, wurde eine Mischung aus X/Y und

A/B gewählt. Das Mikrofon wurde im ersten Drittel des Flurs auf Kopfhöhe so positioniert, dass es in Richtung letztes Drittel schaute, wo sich der Lautsprecher befand. Für die Messung des rechten Kanals wurde es um 30° nach rechts gedreht und umgekehrt. Für die Messung mit dem ECM-DS70P wurde eine mittlere Position gewählt.

Da der Lautsprecher gerichtet abstrahlt, wurde er so flach auf den Boden gestellt, dass er nach oben gerichtet lag. Es wurde auf diese Weise versucht, möglichst in alle Richtungen des Raumes abzustrahlen.

4.3.3 Messung und Mittelung

Es wurde wie in 4.2.4 geschildert vorgegangen, dabei einmal die Türen geöffnet, das Mikrofon oder das Anregungssignal gewechselt. Es wurden jeweils genug Messungen gemacht, um 32 Mittelungen vorzunehmen. Da bei Messungen mit dem K6 für linken und rechten Kanal eigene Messdateien vorlagen, mussten diese zuerst zu einer Stereo-Datei zusammengefasst werden.

Da bei der Messeinrichtung darauf geachtet wurde, dass beide Mikrofon-Positionen die gleich Distanz zum LS-Mitteltöner besaßen, konnten die beiden Dateien so zueinander verschoben werden, dass die ersten Impulse der RIA übereinander lagen. So musste der ganze Schneid-Vorgang pro RIA nicht zweimal durchgeführt werden. Nach der Mittelung war die RIA B fertiggestellt.

4.3.4 Spektrum-Kompensation

Die anderen drei RIAs enthielten immer noch die spektrale Färbung des Anregungssignals, welche in einem weiteren Schritt kompensiert werden musste. Hierzu musste ein Korrektur-Spektrum erstellt werden. Mit dem Signalgenerator wurde, als Signal mit brauner Färbung, ein braunes Rauschen erzeugt. Danach wurde dieses Signal in SAMPLITUDE einer FFT-Analyse unterzogen. Das daraus resultierende (braune) Spektrum wurde invertiert und abgespeichert. Danach konnten die Spektren der übrigen RIAs mit dem invertierten Spektrum des braunen Rauschens multipliziert werden.

4.3.5 Ergebnisse

Der Autor konnte nach Gehör keine Unterschiede feststellen zwischen den Ergebnissen, welchen unterschiedliche Anregungssignale zu Grunde lagen (RIA A und B). Im Ausklang der RIA A waren höherfrequente Störanteile vorhanden gegenüber der RIA B, was mit einer durch die Kompensation verursachten Verlagerung des Rauschens in die höheren Frequenzen erklärt werden kann. Dagegen sind diese Störanteile in RIA A leiser als in RIA B.

Für einen Vergleich der Spektren wurde die Datei „Stimme-Orig“ mit beiden RIAs (A und D) verhallt und beide Ergebnisse mit dem Original verglichen. Das Spektrum aus RIA D wies einen minimalen Anstieg zu den tiefen Frequenzen hin auf gegenüber dem Original. Das Ergebnis aus RIA A wies praktisch keine Unterschiede gegenüber dem Original auf. Ein Kompromiss zwischen braun und weiss, ein rosa Anregungssignal, hätte vermutlich ein optimierteres Ergebnis liefern können.

Bemerkenswert ist, wie RIA C einen deutlich kürzeren Nachhall besitzt als RIA A. Dies zeigt deutlich, wie sich die Raumakustik mit relativ einfachen Mitteln weitgehend gestalten lässt.

Bei der Verhallung mit RIA D werden die Unterschiede im Frequenzgang der beiden Mikrofone gut hörbar. Beim Anhören von Track 5 (RIA D) hat man aber eher das Gefühl, inmitten des Raumes zustehen, obwohl es nach „Fernseher-Lautsprecher“ klingt.

5 TEST: KANN FALTUNG HERKÖMMLICHEN HALL KOPIEREN?

Die Vermutung, dass bei der Möglichkeit, Räume zu messen, es auch möglich sein sollte, eine „RIA“ eines Hallgerätes zu ermitteln, drängt sich zwangsläufig auf. Falls dies möglich ist, würde das bedeuten, dass man – zumindest in der Theorie – mit einem Faltungshall und entsprechenden RIAs ausreichend bedient wäre und auf weitere Hallgeräte verzichten könnte.

In diesem Abschnitt soll getestet werden, ob der durch Faltung erzeugte Hall identisch ist mit dem Originalhall. Da die Theorie besagt, dass ein LTI-System eindeutig durch seine Impulsantwort beschrieben wird und der Autor annimmt, dass Hallgeräte die Bedingungen eines LTI-Systems erfüllen, soll versucht werden, diese Vermutung mit Hilfe eines Dateivergleichs zumindest empirisch zu bestätigen.

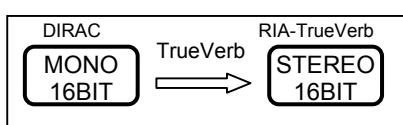
5.1 MESSUNG DER IMPULSANTWORT *TRUEVERB*

Die Impulsantwort eines digitalen Hallgerätes zu messen, gestaltet sich relativ einfach. Falls das zu messende Objekt auf digitalem Wege mit der Anregung gespiesen werden kann, müsste die resultierende Impulsantwort völlig frei von allfälligen Übertragungsfehlern sein. Eine spätere Nachbearbeitung erübrigt sich in diesem Fall.

Da kein externes Hallgerät zur Verfügung stand, wurde ein Hall-PlugIn, der *TrueVerb* von WAVES, der als relativ hochwertig gilt, als erstes Messobjekt gewählt. Es spielt jedoch keine Rolle, ob ein natives PlugIn oder eine externes Gerät gemessen wird, da das Funktionsprinzip der Hallerzeugung grundsätzlich dasselbe ist.

Als Anregung diente wiederum der Dirac-Impuls, sodass das Ergebnis nicht weiter bearbeitet werden musste. In WAVELAB wurde der *TrueVerb* mit der Funktion *APPLY* in die Mono-Anregungsdatei gerechnet, sodass schliesslich eine Stereo-Datei, die Impulsantwort, entstand. Um die Einstellungen des PlugIns reproduzierbar zu halten, wurde das Preset „Studio A“ unverändert übernommen. Sowohl der Effekt-, als auch der Originalanteil sind bei diesem Preset auf 0dB (also ohne Absenkung) eingestellt.

Trotzdem fiel der erste Impuls der Impulsantwort, also der Originalanteil, leiser aus als Vollausschlag. Zudem wies er „Verformungen“ auf, was darauf schliessen lässt, dass



TrueVerb auch den Originalanteil verändert, sprich filtert.
Die erhaltene Impulsantwort wurde gespeichert.

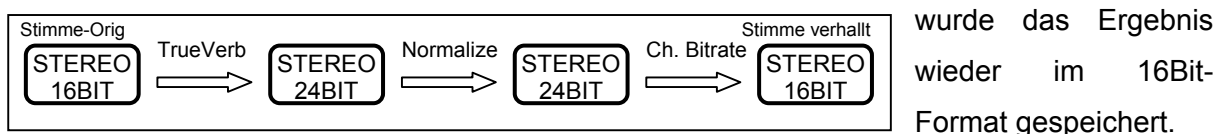
Abb. 5-1: Erzeugung der Impulsantwort

5.2 ERZEUGUNG DER VERGLEICHSDATEIEN

5.2.1 Stimme, verhallt durch *TrueVerb*

Für diesen Test wurde das schon vorher eingesetzte Sprachsample „Stimme-Orig“ verwendet. Diese Datei wurde zwar im Stereo-Format gespeichert, ist aber eigentlich mono (links und rechts sind identisch). Um einen späteren Dateivergleich zu vereinfachen, wurde der linke Kanal am Anfang mit einem Dirac-Impuls versehen, um allfällige, bei der Verhallung auftretende zeitliche Verschiebungen sichtbar zu machen. Zugeich stellt der Impuls das lauteste Sample dar und dient somit als Referenz für eine spätere Normalisierung. Da sowohl bei der Verhallung mit *TrueVerb*, als auch bei der Faltung keine Voraussagen über den Pegel des Ergebnisses gemacht werden können, ist eine anschließende Normalisierung das einfachste Mittel, identische Pegel zu erhalten (Es können nur Dateien mit gleichem Pegel verglichen werden).

Der Hall wurde mit einer Tiefe von 24Bit in die Datei verrechnet, um den durch die Normalisierung entstehenden Fehler möglichst gering zu halten. Nach der Normalisierung



wurde das Ergebnis wieder im 16Bit-Format gespeichert.

Abb. 5-1: Erstellung der Vergleichsdatei mit *TrueVerb*

5.2.2 Stimme, verhallt durch Faltung

Weil nur RIAs mit 16Bit verwendet werden können, war es notwendig, die Ursprungsdatei „Stimme-Orig“ ins 24Bit-Format zu wandeln, um ein Ergebnis mit 24Bit zu erhalten¹. Danach wurde die Faltung mit der aus 5.1 erhaltenen Impulsantwort durchgeführt. Dazu wurde die Funktion *Convolution* angewandt, wobei der *Result*-Regler auf das Maximum (+20dB), der *Original*-Regler auf 0% eingestellt wurde. Das Ergebnis, welches ebenfalls in 24Bit vorlag, wurde anschliessend normalisiert und im 16Bit-Format gespeichert.

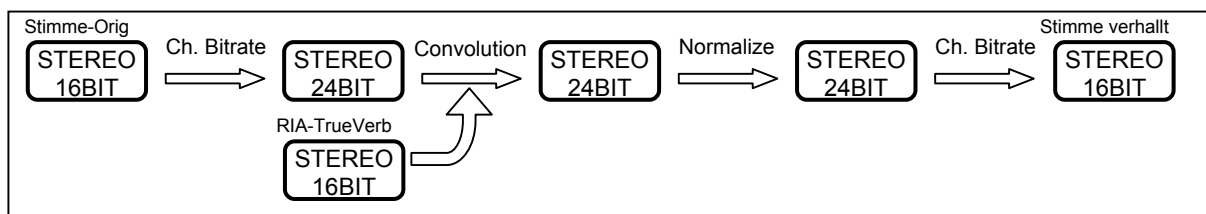


Abb. 5-2: Erstellung der Vergleichsdatei durch Faltung

Die Ergebnisse können auf der beigelegten CD angehört werden: Nr. 7

¹ In SAMPLITUDE 6.0 ist es möglich, 24Bit-Signale mit 16Bit-RIAs zu falten, aber nicht umgekehrt.

5.3 VERGLEICH

5.3.1 Voraussetzungen für einen Vergleich

Um einen Vergleich der beiden Dateien durchführen zu können, müssen zwei Bedingungen erfüllt werden:

- Beide Dateien müssen exakt gleich laut sein.
- Beide Dateien müssen sich samplegenau im gleichen zeitlichen Raster befinden.

Angenommen, beide Dateien sind identisch, so müssten sie auch nach einer Normalisierung identisch sein, womit die erste Bedingung erfüllt wäre. Dank des gesetzten Impulses in der Ursprungsdatei war es leicht sich zu vergewissern, dass keine zeitliche Verschiebung aufgetreten war, weder durch die Faltung, noch durch die Verhallung mit *TrueVerb*. Einem Vergleich stand also nichts mehr im Weg.

In WAVELAB ist es möglich, zwei Dateien miteinander zu vergleichen und eine sogenannte Delta-Datei zu erstellen. Die Delta-Datei wird aus der Differenz der beiden Vergleichsdateien erstellt, d.h. zwei identische Dateien ergeben eine Delta-Datei, welche nur Stille enthält.

5.3.2 Deutung des Ergebnisses

Die Delta-Datei aus dem Vergleich aus 5.2 (CD-Track 8) besitzt (ohne den Anfang) einen durchschnittlichen Pegel von unter 60dB. Die Art des Signals ist aber kein Rauschen, sondern hört sich bei entsprechender Verstärkung (Vorsicht, Regler danach wieder runterdrehen!!!) an wie der reine Hallanteil. Durch das ganze Prozedere (Faltung, Normalisierung) wurde das Ergebnis zwar verfälscht, aber in einer Form, die das Endergebnis nicht hörbar stört. Der Nachhall des Dirac-Impulses war in der Delta-Datei aber immer noch zu hören, und zwar nicht nur im linken, sondern auch im rechten Kanal. Bei genauerer Betrachtung der beiden Vergleichsdateien fiel auf, dass bei der *Trueverb*-Datei der Nachhall des Diracs in beiden Kanälen zu hören war, bei der Faltungs-Datei aber nur im linken.

Diese Tatsache legt folgende Vermutung nah: Der *TrueVerb*, welcher wie gesagt als hochwertig gilt, arbeitet sozusagen mit vier Prozessoren. Je zwei werden vom linken und zwei vom rechten Input gespiesen. Die vier resultierenden Signale werden dann wieder zu einem Stereo-Signal zusammengemischt. Diese Verknüpfung wird dem Umstand gerecht, dass das linke Signal den Hall vom rechten mitbeeinflusst wird und umgekehrt. Da aber das zu verhallende Signal „Stimme-Orig“, abgesehen vom Dirac, mono ist, erzeugte die Faltung das gleiche Ergebnis wie der *TrueVerb*.

Fazit: Mit der von SAMPLITUDE angebotenen Funktion *Convolution* ist es im Prinzip nur zulässig, Mono-Dateien mit Impulsantworten von Stereo-Hallgeräten zu falten. Für die Verhallung von Stereo-Signalen wären zwei Stereo-Impulsantworten erforderlich, es müssten also getrennte Messungen für linken und rechten Input des Hallgerätes durchgeführt werden.

Um den *TrueVerb* korrekt „nachahmen“ zu können, müssten insgesamt vier Faltungen gleichzeitig durchgeführt werden können. Dagegen ist es mit der Faltungsfunktion möglich, zwei Mono-Signale gleichzeitig zu verhallen, ohne das sie sich gegenseitig beeinflussen. Ob man in der Praxis dafür Verwendung findet, sei dahingestellt.

Folgende Tabelle zeigt eine Gegenüberstellung, welche Konfigurationen (Input→Output) zur Verhallung zulässig sind:

TRUEVERB	mono→stereo stereo→stereo	FALTUNG IN SAMPLITUDE	mono→mono mono→stereo 2xmono→2xmono
----------	--	--------------------------	--

Tab. 5-1: Gegenüberstellung der möglichen Konfigurationen

6 ABSCHLIESSENDE ÜBERLEGUNGEN ZUR PRAXIS

Obwohl die Technik zur Hallerzeugung durch Faltung relativ neu ist, hat sie ihre Tauglichkeit bewiesen. Die dem Autor bekannten Geräte arbeiten alle 4-kanalig, sodass auch echter Sterero-Hall möglich ist. Beim Kauf von Hardware wird sogar geeignete Software, um selbständig RIAs zu messen, und eine Palette diverser RIAs mitgeliefert. Wie in dieser Arbeit gezeigt wurde, ist es auch mit relativ einfachen Mitteln machbar, einigermassen brauchbare RIAs zu erhalten. Die Frage ist, ob man bereit ist, die nötige Zeit dafür aufzuwenden.

Im Gegensatz zu herkömmlichen Geräten hat man beim Faltungshall nicht dieselbe Vielfalt an Parametern zur Verfügung. Man ist also auf einen gewissen Grundstock an RIAs angewiesen, um flexibel damit arbeiten zu können. Es ist mittlerweile auch möglich, RIAs käuflich zu erwerben, sei es im Internet oder auf CD-ROM.

Faltungshall-Geräte sind immer noch sehr teuer, die Qualität des erzeugten Halls jedoch wird aus technischer Sicht nur von der Sampling- und Bitrate begrenzt. Mit entsprechenden RIAs steht diese Technik der herkömmlichen Hallerzeugung in nichts nach, was Dichte, „Wärme“ usw. betrifft.

Ob diese Technik für Musikproduktionen praktikabel ist, hängt zum grossen Teil von der Arbeitsweise des Produzenten ab. Falls oft mit nur wenig veränderten Presets gearbeitet wird, gelangt man mit Faltungshall wahrscheinlich schnell zu guten Ergebnissen. Für Soundtüftler, die Stunden damit verbringen, solange an den Parametern zu schrauben bis alles stimmt, bietet Faltungshall eventuell zu wenig Eingriffsmöglichkeiten. Andererseits können beliebige Signale als RIAs missbraucht werden, was ein weites Experimentierfeld eröffnet.

Im Film-Bereich ermöglicht die Hallerzeugung durch Faltung eine ganz neue Arbeitsweise. Falls es nämlich nicht möglich ist, den O-Ton zu verwenden, muss nur die RIA vom Set¹ gemessen werden, um die spätere Synchronisation akustisch in die Umgebung einbetten zu können. Auch künstliche Geräusche können im Nachhinein hinzugefügt werden, ohne Zeit für langwierige Einstellungen eines Hallgerätes aufwenden zu müssen. Viele akustische Umgebungen (im Auto, Häuserschluchten, sehr kleine Räume) lassen sich mit herkömmlichen Geräten nicht oder nur schwer realisieren. Laut der Diplomarbeit von Andreas Turnwald [1] ist die Verwendung von Faltungshall sogar wirtschaftlicher als das Arbeiten mit herkömmlichen Geräten. Für die Synchronisation einer Fremdsprache müssen ebenfalls nur die RIAs mitgeliefert werden, und eine Anpassung des Gesprochenen kann relativ schnell geschehen.

¹sofern es sich bei der Umgebung nicht um Kulissen handelt

7 ANHANG

7.1 QUELLENVERZEICHNIS

- [1] ANDREAS TURNWALD 2002 Diplomarbeit an der FH Düsseldorf zum Thema *Anwendung von Raumimpulsantworten in der Filmtone-Produktion*, Internet (Okt. 2002): www.turnwald.de/andreas/RIAFilm.pdf
- [2] PROF. DR. P. JENSCH (Lose Blattsammlung) *LTI-Systeme*, Internet (Okt. 2002): http://pi.informatik.uni-oldenburg.de/lehre/WS02-SigBildVer/1_LTI-Systeme.pdf
- [3] SVEN MÜLLER 2001 *Messen von Raum-Impulsantworten* Production Partner, Heft 12/2001 S. 106ff., MM-Musik-Media-Verlag, Köln
- [4] ISO 3382 1997 *Norm für die Messung der Nachhallzeit von Räumen mit Hinweis auf andere akustische Parameter* Internet (Nov. 2002): www.virtuallyaudio.ch/publikationen/iso_3382/iso_3382.htm
- [5] MICHAEL DICKREITER 1997 *Handbuch der Tonstudioteknik* Band 1, K.G. Saur Verlag KG, München
- [6] SVEN MÜLLER 2001 *Raumaufnahmen mit dem Yamaha SREV 1* Production Partner, Heft 11/2001 S. 106ff., MM-Musik-Media-Verlag, Köln
- [7] *Samplitude Help Index* Internet (Dez. 2002): www.samplitude.com/de/downloads.htm
- [8] SVEN MÜLLER, PAULO MASSARANI 2001 *Transfer-function measurements with sweeps* AES-Journal, Vol 49, Nr. 6, Aachen
- [9] KURT EGGENSCHWILER, KURT HEUTSCHI *Raumakustische Planungs- und Messverfahren* EMPA Dübendorf, Internet (Dez. 2002): www.arch.ethz.ch/eggenschwiler/sia1.pdf
- [10] ANDRZEJ GOLAS, LOPACZ HENRYK U.A. 1993 *Convolution of simulated room impulse response* AES Convention preprint Nr. 3287, Wien
- [11] WAVES *TrueVerb* Manual

7.2 CD-INHALT

Alle Hörbeispiele beruhen auf der Datei „Stimme-Orig“, welche auf Track 1 zu hören ist.

<i>Track-Nr.</i>	<i>1. Beispiel</i>	<i>2. Beispiel</i>
1	Stimme-Original	-
2	LS-Mic Zimmer	Zimmer Faltung
3	LS-Mic Flur TürG K6	Flur Faltung RIA A
4	LS-Mic Flur TürG K6	Flur Faltung RIA B
5	LS-Mic Flur TürO K6	Flur Faltung RIA C
6	LS-Mic Flur TürG ECM-DS70P	Flur Faltung RIA D
7	TrueVerb	Faltung RIA-TrueVerb
8	Delta-Datei aus Vergleich 5.3	
9	Dirac-Impuls (weiss)	
10	Impuls (braun)	
11	RIA-Zimmer	
12	RIA A (Flur)	
13	RIA B (Flur)	
14	RIA C (Flur)	
15	RIA D (Flur)	
16	RIA-TrueVerb	

Bezeichnungen:

LS-Mic	Originaldatei, nachdem sie durch die gesamte Messstrecke geschickt wurde
Faltung	Originaldatei, die mit entsprechender RIA gefaltet wurde
TürO/G	Raum mit offenen/geschlossenen Türen
K6	mit SENNHEISER K6 aufgenommen
ECM-DS70P	mit SONY ECM-DS70P aufgenommen
RIA A-D	Faltung mit entsprechender RIA aus 4.3 (s. <i>Tab. 4-1</i>)
TrueVerb	Originaldatei, durch TrueVerb verhallt

Umfang der Arbeit: **11'215** Wörter