

Masterarbeit

zur Erlangung des akademischen Grades Master of Science

Perzeptive Bewertung der nichtlinearen Verzerrungen von Lautsprechern

Erstgutachter:

Prof. Dr. Stefan Weinzierl (TU Berlin)

Zweitgutachter:

Prof. Dr. Ing. habil. Wolfgang Klippel (Klippel GmbH)

vorgelegt an der Technischen Universität Berlin Fachgebiet Audiokommunikation

| von: Ben | ijamir | n Henz | ze |
|------------|--------|--------|------|
| | | | |
| Abgabetern | nin: 1 | 8.07.2 | 2017 |

DANKSAGUNG

An dieser Stelle geht mein Dank an alle, die an der Realisierung und dem Gelingen dieser Arbeit beteiligt waren. Dieser gilt den Mitarbeitern der Klippel GmbH; insbesondere möchte ich mich bei Uta und Wolfgang Klippel für das entgegengebrachte Vertrauen, sowie Marian Liebig, Robert Schulze und Robert Krampe für die vielen angeregten Diskussionen bedanken.

Weiterhin möchte ich mich herzlich bei meiner Familie für die Unterstützung während der zahlreichen Reisen nach Dresden bedanken.

Berlin, im Juli 2017

EIDESSTATTLICHE VERSICHERUNG

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und eigenhändig sowie ohne unerlaubte fremde Hilfe und ausschließlich unter Verwendung der aufgeführten Quellen und Hilfsmittel angefertigt habe.

Berlin, den 18. Juli 2017

.....

Benjamin Henze

KURZFASSUNG

Elektrodynamische Lautsprecher werden häufig nahe ihrer Leistungsgrenze betrieben. Für solche Lastfälle enthält das akustische Ausgangssignal einen signifikanten Anteil nichtlinearer Verzerrungen, die vom Wirkprinzip und der Bauform eines Treibers abhängen.

Diese Arbeit untersucht ein perzeptives Evaluationsmodul hinsichtlich der Eignung zur Bewertung nichtlinearer Verzerrungen. Dafür wurde die Software aktualisiert und Fehler in der Implementierung wurden behoben. Anschließend wurden die Modellberechnungen für bestimmte Testsignale den Ergebnissen empirischer Erhebungen gegenübergestellt und die Abbildung psychoakustischer Zusammenhänge wurde verifiziert.

Weiterhin wurde für das Klippel R&D-System eine Automationsroutine entwickelt, um für eine Auswahl an Treibern und Musiktiteln kritische Kombinationen zu identifizieren, bei denen nichtlineare Verzerrungen hörbar sind. Es wurden bestehende Methoden zur Simulation, Auralisation und perzeptiven Evaluation der Lautsprecherwiedergabe miteinander verknüpft. Durch die Verwendung umfangreicher Datensätze aus Lautsprechermessungen konnten realitätsnahe Ergebnisse erzielt werden.

Abschließend wurde ein Hörversuch zur Bestimmung der Detektionsschwelle nichtlinearer Verzerrungen erweitert. Das Ziel der Erweiterung ist den Einfluss nichtlinearer Verzerrungen auf die Klangqualität der Lautsprecherwiedergabe abschätzen zu können. Es wurde ein Konzept erarbeitet, das auf Basis der Detektionsschwelle mit wenigen Messpunkten die Änderung der Klangqualität über der Pegeländerung der nichtlinearen Verzerrungen erfassen kann. Zuletzt wurde das Testverfahren einem ersten Funktionstest unterzogen.

ABSTRACT

Electrodynamic speakers are often operated near their power limit. Under these conditions the radiated acoustic output contains a considerable amount of nonlinear distortions. The character of these distortions is related to the working principle and the design of the transducer.

The current thesis examines the ability of a perceptual evaluation module to assess nonlinear loudspeaker distortions. Therefor the software was updated and bugs were fixed. The model output variables were checked against the empirical results of psychoacoustic research. Additionally the reproduction of potential psychoacoustic relationships related to certain nonlinear distortions was reviewed.

Further on a concept to identify combinations of stimuli and transducers that produce an audible amount of nonlinear distortions is developed and implemented within the Klippel R&D-System. The method used state of the art technologies to simulate, auralize and evaluate the acoustic outputs of speakers. By using measured data of actual speakers realistic results were obtained.

Lastly the recent listening test that determines the audibility threshold for nonlinear distortions was extended. It is the goal of the extension to gather data to predict the impact of certain nonlinear distortions on the sound quality of transducers. Therefor a concept to capture the change in sound quality that is caused by a change in level of certain nonlinear distortions was developed. Finally a first usability test of the new approach was successfully conducted.

INHALT

| Ku | KURZFASSUNGI | | |
|-----|--------------|---|-----|
| AB | BILDUNG | SVERZEICHNIS | IV |
| TA | BELLENV | /ERZEICHNIS | VI |
| AB | KÜRZUN | GSVERZEICHNIS | /II |
| SYI | MBOLE | V | III |
| 1 | Einleitu | ng | 1 |
| 1.1 | Motiva | ation | 1 |
| 1.2 | Aufga | be und Zielsetzung | 2 |
| 2 | Theoret | ische Grundlagen | 3 |
| 2.1 | Ausge | wählte Eigenschaften der Hörwahrnehmung | 3 |
| | 2.1.1 | Hörschwelle | 3 |
| | 2.1.2 | Maskierung | 4 |
| | 2.1.3 | Frequenzgruppen & Erregung | 4 |
| 2.2 | Wichti | ige Psychoakustische Größen | 5 |
| | 2.2.1 | Lautstärkepegel und Lautheit | 5 |
| | 2.2.2 | Schärfe | 6 |
| | 2.2.3 | Rauigkeit | 7 |
| 2.3 | Der El | ektrodynamische Lautsprecher | 9 |
| | 2.3.1 | Wirkprinzip | 9 |
| | 2.3.2 | Reguläre Nichtlinearitäten | 11 |
| | 2.3.3 | Thermische Effekte | 14 |
| | 2.3.4 | Irreguläre Nichtlinearitäten | 15 |
| | 2.3.5 | Psychoakustische Relevanz | 15 |
| 3 | Stand d | er Technik | 17 |
| 3.1 | Model | lierung und Messung des Übertragungsverhaltens | 17 |
| | 3.1.1 | Adaptives Fitting im Zeitbereich | 17 |
| | 3.1.2 | Offline-Identifikation zur Separierung von Verzerrungen | 18 |
| | 3.1.3 | Online-Identifikation zur Separierung nichtlinearer Verzerrungskomponenten. | 19 |
| 3.2 | Hörba | rmachung | 19 |
| | 3.2.1 | Auralisation durch Simulation der Zustandsvariablen | 20 |
| | 3.2.2 | Differenzauralisation | 21 |
| 3.3 | Klippe | el R&D System | 21 |
| | 3.3.1 | Linear Parameter Measurement (LPM) | 22 |
| | 3.3.2 | Large Signal Identification (LSI) | 22 |
| | 3.3.3 | Simulation (SIM2) | 22 |
| | 3.3.4 | Simulation / Auralization (SIM-AUR) | 23 |
| | 3.3.5 | Difference Auralization (DIF-AUR) | 23 |
| 3.4 | Perzep | otive Evaluationsmethoden | 24 |
| | 3.4.1 | Perceptual Evaluation of Audio Quality (PEAQ) | 24 |
| | 3.4.2 | Perceptual Evaluation of Quality (PEQ) | 24 |
| 3.5 | Online | ehörversuch | 33 |
| 4 | Optimie | erung & Evaluation des Perzeptiven Evaluationsmoduls | 34 |
| 4.1 | Update | e auf MAT2 | 34 |

| | 4.1.1 | Anpassung der Struktur | .34 |
|-------------|----------|--|-----------|
| | 4.1.2 | Implementierung von Zeit- und Frequenzcursor | .35 |
| | 4.1.3 | Optimierung der Migrationsfähigkeit | .36 |
| 4.2 | Anpas | sungen des Filterbankmodells | .36 |
| | 4.2.1 | Skalierung der Eingangssignale | .36 |
| | 4.2.2 | Hochpassfilter | .37 |
| 4.3 | Rauig | keitsmodell | .38 |
| | 4.3.1 | Änderungen in der Implementierung | .38 |
| | 4.3.2 | Variation des Modulationsindexes bei fester Trägerfrequenz | . 39 |
| | 4.3.3 | Variation der Modulationsfrequenz bei festen Trägerfrequenzen und konstant | tem |
| | | Modulationsindex | .40 |
| 4.4 | Berecl | nnung der Schärfe | .42 |
| 4.5 | Distor | tion To Mask Ratio | .45 |
| 4.6 | Discol | oration, Treble Stressing, General Bass Enhancement | .47 |
| 4.7 | Anwel | ick | .47 |
| - .0 | Tuson | | . – / |
| 5 | Konzep | t zur Suche kritischer Stimuli | .51 |
| 5.1 5.2 | A blau | 0 f und Funktionsblöcke | .51 |
| 5.2 | Anwei | ndungsbeisniel | .52 56 |
| 5.5 | 5.3.1 | Auswahl der Stimuli | .56 |
| | 5.3.2 | Auswahl der Treiber | .56 |
| | 5.3.3 | Benötigte Rechenzeit | .64 |
| | 5.3.4 | Auswertung der Daten | .65 |
| 5.4 | Ausbli | ick | .66 |
| 6 | Francito | rung das Hörvarsuchs | 68 |
| 61 | Forsch | nungsinteresse | .00 68 |
| 0.1 | 6.1.1 | Konzept des Begriffs Klangqualität | .68 |
| | 6.1.2 | Reduktion multivariater Abhängigkeiten | .70 |
| | 6.1.3 | Forschungshypothesen | .72 |
| 6.2 | Anfor | derungen an das Testverfahren | .73 |
| 6.3 | Testve | erfahren zur Bewertung der Klangqualität | .74 |
| | 6.3.1 | Beschreibung | .74 |
| | 6.3.2 | Anschluss an den bestehenden Hörversuch | .75 |
| | 6.3.3 | Robustheit | .76 |
| | 6.3.4 | Umsetzung mit Webinterface | .76 |
| 6.4 | Anwei | ndungsbeispiel | .82 |
| | 6.4.1 | Verwendete Signale | . 82 |
| | 6.4.2 | Urheberrechtliche Bedenken | . 82 |
| | 6.4.3 | Auszug aus den Ergebnissen | .83 |
| | 6.4.4 | Auswertung des Testlaufs | .85 |
| 6.5 | Ausbli | ick | . 85 |
| 7 | Zusamr | nenfassung | .87 |
| LI | TERATUR | VERZEICHNIS | .89 |
| AN | HANG 1. | | .92 |
| Dic | GITALER | ANHANG | .95 |
| ` | | | - |

ABBILDUNGSVERZEICHNIS

| Abbildung 1: Perzentile der Hörschwellenverteilung nach ISO 389-7 (Maschke & Jakob, 2010, S. 601). |
|---|
| Abbildung 2: Schnittbild eines Konuslautsprechers. (Möser, 2012, S. 374) |
| Abbildung 3: elektromechanisches Ersatzschaltbild für elektrodynamische Lautsprecher (Klippel GmbH, 2013a, S. 4) |
| Abbildung 4: Ersatzschaltbild für Großsignalbereich (Klippel GmbH, 2015a, S.2)11 |
| Abbildung 5: Steifigkeit von Spinne, Sicke und Aufhängung (Klippel, 2006, S. 7)12 |
| Abbildung 6: Schnitt durch Spule und Magnetkörper (vgl. Klippel GmbH, 2014a, S. 4) 13 |
| Abbildung 7: Modellierung des elektrodynamischen Lautsprechers (vgl. Klippel GmbH, 2014b, S. 2) |
| Abbildung 8: Fitting des Modells, vereinfacht (vgl. Klippel, 2016, S. 360)18 |
| Abbildung 9: Prinzip der Offline-Identifikation, vereinfacht (vgl. Klippel, 2016, S. 360)18 |
| Abbildung 10: Prinzip der Online-Identifikation (vgl. Klippel, 2016, S. 361) |
| Abbildung 11: Prinzip der Auralisation durch Simulation (vgl. Klippel, 2001, S. 2)20 |
| Abbildung 12: Prinzip der Differenzauralisation (vgl. Klippel GmbH, 2014b, S.7)21 |
| Abbildung 13: simple Datenbankstruktur21 |
| Abbildung 14: Struktur des PEQ-Moduls |
| |
| Abbildung 15: Verknüpfung zwischen Eingangsblock, Audioexport und Hörversuch |
| Abbildung 15: Verknüpfung zwischen Eingangsblock, Audioexport und Hörversuch |
| Abbildung 15: Verknüpfung zwischen Eingangsblock, Audioexport und Hörversuch |
| Abbildung 15: Verknüpfung zwischen Eingangsblock, Audioexport und Hörversuch |
| Abbildung 15: Verknüpfung zwischen Eingangsblock, Audioexport und Hörversuch |
| Abbildung 15: Verknüpfung zwischen Eingangsblock, Audioexport und Hörversuch26Abbildung 16: Überblick FFT-Modell27Abbildung 17: Überblick Filterbankmodell (vgl. Straube, 2013, S. 25)29Abbildung 18: Überblick Rauigkeitsmodell, vereinfacht (vgl. Kunz, 2014 S. 28)30Abbildung 19: Vergleich der Eingabemasken. links: MAT1, rechts: MAT2.35Abbildung 20: Vergleich der Hochpassfilter38 |
| Abbildung 15: Verknüpfung zwischen Eingangsblock, Audioexport und Hörversuch26Abbildung 16: Überblick FFT-Modell27Abbildung 17: Überblick Filterbankmodell (vgl. Straube, 2013, S. 25)29Abbildung 18: Überblick Rauigkeitsmodell, vereinfacht (vgl. Kunz, 2014 S. 28)30Abbildung 19: Vergleich der Eingabemasken. links: MAT1, rechts: MAT2.35Abbildung 20: Vergleich der Hochpassfilter38Abbildung 21: Rauigkeit bei variabler Modulationstiefe.40 |
| Abbildung 15: Verknüpfung zwischen Eingangsblock, Audioexport und Hörversuch |
| Abbildung 15: Verknüpfung zwischen Eingangsblock, Audioexport und Hörversuch26Abbildung 16: Überblick FFT-Modell27Abbildung 17: Überblick Filterbankmodell (vgl. Straube, 2013, S. 25)29Abbildung 18: Überblick Rauigkeitsmodell, vereinfacht (vgl. Kunz, 2014 S. 28)30Abbildung 19: Vergleich der Eingabemasken. links: MAT1, rechts: MAT2.35Abbildung 20: Vergleich der Hochpassfilter38Abbildung 21: Rauigkeit bei variabler Modulationstiefe.40Abbildung 22: Rauigkeit bei variabler Modulationsfrequenz für mehrere Trägerfrequenzen. 41Abbildung 23: Schärfe. Vergleich zwischen Modell und Daten nach Fastl & Zwicker.43 |
| Abbildung 15: Verknüpfung zwischen Eingangsblock, Audioexport und Hörversuch26Abbildung 16: Überblick FFT-Modell27Abbildung 17: Überblick Filterbankmodell (vgl. Straube, 2013, S. 25)29Abbildung 18: Überblick Rauigkeitsmodell, vereinfacht (vgl. Kunz, 2014 S. 28)30Abbildung 19: Vergleich der Eingabemasken. links: MAT1, rechts: MAT2.35Abbildung 20: Vergleich der Hochpassfilter.38Abbildung 21: Rauigkeit bei variabler Modulationstiefe.40Abbildung 22: Rauigkeit bei variabler Modulationsfrequenz für mehrere Trägerfrequenzen. 41Abbildung 23: Schärfe. Vergleich zwischen Modell und Daten nach Fastl & Zwicker.43Abbildung 24: Vergleich der Amplitudenspektren zweier Rauschsignale.44 |
| Abbildung 15: Verknüpfung zwischen Eingangsblock, Audioexport und Hörversuch26Abbildung 16: Überblick FFT-Modell27Abbildung 17: Überblick Filterbankmodell (vgl. Straube, 2013, S. 25)29Abbildung 18: Überblick Rauigkeitsmodell, vereinfacht (vgl. Kunz, 2014 S. 28)30Abbildung 19: Vergleich der Eingabemasken. links: MAT1, rechts: MAT2.35Abbildung 20: Vergleich der Hochpassfilter38Abbildung 21: Rauigkeit bei variabler Modulationstiefe.40Abbildung 22: Rauigkeit bei variabler Modulationsfrequenz für mehrere Trägerfrequenzen. 41Abbildung 23: Schärfe. Vergleich zwischen Modell und Daten nach Fastl & Zwicker.43Abbildung 24: Vergleich der Amplitudenspektren zweier Rauschsignale.44Abbildung 25: Masking Pattern und Distortion Pattern. DMR = -3,8 dB.46 |
| Abbildung 15: Verknüpfung zwischen Eingangsblock, Audioexport und Hörversuch26Abbildung 16: Überblick FFT-Modell27Abbildung 17: Überblick Filterbankmodell (vgl. Straube, 2013, S. 25)29Abbildung 18: Überblick Rauigkeitsmodell, vereinfacht (vgl. Kunz, 2014 S. 28)30Abbildung 19: Vergleich der Eingabemasken. links: MAT1, rechts: MAT2.35Abbildung 20: Vergleich der Hochpassfilter38Abbildung 21: Rauigkeit bei variabler Modulationstiefe.40Abbildung 22: Rauigkeit bei variabler Modulationsfrequenz für mehrere Trägerfrequenzen. 41Abbildung 23: Schärfe. Vergleich zwischen Modell und Daten nach Fastl & Zwicker.43Abbildung 25: Masking Pattern und Distortion Pattern. $DMR = -3,8$ dB.46Abbildung 26: Masking Pattern und Distortion Pattern. $DMR = -11,9$ dB.46 |
| Abbildung 15: Verknüpfung zwischen Eingangsblock, Audioexport und Hörversuch26Abbildung 16: Überblick FFT-Modell27Abbildung 17: Überblick Filterbankmodell (vgl. Straube, 2013, S. 25)29Abbildung 18: Überblick Rauigkeitsmodell, vereinfacht (vgl. Kunz, 2014 S. 28)30Abbildung 19: Vergleich der Eingabemasken. links: MAT1, rechts: MAT2.35Abbildung 20: Vergleich der Hochpassfilter38Abbildung 21: Rauigkeit bei variabler Modulationstiefe.40Abbildung 22: Rauigkeit bei variabler Modulationsfrequenz für mehrere Trägerfrequenzen. 41Abbildung 23: Schärfe. Vergleich zwischen Modell und Daten nach Fastl & Zwicker.43Abbildung 24: Vergleich der Amplitudenspektren zweier Rauschsignale.44Abbildung 25: Masking Pattern und Distortion Pattern. $DMR = -3,8$ dB.46Abbildung 27: Versuchsaufbau für irreguläre Nichtlinearitäten.48 |

| Abhildung 20: Funktionaller Ablauf der Automationgrouting | 53 |
|---|------------------|
| | |
| Abbildung 30: Front- und Ruckansicht, Speaker#001 | |
| Abbildung 31: Amplitudenfrequenzgang, Speaker#001 | 57 |
| Abbildung 32: <i>Bl</i> (<i>x</i>), Speaker#001. Vermessen mit LSI | |
| Abbildung 33: Kumulierte Modulationsverzerrung, Speaker#001 | |
| Abbildung 34: Front- und Rückansicht, Speaker#002 | 59 |
| Abbildung 35: Amplitudenfrequenzgang, Speaker#002. | 59 |
| Abbildung 36: $L_e(x)$, Speaker#002. Vermessen mit LSI | 60 |
| Abbildung 37: Kumulierte Modulationsverzerrungen, Speaker#002 | 60 |
| Abbildung 38: $K_{ms}(x)$, Speaker#002. Vermessen mit LSI | 61 |
| Abbildung 39: Relative harmonische Verzerrungen, Speaker#002 | 61 |
| Abbildung 40: <i>Bl</i> (<i>x</i>), Microspeaker. Vermessen mit LSI | 62 |
| Abbildung 41: $K_{ms}(x)$, Microspeaker. Vermessen mit LSI | 62 |
| Abbildung 42: Amplitudenfrequenzgang, Microspeaker. | 63 |
| Abbildung 43: Relative harmonische Verzerrungen, Microspeaker. | 63 |
| Abbildung 44: Kumulierte Modulationsverzerrungen, Microspeaker | 64 |
| Abbildung 45: Klangqualität & Hörbarkeit über Verzerrungsintensität (angelehnt a 2016, S. 503). | n: Klippel 69 |
| Abbildung 46: Beispiel zur Zielperformance. | 70 |
| Abbildung 47: Beispielgrafik zu Gradienten ΔSQ . | 71 |
| Abbildung 48: Verschiedene Gradienten ΔSQ | 72 |
| Abbildung 49: Beispiel zur unterschiedlichen Skalennutzung | 76 |
| Abbildung 50: Schema der Dateistruktur. | 77 |
| Abbildung 51: Benutzeroberfläche auf Smartphones | 78 |
| Abbildung 52: Verlaufsanzeige der Detektionsschwellenbestimmung | 79 |
| Abbildung 53: Histogramm der Hörschwellenverteilung | 79 |
| Abbildung 54: Darstellung der Bewertung der Klangqualität | |
| Abbildung 55: Histogramm der Bewertung der Klangqualität | |
| | |

TABELLENVERZEICHNIS

| Tabelle 1: Kalibrierfaktoren des Rauigkeitsmodells bei variabler Abtastrate | 39 |
|---|----|
| Tabelle 2: Berechnete Schärfe S. | 45 |
| Tabelle 3: MOVs und nichtlineare Verzerrungen. | 48 |
| Tabelle 4: Durchschnittliche Rechenzeit als Funktion der Signallänge | 64 |
| Tabelle 5: Ergebnisse Speaker#001. Aktiv: $Bl(x)$, $K_{ms}(x)$, $L_e(x)$ | 65 |
| Tabelle 6: Ergebnisse Speaker#002. Aktiv: $Bl(x)$, $K_{ms}(x)$, $L_e(x)$ | 65 |
| Tabelle 7: Ergebnisse Microspeaker. Aktiv: $Bl(x)$, $K_{ms}(x)$, $L_e(x)$ | 66 |
| Tabelle 8: Zuordnung von Skalenstufen und Qualitätsbeschreibungen. | 74 |
| Tabelle 9: Übersicht erfasster Daten | 81 |
| Tabelle 10: Häufigkeiten verwendeter Abhörgeräte | 84 |
| Tabelle 11: Häufigkeiten durchgeführter Test, sortiert nach Treibern | 84 |
| Tabelle 12: Häufigkeiten durchgeführter Tests, DuT Microspeaker | 84 |

ABKÜRZUNGSVERZEICHNIS

| Abkürzung | Beschreibung | | |
|-----------|--|--|--|
| BackEMF | Elektromotorische Gegenkraft. | | |
| CSV | Comma Separated Values. Dateiformat. | | |
| DIF-AUR | Difference Auralization. Name eines Klippel Moduls. | | |
| DuT | Device under Test. In dieser Arbeit ist damit stets ein zu prüfender Lautsprecher gemeint. | | |
| FFT | Fast Fourier Transformation. | | |
| JND | Just Noticeable Difference. Gerade wahrnehmbarer Unterschied für eine Reizgröße. | | |
| LPM | Linear Parameter Measurement. Name eines Klippel Moduls. | | |
| LSI | Large Signal Identification. Name eines Klippel Moduls. | | |
| MOV | Model Output Variable. Bezeichnet eine durch ein Modell berechnete Ausgangsgröße. | | |
| РСМ | Pulse Code Modulation. Diskrete Abtastung kontinuierlicher Signale. | | |
| PEAQ | <i>Perceptual Evaluation of Audio Quality</i> ist ein Vorschlag der ITU für einen gehörgerechten Bewertungsalgorithmus der Klangqualität von digitalen Audiocodierverfahren. | | |
| PEQ | <i>Perceptual Evaluation of Quality</i> ist ein Modul der Software dB-Lab und derzeit im Entwicklungsstatus. | | |
| SIM-AUR | Simulation / Auralization. Name eines Klippel Moduls. | | |
| SIM2 | Simulation Version 2. Name eines Klippel Moduls. | | |
| SQL | Structured Query Language. Datenbankformat. | | |
| UEN | Uniform Exciting Noise. Spezielles Rauschsignal. | | |

| Symbol | Einheit | Beschreibung |
|--|-----------|---|
| | Bark | Einheit der Frequenzgruppenbreite |
| | sone/Bark | Lautheit je Frequenzgruppe |
| Δf | Hz | Kritische Bandbreite einer Frequenzgruppe |
| $\Delta SQ_{Pi,AT}, \ \Delta SQ_{Pi,DuT}$ | | Gradienten der Änderung der Klangqualität |
| AT | dB | Detektionsschwelle relativ zu S_{dis} |
| В | Т | Magnetische Flussdichte |
| $b_{\scriptscriptstyle (n,i)}$, $b*_{\scriptscriptstyle (n,i)}$ | | Blöcke im Rauigkeitsmodell, Blocknummer <i>n</i> , Filterbankkanal <i>i</i> , Zeitbereichsdarstellung |
| $B_{(n,i)}, B_{(n,i)}^{*}$ | | Blöcke im Rauigkeitsmodell, Blocknummer <i>n</i> , Filterbankkanal <i>i</i> , Spektraldarstellung |
| BE | % | Bass Enhancement |
| Bl bzw. $Bl(x)$ | N/A | Kraftfaktor |
| cal | | Kalibrierfaktor des Rauigkeitsmodells |
| $C_{ms} bzw.$ $C_{ms}(x)$ | mm/N | Mechanische Nachgiebigkeit, reziproker Wert der Steifigkeit |
| $DC_{(n,i)}$ | | Gleichanteil des Blocks n, Filterbankkanal i |
| $d_{\it dis}$, $y_{\it diff}$ | | Differenzsignal im Zeitbereich |
| DMR, DMR _{tot} | dB | Distortion To Mask Ratio, Total Distortion To Mask Ratio |
| DMR_{max} | dB | Maximum der DMR |
| DP | | Distortion Pattern. Repräsentation des Differenzsignals. |
| e(t) | | Fehlersignal |
| F | Ν | Kraft |
| f_c | Hz | Trägerfrequenz |
| F_{coil} | Ν | Auslenkende Kraft, wirkt auf Schwingspule |
| f_{g} | Hz | Grenzfrequenz eines Filters |
| f_{gw}, f_{go} | Hz | Untere und obere Grenzfrequenzen von Rauschsignalen |
| F_{K} | Ν | Rückstellkraft |
| f_m | Hz | Modulationsfrequenz (Rauigkeit) oder Mittenfrequenz (Schärfe) |
| F_{M} | Ν | Trägheitskraft |
| $F_m(x, i, i_2)$ | Ν | Reluktanzkraft |

| Symbol | Einheit | Beschreibung |
|--|------------------------------------|---|
| F_R | Ν | Kraftwirkung des mechanischen Widerstands |
| f_s | Hz | Resonanzfrequenz oder Abtastfrequenz |
| $f_{s,red\#1}, f_{s,red\#2}$ | Hz | Reduzierte Abtastfrequenzen |
| g(z) | | Wichtungsfunktion im Schärfemodell |
| g(zi) | | Wichtungsfunktion im Rauigkeitsmodell |
| $G_{aur}, G_{aur,init}$ | dB | Auralization Gain. Verstärkungsmaß des SIM-AUR-Moduls |
| $G_{\scriptscriptstyle E}$ | dB | Export Gain, Skalierer des PEQ-Moduls |
| G_{L} | dB | Equalization Gain, Skalierer des PEQ-Moduls |
| G_P | dB | Playback Gain, Skalierer des PEQ-Moduls |
| G_R | dB | Reference Gain, Skalierer des DIF-AUR- und PEQ-Moduls |
| Н | $\mathbf{A} \cdot \mathbf{m}^{-1}$ | Feldstärke |
| $H_{(i)}$ | | Wichtungsfunktion im Rauigkeitsmodell |
| H _{HP,PEAQ} , H _{HP,PEQ} | | Übertragungsfunktionen der Hochpassfilter |
| h_{Spalt} | mm | Höhe des Luftspalts |
| $h_{\scriptscriptstyle Spule}$ | mm | Höhe der Schwingspule |
| <i>i</i> bzw. <i>i</i> (<i>t</i>) | А | Strom |
| $k_{\scriptscriptstyle(n,i)}$ | | Korrelationskoeffizienten im Rauigkeitsmodell |
| K_{ms} bzw. $K_{ms}(x)$ | N/mm | Steifigkeit |
| l | m | Länge des Leiters im Magnetfeld |
| L_2 | mH | Parasitäre Induktivität |
| L_e bzw. $L_e(x,i)$ | mH | Induktivität der Schwingspule |
| L_N | phon | Lautstärkepegel |
| L_p | dB SPL | Schalldruckpegel |
| $L_{p,cal}$ | dB SPL | Schalldruckpegel des Kalibriersignals |
| L _{RMS} | dB | RMS-Pegel |
| $L_{y,RMS}$ | dB SPL | RMS-Schalldruckpegel eines Schalldrucksignals y |
| <i>m</i> , <i>m</i> _(<i>n</i>,<i>i</i>) | | Modulationsindex, <i>m</i> je Block <i>n</i> , Filterbankkanal <i>i</i> |
| M, MP | | Maskierungsschwelle |
| min_x, max_x | | Grenzen der Aussteuerung (Automationsroutine) |

| Symbol | Einheit | Beschreibung |
|---|---------|--|
| M_{ms} | g | Mechanische Masse |
| Ν | sone | Lautheit |
| Ν | | Anzahl der Samples <i>n</i> oder Blöcke <i>b</i> |
| n_{fft} | Samples | FFT-Länge |
| Nl_j | | Nichtlinearität j |
| p(t) | Pa | Schalldrucksignal im Zeitbereich |
| $p_{aur}, p_{aur}(t)$ | Pa | Schalldrucksignal im Zeitbereich (Auralisation) |
| $p_{Bl}(t), p_{K_{ms}}(t),$ $p_{L_e}(t)$ | Ра | Schalldrucksignale im Zeitbereich: $Bl(x)$ -, $K_{ms}(x)$ - und $L_e(x)$ -Verzerrungen |
| p_{cal} | Pa | Schalldrucksignal (Kalibriersignal) |
| $p_{def}(t)$ | Pa | Schalldrucksignal im Zeitbereich (irreguläre Defekte) |
| $p_{\it diff}$ | Pa | Schalldrucksignal (Differenzsignal) |
| $p_{dis}(t)$ | Pa | Schalldrucksignal im Zeitbereich (alle Verzerrungen) |
| P_i | | Proband <i>i</i> |
| $p_{lin.dis}(t),$ | Pa | Schalldrucksignal im Zeitbereich (lineare Verzerrungen) |
| $p_{nlin.dis}(t)$ | Pa | Schalldrucksignal im Zeitbereich (nichtlineare Verzerrungen) |
| $p_{noise}(t)$ | Pa | Schalldrucksignal im Zeitbereich (akustische Störquellen) |
| p_{R}, p_{T} | Pa | Schalldrucksignale (Referenz- und Testsignal) |
| Q_{es} | | Elektrischer Gütefaktor |
| Q_{ms} | | Mechanischer Gütefaktor |
| Q_{ts} | | Gesamtgütefaktor |
| R , $R_{(n)}$ | asper | Rauigkeit nach Fastl & Zwicker, je Block n |
| $r_{(n,i)}$ | | Teilrauigkeit je Block n, Filterbankkanal i |
| R_2 | Ω | Elektrischer Widerstand durch Wirbelstromverluste |
| R_e | Ω | Gleichstromwiderstand der Schwingspule |
| R_{ms} | kg/s | Mechanischer Widerstand |
| S | acum | Absolute Schärfe nach Fastl & Zwicker |
| S_{Bb} , $S_{K_{ms}}$, S_{L_e} | | Skalierer: $Bl(x)$ -, $K_{ms}(x)$ - und $L_e(x)$ -Verzerrungen |
| S_{dis} | dB | Skalierer aller nichtlinearen Verzerrungen bei der Auralisation |
| Sig | | Stimulus |
| S_{lin} | | Skalierer, lineare Verzerrungen |

| Symbol | Einheit | Beschreibung |
|---|------------------|---|
| SQ | | Klangqualität |
| t | S | Zeit |
| T_a | Κ | Umgebungstemperatur |
| TS | % | Treble Stressing |
| t_{sig} | S | Länge eines Stimulus |
| $T_{_{V,}} \Delta T_{_{V,krit}}$ | К | Temperatur der Schwingspule bzw. Änderung der Temperatur der Schwingspule |
| $U, U(t), U_{init}$ | V | Spannung |
| V | % | Discoloration |
| V | m/s | Schnelle |
| x bzw. $x(t)$, x_{max} | mm | Auslenkung, Spitzenauslenkung |
| x_{PSE} | mm | Gültigkeitsbereich der Messdaten (LSI) |
| X_R, X_T | | Referenz- und Testsignal im Zeitbereich |
| Y _A , Y _{aur} | | Auralisiertes Signal im Zeitbereich |
| ${\cal Y}_{cal}$ | | Kalibriersignal im Zeitbreich |
| <i>Y</i> _{<i>R</i>} , <i>Y</i> _{<i>T</i>} | | Zeitsynchrone Versionen von x_R , und x_T |
| Y_{R} , Y_{T} , Y_{diff} | | Zeitsynchrone Spektren von Referenz-, Test- und Differenzsignal |
| Y _{R,Env,i} , Y _{T,Env,i} | | Hüllkurven der Filterbankkanäle |
| $Y_{R,W}$, $Y_{T,W}$ | | Gewichtete Version von Y_R , und Y_T |
| $Z_L(f)$ | Ω | Impedanz der Schwingspule |
| δ | W/(m·K) | Thermische Leitfähigkeit |
| μ | $H \cdot m^{-1}$ | Magnetische Permeabilität |
| τ | | Zeitversatz zwischen Referenz- und Testsignal |

1 Einleitung

1.1 Motivation

Audiosysteme gewinnen im Zuge der fortschreitenden Digitalisierung des Lebensalltags zunehmend an Wichtigkeit und sind auch durch mediale Phänomene der jüngeren Zeit in immer mehr technischen Geräten zu finden. Gemeinsam mit den klassischen Anwendungsbereichen der heimischen Musikwiedergabe und der Beschallung von Sprachbis Konzertveranstaltungen ergibt sich ein großer Umfang verschiedener Wiedergabeanforderungen, die als Gemeinsamkeit die Rezeption durch das menschliche Hörvermögen teilen.

Audioprodukte für professionelle Anwender sollen zumeist einen eingeschränkten und finanzkräftigen Markt bedienen und werden entsprechend für lange Einsatzzeiten und reproduzierbar gute Wiedergabeergebnisse ausgelegt. Gegensätzlich verhält es sich mit den Massenmärkten der Unterhaltungselektronik: Um in diesem Bereichen kompetitive Verkaufspreise bei zu minimierenden Einsatz zu erzielen, ist es notwendig die verwendeten Komponenten häufig in Grenzbereichen zu betreiben. Im Extrembeispiel mobiler Endgeräte bedeutet das bei überschaubarer Akkuleistung und kleinstem Bauraum möglichst hohe Schallleistungen für die Audiowiedergabe bei zumindest akzeptabler Wiedergabequalität zu erzielen.

Elektrodynamische Lautsprecher sind dabei unverändert das sprichwörtliche "schwächste Glied" in der Signalkette mit vergleichbar schlechten Wirkungsgraden und gegenüber der antreibenden Leistungselektronik deutlich erhöhten Verzerrungen. Um die eingangs erwähnten ökonomischen und technischen Anforderungen in Einklang zu bringen, müssen die Lautsprecher dabei jenseits ihres linearen Arbeitsbereiches betrieben werden. Es entstehen nichtlineare Verzerrungen, z.B. durch die Auslenkung der Schwingspule außerhalb des homogenen Magnetfeldes (vgl. Dickreiter, 1997, S. 213). Dies wird in einigen Fällen, wie aktiven Kleinstlautsprechern, nur durch eine vorgeschaltete aktive z.B. bei Regelungselektronik möglich, die Einfluss auf das von der Leistungselektronik zu verstärkende elektrische und somit auf das vom Lautsprecher abzustrahlende Schalldrucksignal nimmt.

Die nichtlinearen Verzerrungen können mit technischen Maßen, z.B. dem Anteil harmonischer Verzerrungen, teilweise erfasst werden (vgl. ebd.). Jedoch besitzen diese Maße verhältnismäßig geringe Aussagekraft bzgl. der Rezeption und Bewertung der Klangqualität beim Hörer. Folglich ist es notwendig die Auswirkungen nichtlinearer Verzerrungen auf die Hörwahrnehmung einer Zielgruppe von potenziellen Hörern zu untersuchen. Eine traditionelle Möglichkeit zur empirischen Erfassung ist der traditionelle Hörversuch. Vor dem Hintergrund von monetären Interessen von Lautsprecherherstellern ist dies nicht mehr zeitgemäß, da zumeist nur geringe Stichprobengrößen erzielt werden und der zeitliche und technische Aufwand sehr hoch ist.

Um die Effektivität dieses Prozesses zu steigern wurden verschiedene Werkzeuge entwickelt: Zunächst kann der Prozess der Lautsprecherentwicklung durch ein umfassendes physikalisches Modell und numerische Simulationsmethoden vom realen zum digitalen Prototypen verlagert werden. Zudem können fortgeschrittene Methoden zur Hörbarmachung in der digitalen Domäne genutzt werden. Mittels perzeptiver Modelle können die hörbargemachten Prototypen hinsichtlich der wichtigsten psychoakustischen Größen numerisch evaluiert werden um final eine Auswahl für schnelle, computergestützte Hörversuche treffen zu können. Dabei findet auf dem Weg von der Entwicklungs- zur Marketingabteilung eines Unternehmens eine Dimensionsreduktion auf einen für alle Beteiligten verständlichen Datensatz statt.

1.2 Aufgabe und Zielsetzung

Die vorliegende Arbeit wurde in der *Klippel GmbH* am Standort Dresden erarbeitet und basiert auf Entwicklungen für die Anwendung in Messsystemen für elektrodynamische Lautsprecher.

Für ein bestehendes perzeptives Evaluationsmodul gilt es die Eignung für die psychoakustische Bewertung nichtlinearer Verzerrungen zu evaluieren und die bestehende Implementierung ggf. zu optimieren. Dabei sind insbesondere bisher nicht geprüfte Entwicklungen wie z.B. das implementierte Modell zur Erfassung der Rauigkeit zu untersuchen.

Weiterhin ist ein Konzept zu entwickeln und zu implementieren, dass die Suche nach besonders kritischen Stimuli aus einer Auswahl von verschiedenem Programmmaterialien ermöglicht. Dabei sollen bestehende Methoden der numerischen Simulation von Zustandsgrößen der Lautsprecherwiedergabe mit bestehenden Auralisationstechniken und dem perzeptiven Evaluationsmodul verbunden werden.

Final gilt es den bestehenden Online-Hörversuch zu erweitern: Es soll ein Konzept entwickelt werden, welches es ermöglicht den Einfluss nichtlinearer Verzerrungen auf die *Klangqualität* zu untersuchen. Gemeinsam mit Webentwicklern ist ein Entwurf umzusetzen.

Um diese Ziele zu erreichen wird zunächst die sequentielle Implementierung des perzeptiven Evaluationsmoduls im MAT1-Format auf das modulare MAT2-Format der Software dB-Lab der Klippel GmbH aktualisiert. Die Ergonomie der Benutzereingaben ist zu verbessern und die Migrationsfähigkeit der Modelldaten zwischen verschiedenen Nutzern und Maschinen herzustellen. Anschließend werden die Ausgaben des perzeptiven Evaluationsmoduls mit zu erzeugenden Testsignalen geprüft und ggf. Optimierungen in der Implementierung vorgenommen.

Mit dem Automation Interface der Software dB-Lab wird eine Routine implementiert, die es ermöglicht die Zustandsgrößen der Lautsprecherwiedergabe für digital vorliegendes Programmaterial numerisch zu simulieren und zu evaluieren. Als Basis dienen gemessene oder synthetische Daten von elektrodynamischen Lautsprechern, die auch für die Simulation und Hörbarmachung genutzt werden. Die auralisierten Zeitsignale werden dann von der Routine dem perzeptiven Evaluationsmodul übergeben und ausgewertet. Die gewonnen Erkenntnisse über spezifische Teile des ausgewerteten Programmaterials werden für die Erweiterung des Onlinehörversuchs genutzt.

Der fachliche Rahmen der Arbeit wird in den folgenden Kapiteln zunächst mit einem Überblick über die wichtigsten psychoakustischen Größen und das Wirkprinzip der Lautsprecherwiedergabe eröffnet. Dabei werden die Ursachen nichtlinearer Verzerrungen beleuchtet und ihre psychoakustische Relevanz dargestellt. Anschließend wird der Stand der Technik dargelegt und die verwendeten Methoden erläutert. Es folgt die Evaluation des perzeptiven Evaluationsmoduls, die Vorstellung der Automationsroutine und die Konzeption der Erweiterung des Hörversuchs. Die Ergebnisse einzelner Abschnitte werden innerhalb dieser direkt dargestellt und diskutiert, da dies für das Verständnis nachfolgender Kapitel nötig ist. Zuletzt erfolgt eine Zusammenfassung hinsichtlich der Erreichung der definierten Ziele. Weiterhin wird ein Ausblick auf weitere Forschungsmöglichkeiten vorgenommen.

2 Theoretische Grundlagen

Die grundlegende Physiologie des menschlichen Hörvermögens ist bekannt und seine Funktion kann mit der Umsetzung eines Schallreizes in Empfindungen umrissen werden (vgl. Dickreiter, 1997, S. 107 ff.). Die Transformation von der mechanischen Ebene der Schallwelle auf die elektrische Ebene der Impulse beteiligter Nervenzellen hat dabei erheblichen Einfluss auf die Gestalt des Reizes.

Um die Anforderungen an perzeptive Modelle erfassen zu können ist Verständnis und Kenntnis ausgewählter Eigenschaften der Hörwahrnehmung und psychoakustischer Größen erforderlich.

2.1 Ausgewählte Eigenschaften der Hörwahrnehmung

2.1.1 Hörschwelle

Der Schalldruckpegel, ab dem ein Schallreiz eine Hörempfindung hervorruft, ist durch die Hörschwelle gekennzeichnet (vgl. ebd., S. 110). Diese Hörschwelle ist nicht über das gesamte hörbare Spektrum konstant, sondern verhält sich frequenzabhängig (vgl. Maschke & Jakob, 2010, S. 601). Darüber hinaus verändert sie sich durch individuelle Einflüsse wie Gehörschäden, Tagesform eines Hörers oder dessen Alter: Beispielsweise steigt mit zunehmendem Alter die Hörschwelle für hohe Frequenzen um teilweise bis zu 20 dB SPL an (vgl. Dickreiter, 1997, S. 110 & vgl. Fastl & Zwicker, 2007, S. 21).

Um die individuelle Ruhehörschwelle zu ermitteln wird ein Stimulus mit einem Sinuston mit sehr langsam ansteigender Signalfrequenz präsentiert. Der Proband kann den Pegel des Stimulus verändern und ist aufgefordert diesen um den Punkt zu variieren, an dem der Ton gerade hörbar wird. Die Pegeländerung über der Änderung der Frequenz wird aufgezeichnet und stellt die individuelle Ruhehörschwelle dar. Der Vorgang wird Békésy-Verfahren genannt. (vgl. zu diesem Absatz Fastl & Zwicker, 2007, S. 19)

2005 wurde eine gültige Normalhörschwelle in der Norm ISO 389-7standardisiert, die als Mittelwert aller individuellen Hörschwellen gilt (vgl. Maschke & Jakob, 2010, S. 601). Abbildung 1 illustriert die Normalhörschwelle als 50%-Perzentil der Hörschwellenverteilung:



Abbildung 1: Perzentile der Hörschwellenverteilung nach ISO 389-7 (Maschke & Jakob, 2010, S. 601).

2.1.2 Maskierung

Die Wahrnehmbarkeit eines Schallreizes ist nicht allein von der Hörschwelle bestimmt. Darüber hinaus kann ein Schallreiz von einem weiteren Schallreiz maskiert werden, wenn dieser *Maskierer* entweder einen entsprechenden Schalldruckpegel aufweist oder spektral nahe dem ursprünglichen Reiz ist.

Die Maskierungsschwelle ist der Schalldruckpegel, den ein Testton haben muss, um neben einem simultan angebotenen Maskierer wahrgenommen zu werden (vgl. Fastl & Zwicker, 2007, S. 61). Sie lässt sich auch als Anhebung der Ruhehörschwelle für den betreffenden Testton beschreiben (vgl. Maschke & Jakob, 2010, S. 603). Der Einfluss der Maskierung bewirkt jedoch nicht, dass nur zwischen "hörbaren" und "nicht hörbaren" Reizen unterschieden wird: Wird ein Testton simultan mit einem im Schalldruckpegel ansteigenden Maskierer angeboten, so entsteht bei größerer Pegeldifferenz der beiden Signalpegel zunächst der Eindruck, dass der Testton leiser wird, bevor er bei geringen Differenzen nicht hörbar bzw. vollständig maskiert wird (vgl. Fastl & Zwicker, 2007, S. 61). Die Maskierungsschwelle ist dabei frequenz- und pegelabhängig und wird auch als Verdeckungsschleppe bezeichnet (vgl. Maschke & Jakob, 2010, S. 602). Testtöne, deren Frequenz deutlich niedriger als die des Maskierers ist, haben durch die spektrale Verarbeitung des Gehörs eine niedrigere Maskierungsschwelle als solche, deren Frequenz höher ist.

Neben Maskierungseffekten, die auf der simultanen Darbietung zweier Stimuli beruhen, gibt es auch einen Einfluss des zeitlichen Bezugs versetzt dargebotener Stimuli. Es wird neben der erläuterten Simultanverdeckung auch in Vorverdeckung und Nachverdeckung unterschieden (vgl. ebd.). Vorverdeckung beschreibt die Maskierung, die entsteht, wenn ein Testton kurz vor dem Maskierer angeboten wird, während die Nachverdeckung den gegensätzlichen zeitlichen Bezug betrachtet. Fastl & Zwicker ordnen der Nachverdeckung einen stärkeren Einfluss zu (vgl. 2007, S. 62). Moore stellt die Abhängigkeit der Stärke der zeitlichen Verdeckungseffekte von der individuellen Hörerfahrung heraus und betont das geübte Hörer weniger Vorverdeckung anfällig für die sind (vgl. 2007. S. 465). Neben dem zeitlichen Abstand zwischen Maskierer und Testton ist auch das Verhältnis der Signalpegel zu berücksichtigen: Beispielsweise ist für die Nachverdeckung spätestens ab einem Abstand von 100 bis 200 ms die Maskierungsschwelle wieder auf dem Niveau der Ruhehörschwelle angelangt. Gleichzeitig ändert eine Erhöhung des Schallpegels des Maskierers um 10 dB SPL die Maskierungsschwelle bei konstantem zeitlichen Abstand der Stimuli i.d.R. aber nur um 3 dB (vgl. ebd.).

Dickreiter ordnet diese Phänomene der Anpassung des Gehörs an seine akustische Umgebung zu. Dieses bilde ein Bezugssystem, welches in gewissen Grenzen ähnliche qualitative Bewertungen von Schallreizen ermöglichen kann. (vgl. zu diesem Absatz 1997, S. 113).

2.1.3 Frequenzgruppen & Erregung

Die spektrale Arbeitsweise des Gehörs wurde u.a. durch die Erforschung der Maskierung verständlich. Das Kernelement ist die Basiliarmembran im Innenohr: Nachdem die Schallwelle mechanisch vom Trommelfell durch die Gehörknochen auf das ovale Fenster der Hörschnecke übertragen wurde, regt die Flüssigkeit in der Hörschnecke abhängig von der Wellenlänge des Schallreizes verschiedene Abschnitte auf der Basiliarmembran an. Die Sinneszellen auf der Basiliarmembran leiten den Reiz dann elektrisch ans Nervensystem weiter. (vgl. zu diesem Absatz ebd., S. 107)

Jeder Abschnitt der Basiliarmembran reagiert somit nur auf einen eingeschränkten Frequenzbereich: Diese Bereiche werden *Frequenzgruppen* genannt (vgl. Moore, 2007, S. 462). Verschiedene Versuche mit schmalbandigem Rauschen als Maskierer und reinen Testtönen wiesen nach, dass die Verarbeitung der Frequenzgruppen zur Hörempfindung dabei einer Filterbank aus Bandpässen mit verschiedenen Mittenfrequenzen ähnelt (vgl. ebd., S. 462 ff. & vgl. Fastl & Zwicker, 2007, S. 149 ff.).

Fastl & Zwicker spezifizierten dieses Verhalten in dem Konzept der *Critical Bands*: Für das hörbare Spektrum wird in 25 Frequenzgruppen bzw. Critical Bands unterschieden, deren Breite in der Einheit *Bark* gemessen wird und abhängig von der Mittenfrequenz einer jeden Frequenzgruppe ist (vgl. 2007, S. 158 ff.). Eine ähnliche Skala ist die *ERB*-Skala nach Moore (vgl. 2007, S. 464), die in dieser Arbeit nicht weiter betrachtet wird, da sie keine Verwendung in den folgenden Implementierungen findet.

Die Gestalt der Bandpässe wurde mit verschiedenen Methoden untersucht und kann numerisch mit einer Reihe von Filterfunktionen angenähert werden. Allen Untersuchungen gemein ist, dass die Funktion jedes Bandpassfilters eine zum Nähe zum Maskierungsverhalten der jeweiligen Frequenzgruppe hat (vgl. ebd.) und dass die Filterflanke unterhalb der Mittenfrequenz steiler abfällt als oberhalb, sobald der Schallreiz Schallpegel über 40 dB SPL erreicht (vgl. Maschke & Jakob, 2010, S. 605). Während die untere Flanke z.B. mit 27 dB je Bark fällt (vgl. Fastl & Zwicker, 2007, S. 170) bestehen verschiedene numerische Näherungen für das nichtlineare Verhalten der oberen Flanke (vgl. Maschke & Jakob, 2010, S. 606). Die Filter werden als *Auditory Filter* bezeichnet (vgl. Moore, 2007, S. 464). Für Reize, die mehrere Frequenzgruppen bzw. Auditory Filter erregen, wird aus der Ortsverteilung der Frequenzgruppenpegel ein Erregungsmuster bzw. *Excitation Pattern* gebildet (vgl. Maschke & Jakob, 2010, S. 605).

Die Kenntnis der beschriebenen Sachverhalte ist die Basis für funktionelle Modelle zur numerischen Annäherung der menschlichen Hörfunktion.

2.2 Wichtige Psychoakustische Größen

Teilnehmer von Hörversuchen sind in der Lage die durch die Schallwahrnehmung hervorgerufenen Hörempfindungen distinktiv zu unterscheiden. Die wichtigsten Empfindungen für die perzeptive Untersuchung nichtlinearer Lautsprecherverzerrungen sind durch die psychoakustischen Größen *Lautheit*, *Schärfe* und *Rauigkeit* repräsentiert. Verschiedene Domänen des Schallreizes sprechen dabei unterschiedliche Größen an: Beispielsweise unterliegt die Schärfe hauptsächlich spektralen Einflüssen, während die Rauigkeit vom Zeitverlauf der Einhüllenden eines Signals bestimmt ist.

2.2.1 Lautstärkepegel und Lautheit

Betrachtet man die Abhängigkeit der Hörschwelle von der Frequenz eines Schallreizes liegt der Schluss nahe, dass die Lautstärkeempfindung ebenfalls frequenzabhängig variable Schalldruckpegel erfordert. Um diesen Umstand zu erfassen wurden für reine Töne die *Kurven gleicher Lautstärkepegel* erfasst und in DIN ISO 226 genormt: Sie spiegeln abhängig von der Frequenz denjenigen Schalldruck wieder, der eine gleiche Lautstärkeempfindung hervorruft (vgl. ebd., S. 607 ff. & vgl. Dickreiter, 1997, S. 110). Davon wurde der in *phon* angegebene Lautstärkepegel L_N abgeleitet, welcher für einen Sinuston mit 1 kHz gleich dem Schalldruckpegel ist (vgl. Dickreiter, 1997, S. 110). Der Lautstärkepegel dient daher als Vergleichsmaß. Einem zu qualifizierenden Stimulus kann somit der Schalldruckpegel eines Referenztones mit 1 kHz zugeordnet werden, der die gleiche Lautstärkeempfindung erzeugt. Ein alternatives Maß stellt die *Lautheit N* dar: Ausgehend von der Annahme, dass über 40 phon eine Steigerung um 10 phon eine Verdopplung der Lautheitsempfindung erzeugt, kann der Lautstärkepegel direkt in die in der Einheit *sone* angegebene Lautheit umgerechnet werden (vgl. Maschke & Jakob, 2010, S. 609 f.). Die Erläuterungen beziehen sich auf frontale Beschallung von Versuchspersonen im Freifeld und können mit einer Wichtungsfunktion auf die Beschallung im Diffusfeld angewendet werden (vgl. Dickreiter, 1997, S. 112). Weiterhin ist die Reizdauer von Bedeutung: Für quasi-stationäre Signale stellt sich ab einer Gewissen Länge des Stimulus ein konstanter Lautstärkepegel ein, der bei Unterschreitung einer Mindestdauer linear abnimmt. Maschke & Jakob geben hierfür als kritische Grenze 200 ms an (vgl. 2010, S. 606).

Der Bezug zu reinen Tönen ist problematisch, da für spektral komplexe und zeitlich dynamische Stimuli wie Musik die Lautstärkepegel nur begrenzt anwendbar sind. Dieses Problem versuchten Fastl & Zwicker zu lösen, indem sie Probanden den Schalldruckpegel eines 1-kHz-Referenztones mit dem Pegel eines Uniform Exciting Noise (im folgenden UEN) verglichen ließen, der die gleiche Lautstärkeempfindung hervorruft. Das UEN stellt dabei ein Breitbandrauschen dar, welches spektral gewichtet wurde, um für jede Frequenzgruppe die gleiche Lautstärkeempfindung zu erzeugen. Zunächst waren die Versuchsteilnehmer aufgefordert den Pegel des Referenztones entsprechend ihrer Lautstärkeempfindung an den des UEN anzupassen. In einer weiteren Testreihe musste der UEN an den Referenzton angepasst werden. Für diese gegensätzlichen Stimuli stellten sie mit der verwendeten dabei gravierende Abweichungen von der Annahme Methode fest. dass ein 1-kHz-Referenzton mit entsprechendem Schalldruckpegel immer den Lautstärkepegel eines komplexen Stimulus widerspiegelt. (vgl. zu diesem Absatz Fastl & Zwicker, 2007, S. 208 ff.)

Diese Erkenntnisse und weitere Untersuchungen zum Einfluss der Bandbreite von Rauschsignalen und Verdeckungseffekten führten zu einem fortgeschritten Lautheitsmodell (vgl. ebd., S. 210 ff.). Dieses Lautheitsmodell zieht die Erregung der Frequenzgruppen in gehörgerechte Lautheitsbewertung Betracht: Eine numerische ist demnach das Flächenintegral der Erregungsmuster über alle Frequenzgruppen (vgl. ebd., S. 221 ff.). Das entspricht einer einfachen Summation, wenn die spezifischen Lautheiten für jede Frequenzgruppe vorliegen. Die Einheit dieser Gesamtlautheit ist dann sone/Bark (vgl. ebd., ausführliche Beschreibung Rechenmodells S. 222). Eine eines ist neben den Veröffentlichungen von Fastl & Zwicker auch bei Maschke & Jakob zu finden (vgl. 2010, S. 610 ff.).

Für die perzeptive Evaluation der Lautsprecherwiedergabe ist die Lautheit eine relevante Größe: Beispielsweise kann ein verzerrter Amplitudenfrequenzgang die spezifischen Lautheiten einiger Frequenzgruppen beeinflussen und somit auch einen Einfluss auf die Differenz der Gesamtlautheit gegenüber einem Referenzsignal haben.

2.2.2 Schärfe

Einige deskriptive Komponenten der Hörempfindung werden häufig zusammen in dem Begriff Klangfarbe (vgl. Dickreiter, 1997, S. 115) bzw. Timbre (vgl. Moore, 2007, S. 483) ausgedrückt. Die *Schärfe* eines Klanges kann dort eingeordnet werden, wird aber zumeist separat davon betrachtet, da für diese psychoakustische Größe Verhältnismaße bestimmt werden können. So können nach Fastl & Zwicker Stimuli verglichen werden, die doppelt oder halb so scharf wie ein Referenzschall klingen (vgl. Fastl & Zwicker, 2007, S. 239).

Als Formelzeichen wird S und als Einheit der vom Lateinischen abgeleitete Begriff acum verwendet. Die Normierung auf 1 acum erfolgt mit einem Schmalbandrauschen mit der

Mittenfrequenz 1 kHz und einer Bandbreite von 160 Hz bei 60 dB SPL Schalldruckpegel (vgl. Fastl & Zwicker, 2007, S. 239 & vgl. Maschke & Jakob, 2010, S. 615).

Die Normierung spiegelt bereits wichtige Eigenschaften der Schärfe wider: Die Größe ist stark davon abhängig, wie die spezifischen Lautheiten innerhalb der Frequenzgruppen ausfallen. Dabei stellten Fastl & Zwicker heraus, dass es vernachlässigbar ist, ob der Klang ein dichtes Rauschen ist oder nur aus einzelnen spektralen Linien besteht, solange die spezifischen Lautheiten konstant gehalten werden (vgl. Fastl & Zwicker, 2007, S. 239). Ferner ist die Schärfe vor allem von der Form der spektralen Einhüllenden eines Schalles bestimmt (vgl. Maschke & Jakob, 2010, S. 615). Die Bandbreite schmalbandiger Schalle spielt dabei keine Rolle, solange die kritische Bandbreite der zugehörigen Frequenzgruppe nicht überschritten wird (vgl. Fastl & Zwicker, 2007, S. 239). Weiterhin sind kleine Differenzen der Schalldruckpegel zweier Stimuli vernachlässigbar, da eine Verdopplung der Schärfe für das zur Normung genutzte Rauschen eine Pegeldifferenz von immerhin 60 dB erfordert (vgl. ebd.).

Das spektrale Zentrum eines Schalles hat hingegen deutlich größeren Einfluss: Für gängige Berechnungsmodelle, wie z.B. nach Bismarck (vgl. Maschke & Jakob, 2010, S. 616 f.) oder Fastl & Zwicker (vgl. 2007, S. 241 ff.), werden für spektrale Anteile bei hohen Frequenzen Wichtungsfunktionen benutzt. Die Wichtungsfunktionen erhöhen den Einfluss auf die berechnete Schärfe gegenüber Anteilen bei niedrigeren Frequenzen. Die Schärfe eines hochfrequenten Testschalls kann trotzdem durch einen simultan präsentierten niederfrequenten Schall reduziert werden kann (vgl. ebd., S. 241). Die Berechnung der absoluten Schärfe eines Schalles wurde in der Norm E DIN 45692 festgelegt (vgl. Maschke & Jakob, 2010, S. 617). Ausführliche Beschreibungen einiger Rechenmodelle können bei Fastl & Zwicker (vgl. 2007, S. 241 ff.) und Maschke & Jakob (vgl. 2010, S. 615 ff.) gefunden werden.

Die Schärfe ist hinsichtlich der perzeptiven Bewertung von Lautsprechern relevant, da durch verschiedene Nichtlinearitäten Verzerrungen erzeugt werden, welche die Differenz der Schärfe gegenüber einem Referenzsignal beeinflussen können.

2.2.3 Rauigkeit

Die *Rauigkeit* stellt gemeinsam mit der Schwankungsstärke eine psychoakustische Größe dar, die von Veränderungen der Einhüllenden des Schallreizes im Zeitbereich bedingt ist. Ähnlich der Schärfe können auch für die Rauigkeit Verhältnismaße bestimmt werden (vgl. Maschke & Jakob, 2010, S. 618). Weiterhin eignet sich die recht eindeutige Konnotation des Adjektivs *rau* gut zur Beschreibung von Klangeigenschaften (vgl. Dickreiter, 1997, S. 115).

Die folgenden Ausführungen basieren auf der Veröffentlichung von Fastl & Zwicker (vgl. 2007, S. 257 ff.). Veränderungen in der Hüllkurve eines Zeitsignals können unterschiedlich entstehen. Der zunächst naheliegende Fall ist die Amplitudenmodulation eines Testtones mit der Frequenz f_c (folgend: Träger) durch einen weiteren Ton mit der Frequenz f_m (folgend: Modulator). Wird die Frequenz des Modulators von 0 Hz beginnend erhöht, stellen sich verschiedene Hörempfindungen ein: Für einen Träger mit 1 kHz erzeugen Modulatoren mit $0 \le f_m \le 15$ Hz die Hörempfindung, dass die Lautheit zeitlich schwankt. Diese Empfindung wird als Schwankungsstärke bezeichnet und hat ihr Maximum für den beschrieben Fall bei $f_m = 4$ Hz. Wird die Modulationsfrequenz weiter erhöht, verändert sich die Hörempfindung zur Rauigkeit, die ihrerseits für $f_c \ge 1$ kHz ihr Maximum bei $f_m = 70$ Hz

hat. Für tiefere Trägerfrequenzen verschieben sich die Maxima zu tieferen Modulationsfrequenzen.

Bei höheren Modulationsfrequenzen nimmt der Rauigkeitseindruck wieder ab. Hier liegen die durch Intermodulation entstehenden Seitenbänder des Spektrums weit genug auseinander, sodass der Höreindruck nicht mehr rau ist. Vielmehr werden mehrere Töne gehört. Dies ist einerseits in der begrenzten zeitlichen Auflösung des Gehörs als auch in der Bandbreite der betreffenden Frequenzgruppen begründet. Darüber hinaus ist die Modulationstiefe gekennzeichnet durch den Modulationsindex m von Bedeutung: m = 1 ist äquivalent zu einem vollständig amplitudenmodulierten Signal. Ab einem Wert von $m \approx 0.07$ kann Rauigkeit von Hörern detektiert werden. (vgl. zu diesem Absatz Fastl & Zwicker, 2007, S. 257 ff.).

Während die Modulationsfrequenz und der Modulationsindex starken Einfluss auf die Rauigkeit haben, wirkt sich die Änderung des Schalldruckpegels ähnlich der Schärfe relativ gering aus: Maschke & Jakob geben an, dass erst bei einer Erhöhung des Schalldruckpegels um 40 dB eine verdoppelte Rauigkeit erfasst werden kann (vgl. 2010, S. 618). Darüber hinaus erzeugen frequenzmodulierte Signale drastisch höhere Rauigkeiten.

Die Rauigkeit trägt das Formelzeichen *R* und kann in der Einheit *asper* gemessen werden. Sie ist nach Zwicker & Fastl mit einem Referenzschall mit $f_c = 1$ kHz, $f_m = 70$ Hz und m = 1 bei einem Schalldruckpegel von 60 dB SPL auf 1 asper zu normieren (vgl. 2007, S. 257).

Die Funktionsmechanismen der Rauigkeit scheinen noch nicht vollständig erfassbar zu sein, was sich in der Vielzahl existierender Berechnungsmodelle zeigt: Fastl & Zwicker führten ein Modell ein, welches ausgehend von einer zeitlichen Maskierungsschwelle und der Modulationsfrequenz Rauigkeiten berechnet (vgl. ebd., S. 261). Aures stellte einen Ansatz mit einer gehörgerechten Filterbank gemäß der Frequenzgruppen nach Zwicker & Fastl vor, welcher neben anderen Neuerungen auch eine Korrelation benachbarter Frequenzgruppen berücksichtigt (vgl. Aures, 1985). Daniel & Weber erweiterten dieses Modell, welches anstatt Zwickers Frequenzgruppen der Barkskala mit 1 Bark Bandbreite ein Modell mit 47 Frequenzgruppen (Breite 0.5 Bark) und einer zeitlichen Diskretisierung der Signale in Blöcken der Länge t = 0.2 s arbeitet (vgl. 1997). Duisters stellte darüber hinaus gravierende Änderungen anderer Autoren am Modell von Aures zusammen und implementierte ein Maß Modulationsfrequenz zur Feststellung der (vgl. 2005). Weiterhin wurden anwendungsbezogene Modelle entwickelt: So implementierte Schulze ein Modell zur Bewertung der Betriebsgeräuschen von Haushaltsgeräten (vgl. 2013) und Martner et. al. Berechnungsverfahren entwickelten ein zur Bestimmung der Rauigkeit bei Motorengeräuschen (vgl. 2000).

Rauigkeiten sind bei der Lautsprecherwiedergabe immer dann relevant, wenn ein Treiber hohe und Tiefe Frequenzen bei größeren Auslenkungen wiedergeben soll: Die entstehenden Doppler- und Intermodulationsverzerrungen erzeugen einen rauen Klangeindruck. (vgl. Dickreiter, 1997, S. 213)

2.3 Der Elektrodynamische Lautsprecher

Der elektrodynamische Lautsprecher ist ein weit verbreiteter Schallwandler der elektrische Signale in akustische Signale überführt. Das Konstruktionsprinzip ist wirtschaftlich zu fertigen, erzielt aber im Vergleich zu anderen elektromechanischen Systemen nur mäßige Wirkungsgrade (vgl. Dickreiter, 1997, S. 205). In Abbildung 2 werden die wesentlichen Komponenten eines elektrodynamischen Lautsprechers am Beispiel eines Konuslautsprechers benannt:



Abbildung 2: Schnittbild eines Konuslautsprechers. (Möser, 2012, S. 374)

Eine konusförmige Membran ist an einer Begrenzung mit Sicken in einem Korb weich aufgehängt. An die andere Begrenzung schließt sich der Schwingspulenträger mit der Schwingspule an. Die Membran kann sich linear bewegen und wird von der Zentrierspinne geführt. Der Schwingspulenträger ragt durch einen Spalt in das am Korb befestigte Magnetgehäuse bestehend aus den Polplatten und dem Permanentmagneten hinein.

Weitere Bauformen existieren, welche auf das gleiche Prinzip zurückzuführen sind: Die Bewegung der Membran entsteht durch die Kraftwirkung, die ein stromdurchflossener Leiter in einem Magnetfeld erfährt (vgl. ebd.).

2.3.1 Wirkprinzip

Um die Auslenkung der Membran zu erzeugen wird die Lorentzkraft genutzt, die auf stromdurchflossene Leiter in Magnetfeldern wirkt (vgl. Goertz, 2008, S. 424). Das Magnetfeld hat die magnetische Flussdichte B, der Leiter die vom Magnetfeld durchflossene Länge l und der Strom die Größe i. Die Größen sind vektoriell zu betrachten und für den Fall, dass sie orthogonal zueinander wirken wird die Kraft F nach Gl. 1 maximal:

$$F = B \cdot l \cdot i$$
 Gl. 1

Wird als Strom i eine zeitlich veränderliche Größe mit Vorzeichenwechsel angelegt, so ändert sich nicht nur der Betrag, sondern auch die Wirkungsrichtung der Kraft F proportional zu i. Es gilt hierbei zur Vereinfachung zunächst die Annahme dass die Flussdichte B und die wirksame Leiterlänge l konstant sind.

Im Falle des Konuslautsprechers aus Abbildung 2 ist der Leiter die Schwingspule im Spalt zwischen Polkern und Polplatte. Die wirksame Länge l ist die Länge der Windungen der Schwingspule, die sich zum jeweiligen Zeitpunkt im homogen angenommen Magnetfeld

befindet. Das Magnetfeld mit der Flussdichte B wird vom verwendeten Permanentmagneten bestimmt und der Strom i von der an den Lautsprecher angeschlossenen Leistungselektronik.

Die Beschreibungen umreißen nur einen sehr grundlegenden Teil der involvierten Mechanismen. Die Vielzahl der involvierten Komponenten erfordert ein komplexes Modell um die mechanische und elektrische Domäne sowie die Transformation zwischen beiden ganzheitlich zur erfassen. Beide Domänen sind Schwingkreise, für die sich die Betrachtung in einem Ersatzschaltbild gemäß Abbildung 3 eignet.



Abbildung 3: elektromechanisches Ersatzschaltbild für elektrodynamische Lautsprecher (Klippel GmbH, 2013a, S. 4)

Die folgenden Ausführungen lehnen sich an den Besuch der Blockvorlesung "Sound Quality of Audio Systems" und den zugehörigen Unterlagen (vgl. Klippel, 2016) sowie den Onlinetrainings der Klippel GmbH (vgl. 2013a, S. 2 ff.) an:

Die linke Hälfte des Ersatzschaltbildes stellt die elektrische Domäne dar. Die Größe R_e ist der Gleichstromwiderstand der Schwingspule bzw. entspricht dem Wert der komplexen Impedanz $Z_L(f)$ bei f = 0 Hz. Diese Impedanz ist frequenzabhängig und berücksichtigt neben der Induktivität der Schwingspule auch Verluste durch Wirbelströme. Die an die Polklemmen angelegte Spannung U mit dem zugehörigen Strom i ist eine Wechselgröße und wird über den Kraftfaktor Bl an die mechanische Domäne gekoppelt. Der Kraftfaktor stellt dabei einen Transformator zwischen den Schwingkreisen dar.

Die mechanische Domäne ist ein komplexes Masse-Feder-System. Die bewegte Masse M_{ms} setzt sich aus den Massen aller bewegten Teile zusammen und entspricht einer Kapazität. Für das Kräftegleichgewicht ist der bewegten Masse bei Beschleunigung die Trägheitskraft F_M zugeordnet. Die Federsteife wird durch die Steifigkeit der Zentrierspinne und der Aufhängung der Membran bestimmt und entspricht im Ersatzschaltbild der induktiven Größe K_{ms} . Diese erzeugt eine Rückstellkraft F_K , welche die Schwingspule zur Ruheposition streben lässt. Gleichbedeutend ist das Reziproke von K_{ms} die mechanische Nachgiebigkeit der Anordnung C_{ms} (vgl. Schulze , 2014, S. 8). Weiterhin ist der mechanische Widerstand R_{ms} zu berücksichtigen, dessen zugehörige Kraft F_R zu einer Dämpfung des Systems führt. Die auslenkende Kraft F_{coil} entspricht der Lorentzkraft.

Die mechanische Auslenkung x bzw. die Schnelle v der Spule ist das Ergebnis der Kopplung der Domänen durch den Kraftfaktor *Bl*. Es besteht eine beidseitige Wechselwirkung, denn durch die Auslenkung der bewegten Teile entsteht eine Rückwirkung in die elektrische Domäne: Der Kraftfaktor erzeugt mit der Schnelle an den Polklemmen eine Spannung, die der angelegten Spannung U entgegenwirkt. Dieser Effekt wird als *elektromotorische Gegenkraft* bezeichnet bzw. im Englischen mit *BackEMF* abgekürzt.

Ausgehend von den Komponenten der Teilsysteme lassen sich wichtige Parameter nach **Tma Mode ing**chnen, die als Kenndaten zur Beschreibung von elektrodynamischen Lautsprechern genutzt werden. Beispielsweise repräsentiert die Resonanzfrequenz f_s die **med Mada**zfrequenz des mechanischen Schwingkreises, während die Gütefaktoren Q_{es} und Q_{ms} die Verluste der elektrischen und mechanischen Teilsysteme beschreiben und im Gesamtgütefaktor Q_{ts} zusammengefasst werden. (vgl. Klippel, 2016 & vgl. Klippel GmbH 2013a, S. 2 ff.)

2.3.2 Reguläre Nichtlinearitäten

Die zuvor eingeführten Größen des elektrodynamischen Lautsprechers ermöglichen es dessen Übertragungsverhalten umfassend zu beschreiben. Sofern diese nur als Konstanten angenommen werden beschränkt sich dies aber auf den linearen Teil der Lautsprecherwiedergabe. Dies genügt für die Prädiktion der Zustandsgrößen im Kleinsignalbereich mit typischen Auslenkungen $x \leq 1$ mm; für den Großsignalbereich ist darüber hinaus die Kenntnis über Verlauf und Abhängigkeiten der Kennlinien einzelner beteiligter Größen von Bedeutung (vgl. Klippel, 2016, S. 10). Die eingangs erläuterten häufäge Betriebsbedingungen erfordern Plas Betreiben von Lautsprechern im Großsignalbereich, indem der Einfluss der Nichtlinearitäten auf das Ausgangssignal signifikant ist. Sie werden als regulär betrachtet, wenn ein unmittelbarer Zusammenhang zwischen der Kennlinie der jeweiligen Nichtlinearität, dem Wirkprinzip des Treibers und seiner Geometrie und dem Grad einer spezifischen Verzerrung festgestellt werden kann und kein stochastischer Charakter vorliegt (Vgl. Klippel, 2006, S. 5).

Die Nichtlinearitäten erzeugen im akustischen Ausgangssignal des Lautsprechers Teiltöne, welche nicht im eingespeisten elektrischen Signal enthalten sind. Die Gestalt der Kennlinie und der Grad der Abweichung vom linearen Teil bestimmt dabei das Amplitudenspektrum der erzeugten Teiltöne. Die Kennlinien sind messtechnisch erfassbar und können durch die Überlagerung der linearen Kennlinien mit quadratischen, kubischen oder weiteren Komponenten höherer Ordnungen mathematisch beschrieben werden. (vgl. Müller, 2008, S. 1143).

Für die Musikwiedergabe in Systemen mit nur einem Treiber sind die Nichtlinearitäten besonders kritisch, da das Tongemisch der Quellfrequenz eine Vielzahl von Obertönen erzeugt (vgl. ebd., S. 1144). Diese addieren "sich mit den Intermodulationsartefakten von hohen Teiltönen durch tiefe Frequenzen und führen zu Amplitudenspektren mit unharmonischen Verhältnissen (vgl. Dickreiter, 1997, S. 213 f.).

Für den Großsignalbereich ist das Ersatzschaltbild aus Abbildung 3 zu adaptieren:



Abbildung 4: Ersatzschaltbild für Großsignalbereich (Klippel GmbH, 2015a, S.2)

Die dominierenden Nichtlinearitäten bei elektrodynamischen Lautsprechern sind primär von den Zustandsgrößen Auslenkung x und fließendem Strom i abhängig. Weiterhin haben auch

Temperaturänderungen und Materialkonstanten erheblichen Einfluss. Nachfolgend wird eine für das Verständnis der folgenden Kapitel notwendige Auswahl nichtlinearer Größen betrachtet. Die Bezeichnungen werden gemäß Abbildung 4 gewählt.

$K_{ms}(x)$

Die Steifigkeit $K_{ms}(x)$ der Aufhängung von Membran und Spulenkörper durch Spinne und Sicke ist der reziproke Wert der mechanischen Nachgiebigkeit $C_{ms}(x)$ und nicht konstant. Gegenüber einer linearen Federkennlinie ist die Steifigkeit eine Funktion der Auslenkung. Die resultierende Rückstellkraft ist gemäß Gl. 2 das Produkt der Auslenkung und der abhängigen Steifigkeit:

$$F_{K} \Box K_{ms}(x) \cdot x \qquad \qquad \text{Gl. 2}$$

Demnach trägt $K_{ms}(x)$ die Einheit N/mm. (vgl. Klippel , 2006, S. 6).

Spinne und Sicke tragen unterschiedlich zur Kennlinie bei. Während die Steifigkeit der Sicke über kleine und mittlere Bereich der Auslenkung annähernd linear bleibt und erst nahe der maximalen Auslenkung stark ansteigt, ist die Steifigkeit der Spinne häufig ähnlich einer quadratischen Funktion und schon in mittleren Auslenkung stark nichtlinear. Beide Teile werden addiert um die Gesamtsteifigkeit der Aufhängung zu erhalten. (vgl. ebd.).

Es besteht ein direkter Zusammenhang zwischen den Größen der Teilsteifigkeiten und der Geometrie der zugehörigen Bauteile: Der Querschnitt der Zentrierspinne gleicht einer Feder, welche durch Rotation um die z-Achse zu einer Art Flächenfeder wird. Diese wird bei Auslenkung kontinuierlich gelängt, was eine ansteigende Kraft erfordert. Folglich verhält sich die Kennlinie der Steifigkeit der Spinne schon bei mittleren Auslenkungen nichtlinear. Die Sicke hingegen wird über weite Bereiche der Auslenkung zwar in ihrer räumlichen Orientierung verändert, aber zunächst nicht gedehnt. Erst bei großer Auslenkung wird die Sicke gedehnt und demzufolge steigt auch erst ab diesem Punkt die zugehörige Steifigkeit signifikant an. Abbildung 5 illustriert diesen Zusammenhang:



Abbildung 5: Steifigkeit von Spinne, Sicke und Aufhängung (Klippel, 2006, S. 7)

spider: Spinne, surround: Sicke, total suspension: gesamte Aufhängung.

Die Asymmetrie der Kennlinie kann einen frequenzabhängigen Gleichanteil in der Auslenkung erzeugen, der bei Steifigkeitsverzerrungen sein Maximum bei der mechanischen Resonanzfrequenz erreicht. Akustische Symptome nichtlinearer Steifigkeit beschränken sich auf harmonische Verzerrungen bei tiefen Frequenzen. (vgl. ebd., S. 42).

Bl(x)

Der Kraftfaktor *Bl* ist abhängig von der magnetischen Flussdichte *B* und der Leiterlänge *l* der Schwingspule im homogen anzunehmenden Magnetfeld im Spalt zwischen Polkern und Polplatte. Wird die Schwingspule durch den Stromfluss ausgelenkt ändert sich u.U. die Leiterlänge im Spalt; speziell bei großen Auslenkungen nimmt diese ab. Entsprechend ist auch der Kraftfaktor abhängig von der Auslenkung *x* und wird mit Bl(x) angegeben. (vgl. Klippel, 2006, S. 8 f.)

Auch hier beeinflusst die Geometrie der Bauteile maßgeblich die Kennlinie, denn das Verhältnis der Höhe des Spaltes und der Höhe der Schwingspule bestimmt den Anteil und den Einfluss der nichtlinearen Komponenten. Sind beide Maße gleich groß nimmt die Leiterlänge im Spalt schon bei kleinen Auslenkungen ab. Ist die Höhe der Schwingspule größer als die des Spaltes bleibt die Leiterlänge über einen bestimmten Bereich der Auslenkung konstant, folglich hat auch die Kennlinie des Kraftfaktors ein Plateau und man spricht von einer Überhangkonfiguration. (vgl. ebd.)

Zur Veranschaulichung zeigt Abbildung 6 ein Schnittbild durch Spule und Magnetkörper:



Abbildung 6: Schnitt durch Spule und Magnetkörper (vgl. Klippel GmbH, 2014a, S. 4)

Die Symptome des nichtlinearen Kraftfaktors sind ein nicht-proportionaler Zusammenhang zwischen der Auslenkung x und dem gespeisten Strom i und eine nichtlineare Dämpfung durch die elektromotorische Gegenkraft (vgl. Klippel GmbH, 2014a, S. 3). Eine asymmetrische Bl(x)-Kennlinie kann ebenfalls einen Gleichanteil in der Auslenkung erzeugen. In diesem Falle richtet sich der Versatz unterhalb der Resonanzfrequenz zum Bl-Maximum aus und hat bei der Resonanzfrequenz den kleinsten Betrag. Sehr starke Asymmetrien können zu instabilem Verhalten des Lautsprechers oberhalb der Resonanzfrequenz führen.

Akustische Symptome sind neben harmonischen Verzerrungen bei niedrigen Frequenzen und großen Auslenkungen auch signifikant Intermodulationsverzerrungen im gesamten hörbaren Spektrum. Bei Zweiton-Signalen können dabei unterschiedliche Phänomene beobachtet werden, je nachdem ob der Basston oder der Stimmton konstant gehalten wird und der jeweils andere variiert wird. Für Stimmtonsweeps werden dabei im gesamten Audioband konstant Intermodulationsverzerrungen erzeugt. Für Basstonsweeps passiert dies bis zur halben Resonanzfrequenz, darüber sinken die Intermodulationsverzerrungen ab. (vgl. Klippel , 2006, S. 43).

 $L_e(x,i)$

Die Induktivität der Schwingspule L_e wird durch einen komplexen Zusammenhang zwischen anregendem Strom *i* und Auslenkung *x* nichtlinear beeinflusst. Die Induktivität sinkt für positive Auslenkungen, da sich die Schwingspule dann in Bereichen mit verringerter magnetischer Flussdichte befindet (vgl. ebd., S. 10).

Gleichzeitig beeinflusst der angelegte Strom die Induktivität, da er die magnetische Permeabilität μ verändert und somit indirekt die magnetische Flussdichte *B* nach Gl. 3 verändert (vgl. Klippel, 2006, S. 11):

$$B = \mu(i) \cdot H$$
 Gl. 3

Der Wechselstrom erzeugt hierbei ein magnetisches Wechselfeld, welches der Grundpermeabilität des Dauermagneten eine reversible Permeabilität überlagert und somit die vom Strom abhängige Permeabilität $\mu(i)$ bildet (vgl. Oberdorfer, 1984, S. 340 f.). *H* sei hierbei die Feldstärke ohne Bestromung und Auslenkung, welche den Arbeitspunkt der Induktionskennlinie charakterisiert (vgl. Klippel GmbH, 2014 S. 5). Die Einheit der Induktivität ist *mH*.

Ein Nebeneffekt der Überlagerung des zunächst statischen Magnetfeldes des Permanentmagneten mit dem magnetischen Wechselfeld der bestromten Schwingspule sind Eisenverluste, die auf der Hystereseschleife des Permanentmagneten bei Ummagnetisierung beruhen und ihrerseits nichtlinear auf die Induktivität einwirken (vgl. Oberdorfer, 1984, S. 446 f.). Diese Effekte können nur durch komplexere Modelle mit den Größe $L_2(x, i_2)$ und $R_2(x)$ aus Abbildung 4 mittel eines physikalischen Ansatzes berücksichtigt werden und verursachen den typischen Anstieg des Impedanzgangs eines Lautsprechers bei höheren Frequenzen (vgl. Schulze, 2014, S. 9 f.).

Die Symptome der nichtlinearen Induktivität $L_e(x,i)$ sind in der elektrischen und akustischen Domäne erfassbar. Die Abhängigkeit von der Auslenkung erzeugt eine Gegeninduktion in der elektrischen Domäne, die ihrerseits wie die elektromotorische Gegenkraft eine nichtlineare elektrische Dämpfung hervorruft. Die durch den Strom verursachten nichtlinearen Effekte erzeugen akustische Verzerrungen über das komplette hörbare Spektrum. Als Nebeneffekt beider Abhängigkeiten entsteht überdies eine Reluktanzkraft $F_m(x, i, i_2)$ entlang der Richtung der Lorentzkraft, wodurch ebenfalls im kompletten hörbaren Spektrum Verzerrungen erzeugt werden. (vgl. Klippel, 2006, S. 14)

Bei den hierdurch erzeugten Verzerrungen handelt es sich in der akustischen wie in der elektrischen Domäne vorwiegend um Intermodulationsverzerrungen. Nichtlineare $L_{e}(x)$ -Kennlinien erzeugen für Stimmtonsweeps im hörbaren Spektrum mit der Frequenz im Pegel ansteigende Intermodulationsverzerrungen. Für Stimmtonsweeps ähnelt das Verhalten den Bl(x)-Verzerrungen, jedoch sind diese hierbei auch im Strom messbar. Für den nichtlinearen Zusammenhang von Induktivität und Strom $L_e(i)$ zeigt sich für Basstonsweeps das gleiche Verhalten wie bei den $L_e(x)$ -Verzerrungen. Stimmtonsweeps erzeugen hier akustische Verzerrungen die unterhalb der Resonanzfrequenz des Lautsprechers maximal sind, bei der Resonanzfrequenz ein Minimum haben und oberhalb der doppelten Resonanzfrequenz verschwinden. Ein Gleichanteil der Auslenkung wird nur durch nichtlineare $L_e(x)$ -Kennlinien erzeugt. Dieser wird oberhalb der Resonanzfrequenz maximal und strebt in Richtung des $L_e(x)$ -Maximums. (vgl. ebd., S. 43 f.)

2.3.3 Thermische Effekte

Bei der Mehrheit elektromechanischer Systeme sind Verluste im Wirkungsgrad nicht unerheblich mit der Umwandlung elektrischer Energie in thermische Energie verbunden. Auch beim elektrodynamischen Lautsprecher haben thermische Verluste signifikante Einflüsse. Eine ausführliche Betrachtung der thermischen Einflüsse sind bei Schulze zu finden (vgl. 2014, S. 22 ff.) und hier nur stark verkürzt wiedergegeben. Als einführendes Beispiel sei die Änderung des elektrischen Widerstandes R_e der Schwingspule erläutert: Dieser ist abhängig von der Temperatur der Schwingspule und ergibt sich gemäß Gl. 4 zu:

$$\mathbf{R}_{e}(T_{v}) = \mathbf{R}_{e}(T_{a}) \cdot (1 + \delta \cdot \Delta T_{v})$$
Gl. 4

 $R_e(T_a)$ sei hierbei der elektrische Widerstand bei Umgebungstemperatur, δ die thermische Leitfähigkeit und ΔT_v die Änderung der Schwingspulentemperatur. Über das Ohm'sche Gesetz führt dieser Zusammenhang zu einer geringeren Stromaufnahme bei steigender Schwingspulentemperatur. In Folge dessen sinkt die auslenkende Kraft. (vgl. Schulze, 2014, S. 14)

Dieses triviale Beispiel illustriert, dass für die Simulation der Zustandsvariablen eines Lautsprechers bei beliebiger Anregung der Einfluss thermischer Verluste nicht zu vernachlässigen ist. Neben der Schwingspulentemperatur ändert sich durch Wärmetransfer und Wirbelströme auch die Temperatur von Magnet und Polplatten. Weiterhin entstehen mit der wechselnden Auslenkung x der Membran durch die bewegten Luftmassen Effekte der Konvektionskühlung.

Die Berücksichtigung aller wichtigen thermischen Effekte geschieht nach Klippel in einem Modell aus verschalteten RC-Gliedern (vgl. ebd., S. 22). Jedes RC-Glied hat dabei ein Teilsystem mit dem zugehörigen thermischen Widerstand, der jeweiligen Wärmekapazität und der entsprechenden Zeitkonstante um die Erhitzung der Teilsysteme abzubilden. Im Gegensatz zu anderen Modellen werden in diesem Modell auch Kühleffekte berücksichtigt.

Die Abhängigkeit des thermischen Modells von verschiedenen Größen wie der Auslenkung x, Membranschnelle v oder der Umgebungstemperatur erzeugt nichtlineares Verhalten. Überdies zeigt bereits das eingangs vereinfachte Beispiel des Einflusses der Schwingspulentemperatur T_v auf den elektrischen Widerstand R_e , dass eine Rückkopplung des thermischen Modells auf die elektromechanischen Zustandsvariablen besteht. (vgl. ebd.)

2.3.4 Irreguläre Nichtlinearitäten

Das Auftreten von Symptomen irregulärer Nichtlinearitäten hat stochastischen Charakter und ist zumeist auf verschiedene Defekte zurückzuführen (vgl. Klippel, 2006, S. 18). Zu den wichtigsten Verursachern zählen vibrierende oder lose Teile, auf die Spinne oder Membran schlagende Anschlusslitzen, lose Partikel auf der Membran oder im Spalt, Undichtigkeiten im Gehäuse und ein durch Kippschwingungen hervorgerufenes Reiben der Schwingspulen bzw. des Schwingspulenträgers im Spalt (vgl. ebd., S. 18, & vgl. Müller, 2008, S. 1151). Die Symptome weisen häufig impulshaften Charakter auf, da sie i.d.R. erst bei höheren Auslenkungen angeregt werden und bei geringen Auslenkungen nicht vorhanden sind.

Diese Phänomene werden unter dem Begriff "Rub & Buzz" zusammengefasst. Ihre Auswirkungen sind zumeist breitbandige Verzerrungen im hörbaren Spektrum wodurch z.B. eine herkömmliche Betrachtung der harmonischen Verzerrungen zur Detektion solcher Defekte nicht ausreicht. (vgl .Müller, 2008, S. 1151 f.)

2.3.5 Psychoakustische Relevanz

Wie bereits festgestellt können $K_{ms}(x)$ und Bl(x) in erheblichem Maße harmonische Verzerrungen hervorrufen. Diese wirken sich in laut Dickreiter erst bei beträchtlichen Anteilen auf eine empfundene Änderung der Klangfarbe aus (vgl. 1997, S. 213).

Losgelöst vom Konstrukt der Klangfarbe kann aber wie zuvor erklärte der Lautheitseindruck betrachtet werden. Da sowohl $K_{ms}(x)$ als auch Bl(x) die als Klirr bekannten harmonischen Verzerrungen hauptsächlich für tiefe Frequenzen erzeugen, kann ein Hörer bis zu einem gewissen Grad der Verzerrung im Paarvergleich einen verzerrten Klang als angenehmer empfinden. In jedem Falle ändert das Addieren der Verzerrung die spezifischen Lautheiten der jeweiligen Frequenzgruppen. Es ist festzuhalten, dass in der Praxis der Stimulus mit dem erhöhten Lautheitseindruck häufig gegenüber dem nicht verzerrten Stimulus präferiert wird.

Eine weitere dominierende Verzerrungsform sind Intermodulationsverzerrungen. Die Nichtlinearitäten Bl(x) und $L_e(x,i)$ erzeugen im akustischen Ausgangsignal des Lautsprechers Summen- und Differenztöne, die durch Intermodulation von verschiedenen spektralen Anteilen entstehen. Darüber hinaus kann ein stark nichtlinearer Kraftfaktor Bl(x) Amplitudenmodulationsverzerrungen hervorrufen (vgl. Klippel, 2006, S. 43). Beide Phänomene sprechen die Empfindungsgröße der Rauigkeit an. Ein Einfluss der empfundenen Rauigkeit durch nichtlineare Kraftfaktor- und Induktivitätsverzerrungen ist somit gegeben.

Der breitbandige Charakter der Verzerrungen, die von irregulären Nichtlinearitäten erzeugt werden, kann die spektrale Zusammensetzung des verzerrten Signals gegenüber dem nicht verzerrten Signal erheblich verändern. Unter Umständen wird die spektrale Einhüllende dadurch vor allem bei hohen Frequenzen verändert, die somit auch außerhalb der Maskierungsflanken des nicht verzerrten Stimulus liegen können. Entsprechend wird einem derart verzerrten Stimulus im Paarvergleich eine höhere Schärfe zugeordnet werden.

3 Stand der Technik

Aktuelle Methoden zur Erfassung und Simulation des Verhaltens elektrodynamischer Lautsprecher im Klein- und Großsignalbereich werden in diesem Abschnitt vorgestellt. Weiterhin werden verschiedene Methoden zur Hörbarmachung von Lautsprecherverzerrungen erläutert und ihre Eignung für die Untersuchung der besprochenen Nichtlinearitäten herausgestellt. Zuletzt werden objektive Methoden zur Bewertung des Einflusses auf die Klangqualität und zur Erfassung der Hörbarkeit von Verzerrungen erläutert. Die ausgewählten Methoden werden vorgestellt, da sie in dieser Arbeit entweder angewendet oder evaluiert und optimiert werden.

3.1 Modellierung und Messung des Übertragungsverhaltens

Um das Übertragungsverhalten elektrodynamischer Lautsprecher modellieren zu können müssen lineare und nichtlineare Parameter in Teilsysteme zusammengefasst werden. Diese Teilsysteme erzeugen jeweils die akustischen Teilsignale $p_{lin.dis}(t)$, $p_{nlin.dis}(t)$ und $p_{def}(t)$. Die Addition aller Teilsignale erzeugt das akustische Gesamtsignal p(t), welches bei der Lautsprecherwiedergabe eines Stimulus vom Gehör verarbeitet oder in einem Messsystem von einem Messmikrofon aufgezeichnet wird. Abbildung 7 veranschaulicht diesen Zusammenhang:



Abbildung 7: Modellierung des elektrodynamischen Lautsprechers (vgl. Klippel GmbH, 2014b, S. 2)

Das lineare Modell stellt dabei das Verhalten des Lautsprechers im Kleinsignalbereich dar. Das Modell für reguläre Nichtlinearitäten erzeugt die in Abschnitt 2.3.2 besprochenen Anteile im Großsignalbereich. Irreguläre Nichtlinearitäten wie in 2.3.4 lassen sich durch ihre stochastische Charakteristik nur eingeschränkt simulieren. Je nach akustischer Umgebung kann ein signifikanter Einfluss durch akustische Störquellen gegeben sein, der in Abbildung 7 unter "Noise" mit dem zugehörigen Schalldrucksignal $p_{noise}(t)$ zusammengefasst wurde.

Folgend wird der prinzipielle Zusammenhang der Anpassung der einzelnen Modelle in einem Messsystem dargelegt. Auf eine tiefergehende Darstellung der einzelnen Mechanismen als in Abschnitt 2.3 wird verzichtet, da sie für das Verständnis der folgenden Kapitel nicht nötig ist und bei weiterem Interesse zum Beispiel in der Diplomarbeit von Schulze (vgl. 2014) gefunden werden kann.

3.1.1 Adaptives Fitting im Zeitbereich

Ein gängiger Ansatz zur Anpassung des Modells an einen Prüfling bzw. Device under Test (DuT) ist das Referenzieren eines Fehlersignals, welches durch Subtraktion eines gemessen

Signals vom modellierten Signal entsteht. Abbildung 8 veranschaulicht die Methode in vereinfachter Form:



Abbildung 8: Fitting des Modells, vereinfacht (vgl. Klippel, 2016, S. 360)

Ziel ist es das Fehlersignal e(t) durch Anpassung der Parameter des linearen Modells zu minimieren. Das Modell ist daher ein adaptives Filter (vgl. Klippel, 2016, S. 360). Das Ausgangssignal p(t) ist im Idealfall die Summe der Teilsignale $p_{def}(t)$, welches dem Fehlersignal entspricht, und $p_{lin,dis}(t)$, welches das linear verzerrte Signal darstellt.

Eine vollständige Separation regulärer linearer und nichtlinearer Verzerrungen ist nicht ohne weiteres möglich. Im Residuum des Fehlersignals können Produkte nichtlinearer Verzerrungen enthalten sein, welche nicht klar von irregulären Defekten zu trennen sind. Zudem erfolgt die Adaption nur langsam. (vgl. ebd. Gleichwohl besitzt das Verfahren für Produktionstests Gültigkeit: Wenn nur defekte Treiber erkannt werden müssen, ist die nicht vollständige Separierung von regulären linearen und nichtlinearen Verzerrungen akzeptabel, denn durch die verbleibenden Defekte im Fehlersignal lassen sich defekte Treiber von funktionierenden unterscheiden (vgl. Klippel, 2003, S. 1)

Um eine Trennung von linearen und regulären nichtlinearen Komponenten zu ermöglichen ist daher eine separate Identifikation der linearen Parameter nötig.

3.1.2 Offline-Identifikation zur Separierung von Verzerrungen

Bei der Offline-Identifikation wird zunächst eine Messung im Kleinsignalbereich durchgeführt. Aus den an den Polklemmen gemessenen Zeitverläufen des Stroms i(t) und der Spannung U(t) können absolute elektrische Daten gewonnen werden, die gemeinsam mit einer Messung der Auslenkung x(t) absolute mechanische Daten berechnen lassen (vgl. Klippel, 2000, S. 5). Damit kann das lineare Modell optimiert werden um den Anteil der linearen Verzerrungen von den nichtlinearen Verzerrungen in einer nachfolgenden Messung im Großsignalbereich separieren zu können. Das Schema wird in Abbildung 9 vereinfacht dargestellt:



Abbildung 9: Prinzip der Offline-Identifikation, vereinfacht (vgl. Klippel, 2016, S. 360)

Durch die Subtraktion des modellierten Ausgangssignals des linearen Modells $p_{lin.dis}(t)$ vom gemessen Signal p(t) lassen sich die nichtlinearen Verzerrungen $p_{nlin.dis}(t)$ insgesamt separieren. Das lineare Modell entspricht hierbei nicht einem simplen adaptiven Filter, sondern ist ein digital implementiertes Abbild des elektromechanischen Modells aus Abschnitt 2.3 (vgl. Klippel, 2000, S. 5).

Der Anwendungsbereich ist auch hier eingeschränkt, da sich die Auswirkungen einzelner Nichtlinearitäten nicht voneinander trennen lassen und zeitvariante Prozesse wie thermische Änderungen vernachlässigt werden (vgl. Klippel, 2016, S. 360).

3.1.3 Online-Identifikation zur Separierung nichtlinearer Verzerrungskomponenten

Zur Separierung verschiedener Komponenten der nichtlinearen Verzerrungen wird das Verfahren der Online-Identifikation genutzt. Zunächst wird das DuT im Kleinsignalbereich vermessen und die Parameter des linearen Systems werden ermittelt. Im weiteren Verlauf der Messung wird die Auslenkung zunehmend gesteigert. Dabei werden die Parameter der einzelnen Teilsysteme der nichtlinearen Komponenten geschätzt, die ihrerseits auch auf das lineare System rückwirken können. Thermische Änderungen werden kontinuierlich erfasst und die entsprechenden Rückkopplungen modelliert. Abbildung 10 verdeutlicht das Prinzip:



Abbildung 10: Prinzip der Online-Identifikation (vgl. Klippel, 2016, S. 361)

Durch die Zerlegung des nichtlinearen Systems aus Abschnitt 2.3.2 in separate Teilsysteme ist es möglich den Anteil jeder Verzerrungskomponente im Ausgangssignal des Lautsprechers zu schätzen. Demnach setzt sich das akustische Ausgangsignal p(t) als Summe der Teilsignale der nichtlinearen Systeme $p_{dis}(t)$ und dem linear verzerrten Signal $p_{lin}(t)$ nach Gl. 5 zusammen.

$$p(t) = p_{lin}(t) + p_{dis}(t) = p_{lin}(t) + p_{Bl}(t) + p_{K_{ms}}(t) + p_{L_e}(t)$$
Gl. 5

Das Verfahren besitzt Gültigkeit für Klein- und Großsignalbereich und kann darüber hinaus zeitvariante Einflüsse der Erwärmung von Magnet und Schwingspule abbilden. Weiterhin lassen sich die Einflüsse der wichtigsten Nichtlinearitäten qualitativ erfassen. Im Gegensatz zum adaptiven Fitting und der Offline-Identifikation kann das Verfahren nicht zur zielsicheren Detektion von irregulären Defekten angewendet werden. (vgl. Klippel, 2016, S. 361)

3.2 Hörbarmachung

Verfahren zur Hörbarmachung akustischer Signale oder Umgebungen werden unter dem Begriff Auralisation zusammengefasst. Für elektroakustische Systeme wird die Auralisation

zumeist genutzt um in konventionellen Hörversuchen oder mit perzeptiven Rechenmodellen Einflüsse von Verzerrungen zu evaluieren.

In dieser Arbeit werden zwei Verfahren zur Hörbarmachung regulärer und irregulärer nichtlinearer Verzerrungen von elektrodynamischen Lautsprechern genutzt und nachfolgend kurz vorgestellt. Beiden ist gemein, dass der zu evaluierende Anteil der nichtlinearen Verzerrungen für nachfolgende Prozesse skaliert werden kann. Somit kann in einem Hörversuch der Einfluss der Pegeländerungen der nichtlinearen Verzerrungen auf die Wiedergabequalität untersucht werden.

3.2.1 Auralisation durch Simulation der Zustandsvariablen

Die Modellierung des elektrodynamischen Lautsprechers im Großsignalbereich nach Abschnitt 2.3.2 kann mit der Zerlegung des nichtlinearen Modells in nichtlineare Teilsysteme wie in Abbildung 10 genutzt werden um beliebige Stimuli zu auralisieren. Die Struktur des Lautsprechermodells entspricht dabei im Wesentlichen dem zur Online-Identifikation nichtlinearer Parameter genutztem Modell. Der Unterschied besteht hierbei darin, dass die Ausgangssignale der nichtlinearen Teilsysteme $p_{L_e}(t)$, $p_{K_{ms}}(t)$ und $p_{Bl}(t)$ mit einem jeweils eigenen Skalierungsfaktor S_{L_e} , $S_{K_{ms}}$, S_{Bl} zum ebenfalls mit S_{lin} skalierbaren Ausgangssignal des linearen Systems $p_{lin}(t)$ addiert werden können um das auralisierte Signal $p_{aur}(t)$ zu erzeugen. Dieses Prinzip entspricht einem Mischer mit jeweils eigenen Kanal und Pegelsteller für jede Verzerrungskomponente. Abbildung 11 veranschaulicht den Signalfluss:



Abbildung 11: Prinzip der Auralisation durch Simulation (vgl. Klippel, 2001, S. 2)

Das auralisierte Ausgangssignal ergibt sich somit nach Gl. 6:

$$p_{aur}(t) = p_{lin}(t) \cdot S_{lin} + p_{L_e}(t) \cdot S_{L_e} + p_{K_{ms}}(t) \cdot S_{K_{ms}} + p_{Bl}(t) \cdot S_{Bl}$$
Gl. 6

Es ist entscheidend, dass die Skalierungsfaktoren nur die Erzeugung des auralisierten Signals beeinflussen. Die Struktur des Modells erfordert es, dass die Rückkopplung der Teilsysteme unabhängig von der Skalierung der Ausgangssignale der Teilsystem erfolgt um das Verhalten des Modell nicht zu verändern (vgl. Schulze, 2014, S. 17).

3.2.2 Differenzauralisation

Ein alternativer Ansatz zur dargestellten Nutzung eines Lautsprechermodells ist die Differenzauralisation. Dabei wird das Signal der nichtlinearen Verzerrungen aus der Differenzenbildung mit einer Referenz- und einer Testaufnahme eines oder mehrerer DuTs gewonnen. Abbildung 12 verdeutlicht das Prinzip:



Abbildung 12: Prinzip der Differenzauralisation (vgl. Klippel GmbH, 2014b, S.7)

Um das Verzerrungssignal d_{dis} zu erhalten wird ein Referenzsignal x_R von einem Testsignal x_T abgezogen. Die beiden Signale werden zuvor von einem möglichen Zeitversatz τ befreit und mit dem Verstärkungsfaktor G_R bei Pegelunterschieden angeglichen. Das auralisierte Signal y_A ist gemäß Gl. 7 dann die Addition des mit dem Faktor S_{dis} skalierten Differenzsignals d_{dis} zum skalierten Referenzsignal:

$$y_A = x_R \cdot G_R + d_{dis} \cdot S_{dis}$$
 Gl. 7

Je nach Referenzsystem lassen sich dabei verschiedene Verzerrungsarten auralisieren: Handelt es sich bei Referenz- und Testobjekt um den selben Lautsprecher, kann das Verzerrungssignal alle nichtlinearen Verzerrungen und evtl. Defekte dieses DuTs enthalten. Dies gilt für den Fall, dass die Referenzaufnahme im Kleinsignalbereich und die Testaufnahme im Großsignalbereich durchgeführt wird. Sollen hingegen nur irreguläre Defekte auralisiert werden, so kann z.B. ein Treiber mit und ohne lose Partikel auf der Membran aufgezeichnet werden. (vgl. Klippel, 2016, S. 464 ff.)

3.3 Klippel R&D System

Das Klippel R&D System bietet Lösungen für Aufgaben und Probleme im Umfeld der Entwicklung und Optimierung von elektroakustischen Anlagen und Treibern von der Komponentenebene bis zur Systemebene. Es besteht aus verschiedenen Hardwarekomponenten (Distortion Messmikrofone. Analyzer, Triangulationslaser. Spannungssensoren, ...) welche die Messung relevanter Parameter mit der Software dB-Lab ermöglichen. dB-Lab verwaltet die gemessenen Daten in Datenbanken und stellt Optionen für Ansicht und Postprocessing der Ergebnisse bereit. Die Software ist modular aufgebaut und mit entsprechenden Modulen lassen sich verschiedene Aufgaben bewältigen. Die simpelste Struktur innerhalb einer Datenbank ist ein Treiberobjekt, dem beliebig viele Operationen zugeordnet werden können und in Abbildung 13 veranschaulicht:



Abbildung 13: simple Datenbankstruktur

Dabei wird jeder Operation ein Modul zugeordnet. Das *Automation Interface* von dB-Lab ermöglicht die Verknüpfung von Modulen sowie das automatisierte Weiterleiten von Daten zwischen Operationen.

Im Folgenden werden die in dieser Arbeit verwendeten Module kurz vorgestellt.

3.3.1 Linear Parameter Measurement (LPM)

Zur Erfassung der linearen Parameter eines Lautsprechers wird das Modul *Linear Parameter Measurement* genutzt. Mittels einer Messung der an den Polklemmen anliegenden Spannung und des fließenden Stromes im Kleinsignalbereich werden die elektrischen Parameter eines Treibers ermittelt. Ein lineares Modell wird dabei auf den gemessenen Impedanzgang des DuTs gefittet. Durch Nutzung eines zusätzlichen Triangulationslasers kann die Membranauslenkung abgetastet werden, wodurch die Bestimmung der mechanischen Parameter ermöglicht wird. Darüber hinaus kann das Messsystem mit einem Messmikrofon erweitert werden, um das Schalldrucksignal simultan zu erfassen. Zur Durchführung der Messung wird mindestens ein Distortion Analyzer 1, ein Verstärker und ein Computer benötigt. (vgl. Klippel GmbH, 2012, S. 2 ff.)

Das Modul wurde in dieser Arbeit benutzt um die linearen Parameter von verschiedenen Treibern zu vermessen.

3.3.2 Large Signal Identification (LSI)

Das *Large Signal Identification* Modul ermöglicht die Messung elektrischer, mechanischer und thermischer Parameter für verschiedene Treibertypen in verschiedenen Einbausituationen in einer Echtzeitmessung. Es ermöglicht das Identifizieren und Quantifizieren nichtlinearerer Verzerrungen und ihrer Ursachen und bietet dabei thermischen und mechanischen Überlastschutz für das DuT. Die Messung startet im Kleinsignalbereich und erhöht schrittweise die Verstärkung bis zum Erreichen der Überlastschutzparameter im Großsignalbereich. Das Modul erfordert zur Durchführung einer Messung neben einem Computer zumindest einen Distortion Analyzer 2 und einen Verstärker. Es kann bei Bedarf mit einem Triangulationslaser erweitert werden. (vgl. Klippel GmbH, 2015b. S 1 ff.)

Das Prinzip nutzt das in Abschnitt 3.1.3 beschriebene Verfahren der Online-Identifikation. In dieser Arbeit wurde das LSI-Modul zur Bestimmung der nichtlinearen Kennlinien verschiedener Treiber genutzt.

3.3.3 Simulation (SIM2)

Das *Simulation* Modul in der Version 2 ermöglicht die numerische Simulation elektrodynamischer Treiber über den kompletten Arbeitsbereich unter Berücksichtigung dominanter Nichtlinearitäten. Der zur Simulation genutzte Datensatz kann entweder aus einer Messung mit den LPM- und LSI-Modulen stammen oder synthetisch generiert worden sein. Da sich die einzelnen Nichtlinearitäten an- und abschalten lassen eignet sich das Modul zur gezielten Optimierung des Verhaltens im Großsignalbereich. Das verwendete Lautsprechermodell berücksichtigt zusätzlich zu den in Kapitel 2.3 beschriebenen Größen auch die Einbausituation des DuTs in konventionelle Gehäusetypen. Es handelt sich um ein reines Softwaremodul. (vgl. Klippel GmbH, 2013b, S. 1 ff.)

Innerhalb des Moduls werden weiterhin thermische Änderungen für zwei Zustände berücksichtigt: Der Kurzzeitmodus simuliert den thermisch gesättigten Zustand der
Schwingspule, während der Langzeitmodus überdies die Komponenten der Magnetanordnung als thermisch gesättigt annimmt (vgl. Schulze, 2014, S. 20).

Das Modul wurde in dieser Arbeit genutzt um die mit LPM und LSI gewonnen Daten zu untersuchen und diese ggf. für den Forschungszweck anzupassen. Ein Schwerpunkt liegt dabei auf der Untersuchung der DuTs hinsichtlich auftretender Intermodulationsverzerrungen und der Eignung für die Auralisation breitbandiger Stimuli.

3.3.4 Simulation / Auralization (SIM-AUR)

Zur Auralisation beliebiger Stimuli mit einem präzisen Modell des elektrodynamischen Lautsprechers wurde das *Simulation/Auralization* Modul entwickelt. Zum Zeitpunkt der vorliegenden Arbeit ist es angekündigt jedoch noch nicht veröffentlicht worden. Folglich wird eine Entwicklungsversion verwendet.

Ähnlich dem SIM2-Modul simuliert das SIM-AUR-Modul numerisch die elektrischen, mechanischen und thermischen Zustandsvariablen eines Lautsprechers unter Berücksichtigung gängiger Einbausituationen. Darüber hinaus verarbeitet es beliebige Stimuli und ist nicht auf einfache Testsignale beschränkt. Auch die Zeitverläufe der simulierten Größen werden ausgegeben. Mit der Simulation kann der Stimulus monaural auralisiert und abgehört werden. Weiterhin steht eine Exportfunktion für die Auralisation zur Verfügung, welche das auralisierte Signal mit variabel skalierten nichtlinearen Verzerrungen zur Weiternutzung bereitstellt. Die Nutzung des Modusl erfordert keine weitere Hardware. (vgl. Klippel GmbH, 2016, S. 1 f.)

Auch wenn das Modul dem Auralisationsschema aus Abbildung 11 folgt, besteht im derzeitigen Entwicklungsstand eine Einschränkung: Die Pegelsteller für die einzelnen Nichtlinearitäten können momentan nicht beeinflusst werden. Derzeit müssen die nichtlinearen Teilsysteme einzeln zu- oder abgeschaltet werden um in der Auralisation nur die gewünschten nichtlinearen Verzerrungskomponenten zu beinhalten. Die Summe der nichtlinearen Verzerrungen lässt sich insgesamt für die Wiedergabe mit dem Verstärkungsfaktor S_{dis} skalieren.

Dieses Verfahren beeinträchtigt die bereits herausgestellte Rückkopplung der Systeme, wodurch beim Abschalten einer Nichtlinearität das Gesamtverhalten des Systems beeinflusst wird. Da der Fokus dieser Arbeit jedoch vorrangig auf dem Einfluss einzelner Nichtlinearitäten auf die empfundene Wiedergabequalität lag, ist diesem Umstand keine negative Bedeutung beizumessen.

Das Modul wurde für diese Arbeit benutzt um beliebige Stimuli mit einer Auswahl an DuTs zu auralisieren.

3.3.5 Difference Auralization (DIF-AUR)

Das *Difference Auralization* Modul ist eine Implementierung der in Abschnitt 3.2.2 vorgestellten Auralisationstechnik.

Das Modul kompensiert den Zeitversatz τ zwischen Referenz- und Testsignal korrelationsbasiert. Das Differenzsignal kann separiert werden und für die Wiedergabe oder den Export skaliert dem Referenzsignal beigemischt werden. (vgl. Klippel GmbH, 2014c, S. 1)

Für diese Arbeit wurde das Modul verwendet um Rub&Buzz-Verzerrungen für Hörversuche zu separieren und entsprechende Audiodateien bereitzustellen.

3.4 Perzeptive Evaluationsmethoden

Das Ziel perzeptiver Evaluationsmethoden ist das Erfassen quantitativer Werte hinsichtlich subjektiver Wahrnehmungen und deren Änderungen. Im Kontext von gehörbezogenen Sinneseindrücken besteht diese Abhängigkeit zumeist zu Verzerrungen eines Schalls. Liebig definiert die Verzerrung als Differenz zwischen einem Referenz- und einem Testsignal bzw. als Signalanteile die nur im Testsignal vorhanden sind (vgl. 2012, S. 23).

In perzeptiven Rechenmodellen werden wahrnehmungsbezogene Größen mit psychoakustischen Modellen ermittelt. Die berechneten Werte werden als *Model Output Variables* (im Folgenden: MOV) bezeichnet. MOVs sind zumeist als Verständnis- und Erklärungshilfen für die Durchführung und Auswertung von Hörversuchen zu nutzen. Je nach technischer Anwendung werden spezifische MOVs ermittelt. Folgend werden kurz die für diese Arbeit wichtigsten Methoden und ihr Zusammenhang erläutert. Dabei wird sich auf die zum Verständnis notwendigen Bestandteile beschränkt.

3.4.1 Perceptual Evaluation of Audio Quality (PEAQ)

Die Empfehlung BS.1387-1 der *International Telecommunication Union* wurde erstmalig 1998 veröffentlicht und stellt eine Messmethode zur Erfassung der perzeptiven Klangqualität eines DuTs vor (vgl. ITU, 2001, S. 1). Der Schwerpunkt der Entwicklung liegt auf einer gehörgerechten Bewertung von Signalverzerrungen, welche durch Codierverfahren für die Distribution von digitalen Audiodaten verursacht werden (vgl. Thiede et. al., 1998, S. 1f.).

Das Grundprinzip des Verfahrens ist ein Vergleich zwischen einem Referenz- und einem Testsignal mit einer objektiven Messmethode. Dabei entspricht das Testsignal dem durch das DuT verarbeitete Referenzsignal und enthält somit die durch das DuT verursachten Differenzen. (vgl. ITU, 2001, S. 8).

Drei Strategien werden in der objektiven Messmethode genutzt um die Klangqualität zu bewerten: Das *Masked Threshold Concept* vergleicht das Differenzsignal zwischen Referenzund Testsignal mit einer aus dem Referenzsignal berechneten Maskierungsschwelle. Das Konzept der *Comparison of Internal Representations* bildet kein Differenzsignal sondern vergleicht gehörgerechte Darstellungen der Eingangssignale. Beide Konzepte erfordern die Nutzung eines *peripheren Gehörmodells* um die jeweiligen Erregungsmuster berechnen zu können. Die dritte Strategie basiert auf dem Vergleich von FFT-Spektren und wird als Ergänzung zu Gehörmodellen betrachtet. (vgl. Thiede et. al., 1998, S. 3 f.)

Die vorgeschlagene Methodik kann anwendungsbezogen in einer Basisversion oder einer höher auflösenden, erweiterten Version genutzt werden (vgl. ITU, 2001, S. 11). Die vorgeschlagenen Methoden wurden u.a. von Liebig (vgl. 2012) für die Anwendung an Lautsprechern implementiert und der aktuelle Entwicklungsstand ist Kern dieser Arbeit. Im folgenden Abschnitt wird daher die Implementierung zum Zeitpunkt der Übernahme durch den Autor dargestellt.

3.4.2 Perceptual Evaluation of Quality (PEQ)

Das *Perceptual Evaluation of Quality* Modul ist derzeit im Entwicklungsstatus und soll als Werkzeug innerhalb des Klippel R&D-Systems veröffentlicht werden. Das Ziel ist es für die Entwicklung von Lautsprechern eine gehörgerechte Evaluation der Klangqualität eines Treibers auf Basis von Rechenmodellen zu ermöglichen. Der erzeugte Datensatz aus MOVs soll dabei sowohl für Entwicklungsingenieure als auch für Entscheider in Vertriebsabteilungen verständlich und informativ sein. Die Entwicklung wurde 2012 mit dem *Perceptual Assessment Module* in der Masterarbeit von Liebig begonnen (vgl. 2012).

Mögliche Anwendungsszenarien sind die Nutzung digitaler Prototypen aus dem SIM2-Modul, die mit dem SIM-AUR-Modul auralisiert wurden, oder die Verwendung von Signalen, die mit echten Treibern angefertigt wurden. Für die letztgenannte Anwendung verfügt das Modul über eine einfache Ausgabe auralisierter Signale, die neben der Bewertung durch die MOVs auch eine subjektive Bewertung des DuTs ermöglichen sollen.

Überblick

Dem PEQ-Modul wird ein Referenzsignal x_R und ein Testsignal x_T innerhalb von dB-Lab entweder als Audiodateien oder als Vektordaten übergeben. Abbildung 14 verbildlicht die Struktur des Moduls, die Kernelemente werden nachfolgend erläutert:



Abbildung 14: Struktur des PEQ-Moduls

Eingangsblock

Innerhalb des Eingangsblocks werden Signale für die weitere Verarbeitung vorbereitet. Der *Reference Gain* G_R ist ein variabler Verstärkungsfaktor und wird exklusiv auf das Referenzsignal appliziert. Dies ist insbesondere dann anzuwenden, wenn die Signale vom gleichen DuT stammen und mit unterschiedlicher Verstärkung angefertigt wurden. Die angewendete Anpassung der Verstärkung ist dann vom Benutzer dem Modul zu übergeben.

Es folgt die Prüfung auf Zeitsynchronizität der Signale. Dieser Prozess sucht innerhalb eines variablen Zeitfensters nach dem Korrelationsmaximum beider Signal und entfernt einen evtl. vorhandenen Zeitversatz τ . Der Prozess findet zweistufig statt: Eine grobe Bestimmung des Zeitversatz erfolgt zunächst im Zeitbereich und wird von einer feineren Prüfung im Frequenzbereich der Signale gefolgt. Der ermittelte Versatz τ wird durch Verschieben der Signale zueinander kompensiert. Nicht überlappende Samples werden verworfen. (vgl. Liebig, 2012, S. 28 f.)

Die perzeptive Evaluation erfordert es, dass die Signale als Schalldrucksignale betrachtet werden. Um die Zeitsignale in die Schalldruckdomäne zu wandeln gilt die Konvention dass der Wert des Zeitsignals in *Pascal* zu verstehen ist. Für die Konvertierung wird der RMS-Wert des Zeitsignals nach Gl. 8 bestimmt:

$$\tilde{y} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} y^2(n)}$$
Gl. 8

Der zugehörige Schalldruckpegel für ein Zeitsignal mit $\tilde{y} = 1$ wird auf den Normschalldruck bezogen und ergibt sich nach Gl. 9 zu:

$$L_{y,RMS} = 20 \cdot \log_{10} \left(\frac{\tilde{y}}{2 \cdot 10^{-5} \operatorname{Pa}} \right) dB = 20 \cdot \log_{10} \left(\frac{1}{2 \cdot 10^{-5} \operatorname{Pa}} \right) \approx 94 \ dB \ SPL$$
Gl. 9

Um verschiedene Schalldruckpegel für die gehörgerechten Modelle der perzeptiven Evaluation zu erzeugen, wird ausgehend von dem Schalldruckpegel des Referenzsignals und einer Zielvorgabe der *Playback Gain G_P* berechnet und auf beide Signale appliziert.

Weitherhin ist es notwendig, dass die synchronisierten Signale auf möglichst identische Pegel skaliert werden. Dafür ist der *Equalization Gain* G_L als Skalierer vorgesehen. Die Berechnung des Skalierers kann in zwei unterschiedlichen Modi erfolgen: Der *Level Mode* zieht zur Berechnung den RMS-Wert der Zeitsignale heran. Der *Loudness Mode* berechnet den Skalierer G_L in abhängig von der Lautheit *N* der Signale. Das Verfahren wurde von Zeller in Anlehnung an Zwicker ausführlich beschrieben (vgl. 2012, S. 194 ff.). Der Equalization Gain wird exklusiv auf das Testsignal angewendet.

Audioexport

Der Audioexport ist ein Block, der dem DIF-AUR-Modul entlehnt ist. Die Signalvorbereitung im Eingangsblock ist zwischen beiden Modulen sehr ähnlich gestaltet. Während das DIF-AUR Modul zur Auralisation der Verzerrung mit einem variablen Verstärkungsfaktor S_{dis} entwickelt wurde, soll der Audioexport des PEQ-Moduls die kalibrierbare Einrichtung eines Hörversuchs ermöglichen. Abbildung 15 zeigt den Signalfluss vom Eingangsblock über den Signalexport zum Hörversuch:



Abbildung 15: Verknüpfung zwischen Eingangsblock, Audioexport und Hörversuch

Der *Export Gain* G_E ist ein weiterer Skalierer, der die Aussteuerungsreserve für den Export ins digitale PCM-Format sicherstellt (vgl. Liebig, 2012, S. 33).

Die Kalibrierung eines Hörversuchs erfolgt mittels des erzeugten Kalibriersignals y_{cal} dessen Zielschalldruckpegel $L_{p,cal}$ vom Benutzer festgesetzt wird. Es handelt sich um ein oktavbreites

Rauschen mit einer Mittenfrequenz von 1 kHz. Bei der Wiedergabe dieses Kalibriersignals wird die Verstärkung des Wiedergabesystems so lang angepasst und durch Messung überprüft bis der Zielschalldruckpegel am Hörplatz erreicht ist. Somit ist sichergestellt, dass die Wiedergabe des Referenz- und Testsignals den identischen Schalldruckpegel erzeugt, der auch dem perzeptiven Evaluationsblock übergeben wird. (vgl. Liebig, 2012, S. 33)

Perzeptive Evaluation

Wie Abbildung 14 zeigt werden innerhalb des perzeptiven Evaluationsblocks mehrere Rechenmodelle angewendet. Die drei Verfahren werden nachfolgend verkürzt beschrieben. Detaillierte Beschreibungen sind den jeweils referenzierten Veröffentlichungen zu entnehmen.

FFT-Modell

Das FFT-Modell ist eine Implementierung der Basisversion des PEAQ Algorithmus und nutzt zur Transformation der Zeitsignale in Erregungsmuster die *Fast Fourier Transformation* und mehrere Wichtungsfunktionen. Die genauen Teilschritte des Verfahrens sind in Abschnitt 2 der Empfehlung der ITU beschrieben (vgl. ITU, 2001, S. 32 ff). Abbildung 16 verdeutlicht den Ablauf:



Abbildung 16: Überblick FFT-Modell

Die Aufgabe des Blocks ist die Transformation der pegelkorrigierten und synchronisierten Zeitsignale y_R und y_T in psychoakustische Erregungsmuster um die gewünschten MOVs berechnen zu können.

Die FFT wird blockweise mit einer Länge von 2048 Samples und einer Überlappung von 50% bzw. 1024 Samples durchgeführt; die Blöcke werden dabei mit einem Hanning-Fenster gewichtet (vgl. ebd., S. 35) und die Fensterverluste werden kompensiert. Während das PEAQ-Verfahren diese Werte für eine Abtastrate von 48 kHz spezifiziert, werden diese innerhalb des PEQ-Moduls auf alle gängigen Abtastraten angewendet.

Die Betragsspektren werden nach der Transformation mit dem Skalierer G_P auf den Zielschalldruckpegel gebracht und das Differenzspektrum Y_{diff} gebildet. Im Falle des Testsignals wird zusätzlich noch der Skalierer G_L appliziert.

Anschließend werden die Spektren Y_R und Y_T mit einer Wichtungsfunktion zur Simulation der Transferfunktion des Außen- und Mittelohres gewichtet. Aus den gewichteten Spektren $Y_{R,W}$ und $Y_{T,W}$ wird das Fehlersignal bzw. *Distortion Pattern DP* gebildet. Die gewichteten Spektren und das Fehlersignal werden in die Domäne der Frequenzgruppen überführt; das Fehlersignal wird danach nicht weiter transformiert. Die Breite einer Frequenzgruppe beträgt 0,25 Bark. (vgl. ITU, 2001, S. 34 & S. 36)

Die Frequenzgruppenspektren für Referenz- und Testsignal werden mit einem frequenzabhängigen Versatz versehen um den internen Rauschteppich und die Ruhehörschwelle des Gehörs zu simulieren (vgl. Liebig, 2012, S. 16 & vgl. ITU, 2001, S. 43).

Zur Bildung der Erregungsmuster werden die Frequenzgruppenspektren zunächst im Frequenzbereich verschmiert. Diese Stufe soll die Anregung benachbarter Frequenzgruppen simulieren und bildet den aus Abschnitt 2.1.2 bekannten Effekt der Simultanverdeckung nach (vgl. Liebig, 2012, S. 16). Temporale Verdeckungseffekte werden zuletzt durch eine Verschmierung im Zeitbereich simuliert. Dies wird mit einem Tiefpassfilter erster Ordnung erzielt, der abhängig von der Mittenfrenquenz der jeweiligen Frequenzgruppe berechnet wird (vgl. ITU, 2001, S. 45). Die Bildung der Erregungsmuster wurde von Liebig umfassend evaluiert (vgl. 2012, S. 44 ff.).

Sowohl die rein frequenzverschmierten als auch die frequenz- und zeitverschmierten Erregnungsmuster werden gemeinsam mit dem Fehlersignal zur Berechnung der MOVs benutzt.

Filterbankmodell

Das Filterbankmodell des PEQ-Moduls ist eine Version des erweiterten PEAQ-Algorithmus und wurde 2013 von Straube implementiert und für die gegebene Anwendung verifiziert (vgl. 2013). Die Notwendigkeit der Nutzung der vergleichbar komplexen Filterbank ist dadurch gegeben, dass sich bestimmte Empfindungsgrößen mit den Kompromissen des FFT-Modells hinsichtlich der zeitlichen Auflösung nicht ausreichend genau berechnen lassen (vgl. ebd., S. 23). Abbildung 17 stellt das Verfahren auf der nächsten Seite dar.



Abbildung 17: Überblick Filterbankmodell (vgl. Straube, 2013, S. 25)

Die Skalierung der Eingangssignale erfolgt wiederum abhängig von den Faktoren G_P und G_L . Zur Entfernung etwaiger Gleichanteile und tieffrequenter Störungen wurde ein Butterworth-Filter vierter Ordnung als Hochpass mit der Grenzfrequenz $f_g = 20$ Hz implementiert (vgl. Straube, 2013, S. 27).

Anschließend werden die Eingangssignale von der Filterbank in je 40 Bandpasssignale zerlegt und somit auf die Frequenzgruppen transformiert (vgl. ITU, 2001, S. 48 ff.). Entgegen der Angabe des PEAQ-Verfahrens beträgt die Breite einer jeden Frequenzgruppe in der vorliegenden Implementierung nicht 0,5 Bark, sondern 0,707 Bark (vgl. Straube, 2013, S.28). Die Ausgänge der Filter werden um den Faktor 32 unterabgestastet (vgl. ITU, 2001, S. 50). Für Eingangssignale mit einer Abtastrate $f_s = 48000$ kHz haben die Filterausgänge demach eine Abtastrate von $f_{s,red\#1} = 1500$ Hz.

Die Außen- und Mittelohrwichtung erfolgt wiederum abhängig von der Mittenfrequenz des jeweiligen Filterbankkanals (vgl. ebd.).

Die Frequenzverschmierung des FFT-Modells wird im Filterbankmodell durch eine Korrektur der Filterflanken jedes Kanals erzielt, die sich abhängig von der Mittenfrequenz des Bandpassfilters und dem Pegel des jeweils gefilterten Signals ändert (vgl. Straube, 2013, S. 47). Straube bezeichnet dies bewusst differenziert als Filterflankenkorrektur, um die Unterschiede der verwendeten Verfahren im FFT- und Filterbankmodell kenntlich zu machen. Diese Stufe modelliert Effekte der Simultanverdeckung.

Die Hüllkurven können basierend auf der verwendeten Filterstruktur durch Adddition der quadrierten Real- und Imaginärteile eines jeden komplexen Signals je Filterbankkanal gebildet werden (vgl. ebd., S. 52).

Die letzten Stufen des Modells simulieren die Vor- und Nachverdeckung und das interne Rauschen des Hörers bzw. die Ruhehörschwelle (vgl. ebd., S. 26). Die erste Verschmierung im Zeitbereich ist als Tiefpassfilter und Unterabtastung um den Faktor 6 realisiert und modelliert die Vorverdeckung (vgl. Straube, 2013, S. 52). Für das zuvor genannte Beispiel mit $f_s = 48000$ kHz mit haben die Signale je Filterbankkanal nun nur noch eine Abtatsrate von $f_{s,red\#2} = 250$ Hz. Der interne Rauschteppich wird durch Addition eines frequenzabhängigen Versatzes nachgestellt (vgl. ebd., S. 53). Die zweite zeitliche Vermschmierung ist ein weiteres Tiefpassfilter, ohne zusätzliche Unterabtastung, und modelliert die Nachverdeckung (vgl. ebd., S. 54).

In der bisherigen Implementierung im PEQ-Modul schließt sich an das Filterbankmodell keine diskrete Bildung der Erregungsmuster an. Da die Filterbank für weitere Anwendungen optimiert wurde, stehen dem Nutzer die Ausgangssignale der 40 Filterbankkanäle in Form der frequenz- und zeitverschmierten Hüllkurven zur Verfügung.

Die verbesserte zeitliche Auflösung der Ausgänge des Filterbankmodells eignet sich für die Berechnung komplexerer psychoakusticher Größen wie der Rauigkeit. Innerhalb des PEQ-Moduls schließt sich derzeit an das Filterbankmodell ein 2014 von Kunz implementiertes Modell zur Bestimmung der MOV Rauigkeit an (vgl. 2014).

Rauigkeitsmodell

Das Rauigkeitsmodell ist mit Teilstufen des Filterbankmodells verknüpft. Da die temporale Maskierung für die Empfindungsgröße Rauigkeit keine Rolle spielt, erfolgt der Abgriff der Signale direkt im Anschluss an die Hüllkurvenbildung des Filterbankmodells (vgl. Kunz, 2014, S. 27). Die zeitlichen Verschmierungen und die Addition des internen Rauschens werden entsprechend zur Berechnung der Rauigkeit nicht berücksichtigt. Abbildung 18 illustriert das Verfahren vereinfacht dargestellt für einen Satz Hüllkurven eines Eingangssignals:



Abbildung 18: Überblick Rauigkeitsmodell, vereinfacht (vgl. Kunz, 2014 S. 28)

Die Implementierung lehnt sich an das von Daniel & Weber vorgeschlagene Modell an (vgl. Kunz, 2014, S. 28); entsprechend ist als erste Stufe eine Neuaufteilung der Zeitsignale in N Blöcke b der Länge t = 200 ms notwendig (vgl. Daniel & Weber, 1997, S. 116). Die Zeitsignale der Hüllkurven für ein Eingangssignale mit $f_s = 48$ kHz sind bereits auf eine Abtastrate $f_{s,red\#1} = 1500$ Hz heruntergetastet. Je Block liegen somit für 40 Filterkanäle je 300 Samples des Zeitsignals der Hüllkurven vor. In Abbildung 18 erfolgt die Auswahl eines jeden Filterbankkanals innerhalb eines Blocks *b* mittels des Index i = 1...40. Die Auswahl des entsprechenden Blocks erfolgt über den Index n = 1...N; die zugehörige zeitliche Position jedes Blocks ist demnach das Produkt $n \cdot t$.

Zur Bestimmung der Modulationsfrequenz wird das Zeitsignal jedes Filterbankkanals mittels gefensterter FFT in den Frequenzbereich überführt. Das Fenster wird nach Blackman-Nuttall gebildet. Die resultierenden Spektren $B_{(n,i)}$ werden dann mit einer Wichtungsfunktion $H_{(i)}$ gewichtet um den Abstand zwischen Modulations- und Trägerfrequenz abzubilden. Die Gestalt der Wichtungsfunktion ist dabei abhängig vom jeweiligen Filterbankkanal *i*. Durch inverse FFT wird das gefilterte Zeitsignal $b^*_{(n,i)}$ erzeugt. (vgl. Kunz, 2014, S. 29 f)

Anschließend wird analog zu Gl. 8 der Effektivwert des gefilterten Signals $b^*_{(n,i)}$ für jeden Filterbankkanal berechnet. Gemeinsam mit dem Gleichanteil $DC_{(n,i)}$, welcher aus dem ungefilterten Zeitsignal berechnet wird, kann nach Gl. 10 der verallgemeinerte Modulationsindex $m_{(n,i)}$ berechnet werden:

$$m_{(n,i)} = g(zi) \frac{\tilde{b} *_{(n,i)}}{DC_{(n,i)}}$$
 Gl. 10

Dabei wird eine spektrale Gewichtung g(zi) appliziert, deren Zweck in der Abbildung des Verlaufs der Rauigkeitswahrnehmung ggü. dem Trägersignal besteht. Der verallgemeinerte Modulationsindex wird in diesem Kontext auch gebildet um den Einfluss transienter Signalverläufe auf die Rauigkeit zu begrenzen. (vgl. ebd., S. 30)

Die Unterscheidung zwischen rauschhaften Signalen und periodisch modulierten Signalen wird durch die Korrelation der gewichteten Spektren der Nachbarkanäle *i*+1 nd *i*-1 mit dem gefilterten Spektrum $b^{*}_{(n,i)}$ ermöglicht. Die zugehörigen Korrelationskoeffizienten $k_{(n,i-1)}$ und $k_{(n,i+1)}$ werden zur Berechnung der spezifischen Teilrauigkeit eines Filterbankkanals innerhalb des Blocks nach Gl. 11 benötigt:

$$r_{(n,i)} = \left(m_{(n,i)} \cdot k_{(n,i-1)} \cdot k_{(n,i-2)}\right)^2$$
Gl. 11

$$R_{(n)} = cal \cdot \sum_{i=1}^{40} r_{(n,i)}$$
 Gl. 12

Die Gesamtrauigkeit R(n) lässt sich in Gl. 12 sich durch Summierung der spezifischen Teilrauigkeiten innerhalb eines Blocks berechnen. Dabei wird ein Kalibrierfaktor *cal* benutzt, um das Modell auf die Bezugsrauigkeit nach Fastl & Zwicker zu kalibieren. Die Gesamtrauigkeit R eines Signals kann durch Mittelwertbildung über die Gesamtrauigkeiten $R_{(n)}$ der Blöcke eines Signals ermittelt werden. (vgl. zu den letzen beiden Absätzen: ebd., S. 31)

Neben der Rauigkeit, deren Berechnung zuvor dargelegt wurde, erzeugt das PEQ-Modul weitere MOVs als Ausgaben der Analyse. Diese werden nachfolgend zusammenfassend erläutert.

Distortion To Mask Ratio - DMR

Das MOV *Distortion To Mask Ratio* beschreibt den Abstand zwischen der Maskierungsschwelle M und dem Distortion Pattern DP in dB. Beide werden innerhalb des FFT-Modells gebildet. Die Größe lässt sich sowohl als Zeitverlauf für jeden Block als auch als Gesamtmaß für ein Signal bestimmen. Im letztgenannten Fall wird sie als *Total Distortion To Mask Ratio* (DMR_{tot})bezeichnet und ist der Mittelwert der DMR-Werte aller Blöcke. (vgl. Liebig, 2012, S. 37)

Lautheit und Lautstärkepegel

Das PEQ-Modul stellt die spezifischen Lautheitsverteilungen bzw. Lautstärkespegelverteilungen für Referenz-, Test- und Differenzsignal für einen gegebenen Zeitpunkt dar. Die Berechnung erfolgt auf Basis der frequenz- und zeitverschmierten Erregungsmuster. Der Nutzer kann dabei den Zeitpunkt wählen und die Art des Spektrums angeben. Die Lautheitsspektren werden als spezifische Lautheiten in der Einheit sone/Bark über der Barkskala dargestellt. Für die Darstellung der Lautstärke hat die Ordinate der Grafik die Einheit phon.

Weiterhin werden die Zeitverläufe der Lautheit bzw. Lautstärke wahlweise dargestellt. Dabei trägt die Ordinate entsprechend die Einheit sone oder phon. Die Berechnungen von MOVs mit Lautheitsbezug wurden von Liebig hinreichend evaluiert (vgl. 2012, S. 44).

Absolute Schärfe

Die absolute Schärfe *S* nach Fastl & Zwicker (vgl. Fastl & Zwicker, 2007, S. 241 ff.) wird für das Referenz- und Testsignal berechnet und der Zeitverlauf in der Einheit *acum* dargestellt. Die Berechnung erfolgt basierend auf den frequenz- und zeitverschmierten Erregungsmustern.

Discoloration, General Bass Enhancement, Treble Stressing

Klippel untersuchte 1990 den Zusammenhang zwischen Höreindruck und Lautsprecherparametern und leitete unter anderem die Größen *Discoloration V, General Bass Enhancement BE* und *Treble Stressing TS* ab um verschiedene Dimensionen des Höreindrucks zu beschreiben. Allen Größen ist gemein, dass zur Berechnung über die Differenz der spezifischen Lautheitsverteilungen über den Frequenzgruppen zwischen Referenz- und Testsignal integriert wird. Dabei werden für die Berechnung von *BE* und *TS* spezifische Wichtungsfunktionen genutzt. Die Angabe erfolgt in Prozent. (vgl. Klippel, 1990, S. 49).

Das Maß Discoloration wird der generellen Verfärbung zwischen Referenz- und Testsignal zugeordnet, während Treble Stressing die Dimension Schärfe und General Bass Enhancement die Dimension Lautstärke beschreibt (vgl. ebd.). Die Versuche wurden nicht mit dem Fokus auf den Einfluss nichtlinearer Verzerrungen durchgeführt, da als Referenz stets das unverzerrte Ursprungssignal genutzt wurde während das Testsignal die Wiedergabe durch den Lautsprecher repräsentiert (vgl. ebd., S. 48). Folglich ist die Bildung und Evaluation der Maße innerhalb der durchgeführten Hörversuche stark von den linearen Verzerrrungen der Testlautsprecher geprägt, sofern diese nicht nahe ihrer Aussteuerungsgrenze betrieben wurden.

3.5 Onlinehörversuch

Zu Beginn der vorliegenden Arbeit war bereits ein webbasierter Hörversuch zur Bestimmung der individuellen Wahrnehmungsschwelle verschiedener Verzerrungen für diverse Kombinationen von DuTs und Stimuli in Betrieb.

Das Prinzip entspricht dem *Two Alternatives Forced Choice* Verfahren: Dem Probanden werden zwei Signale präsentiert: eines enthält lineare und nichtlineare Verzerrungen (*"distorted"*) und das andere nur lineare Verzerrungen (*"undistorted"* bzw. *"linear"*). Der Proband kennt die Zuordnung der Signale nicht und muss auf Basis seiner Wahrnehmung das verzerrte Signal identifizieren. Eine Enthaltung ist nicht möglich. Bei einer richtigen Antwort wird für den nächsten Versuch der Pegel der nichtlinearen Verzerrungen reduziert. Dies erfolgt typischer Weise mit einer Schrittweite von 3 dB. Bei einer falschen Antwort wird der Anteil der nichtlinearen Verzerrungen um zwei Schritte erhöht, z.B. um 6 dB. Jede Umkehr der Richtung der Pegeländerung der nichtlinearen Verzerrungen wird festgehalten. Am dritten Umkehrpunkt wird der Median zwischen den Umkehrpunkten ermittelt. Dieser repräsentiert für diesen Test die Wahrnehmungsschwelle des Probanden. (vgl. Klippel GmbH, 2017)

Die Methode wurde von der *Weighted Up-Down Method* nach Kaernbach adaptiert. Sie stellt gegenüber ähnlichen Methoden eine einfachere Implementierung dar und kann gegen einen beliebigen Punkt der zu untersuchenden psychometrischen Funktion konvergieren. Darüber hinaus eignet sie sich für die Untersuchung von Wahrnehmungsschwellen, die eine Detektionsrate des Unterschiedes zweier Signale von 75% durch den Probanden erfordern. (vgl. Kaernbach, 1991, S. 227)

Der Wert dieser Detektionsschwelle orientiert sich in der vorliegenden Implementierung am Wert des Skalierungsfaktors S_{dis} für das jeweilige auralisierte Beispiel eines DuTs. Sprich wenn ein Teilnehmer des Versuchs die nichtlinearen Verzerrungen eines Treibers bei dessen tatsächlicher Performance ($S_{dis} = 0$ dB) mit 75% Wahrscheinlichkeit richtig detektiert, so ist seine Hörschwelle *AT* für die getestete Kombination aus Treiber und Stimulus 0 dB. Müssen die nichtlinearen Verzerrungen mit $S_{dis} = 6$ dB verstärkt werden um zielischer detektiert zu werden, so ist die Hörschwelle analog AT = 6 dB.

4 Optimierung & Evaluation des Perzeptiven Evaluationsmoduls

Dieses Kapitel stellt die vorgenommenen Veränderungen am PEQ-Modul vor. Die MOVs werden evaluiert und vorgenommene Optimierungen werden veranschaulicht und begründet. Darüber hinaus wird ein Ausblick auf weitere Verbesserungen gegeben, die Gegenstand zukünftiger Entwicklungen sein können.

4.1 Update auf MAT2

Das MAT-Modul dient als Schnittstelle zwischen der Software dB-Lab und der numerischen Rechenanwendung *Scilab*. Es ermöglicht das Einbinden von Skripten, deren Ein- und Ausgaben in der grafischen Oberfläche von dB-Lab erfolgen können. Das Modul unterstützt zwei Versionen, die den Rahmen für die zu entwickelnden Skripte bilden. Innerhalb einer Datenbank von dB-Lab werden die Skripte an Operationen gebunden, die einem Treiberobjekt zugeordnet sind.

4.1.1 Anpassung der Struktur

Das PEQ-Modul lag zu Beginn der Arbeit im MAT1-Format vor. Dieses Format ermöglicht das Einbinden weiterer skriptbasierter Funktionsbibliotheken und alloziert in einem vordefinierten Header Ein- und Ausgaben des Hauptskripts. Die Struktur des Hauptskripts ist dabei sequentiell. Daraus ergeben sich einige Nachteile die den Umgang mit MAT1-Skripten erschweren. Exemplarisch sind die folgenden Probleme zu nennen:

- Die sequentielle Struktur ermöglicht keine Überprüfung der Nutzereingaben vor Aufruf des Hauptskripts. Fehlerhafte Eingaben können zum Abbruch des Hauptskripts führen. Je nach Länge und Komplexität wird damit unnötig Rechenzeit verschwendet.
- Sämtliche konstante Vorgaben werden innerhalb des Hauptskripts bei jedem Aufruf neu angelegt. Dieses Verhalten ist ebenfalls nicht rechenzeiteffizient.
- Parameterübergaben erfolgen über eine textbasierte Eingabemaske. Die Benutzbarkeit leidet enorm darunter.

Das MAT2-Format stellt die Weiterentwicklung von MAT1 dar und ermöglicht die Lösung der genannten Probleme. Die Struktur des Hauptskripts ist dabei in verschiedene Funktionsblöcke untergliedert, die an entscheidenden Zeitpunkten der Nutzerinteraktion mit dB-Lab ausgeführt werden. Bezogen auf die zuvor dargelegten Probleme und des Updates des Moduls sind als wichtigste Blöcke die folgenden zu nennen:

- *OnCheckParams* ist ein Funktionsblock der die Prüfung von Benutzereingaben vor Ausführung der eigentlichen Berechnungen ermöglicht. Diese Prüfung wurde für das PEQ-Modul für alle relevanten Benutzereingaben implementiert.
- OnInitParams stellt einen Funktionsblock dar, der alle konstanten Vorgaben, wie z.B. Fenster für grafische Ausgaben und deren Achsenbeschriftung, beim Laden des Skriptes in dB-Lab alloziert. Entsprechend kann der Nutzer schon vor Ausführung der Berechnung die Art der Darstellung der Ergebnisse erkennen. Dies wurde für sämtliche Darstellungen implementiert.

- OnLoad alloziert in einem separaten Funktionsblock Konstanten wie z.B. den Bezugschalldruck p_0 unmittelbar vor Ausführung der Berechnungen. Der Vorteil der separaten Allokation besteht in einer für den Entwickler erheblich vereinfachten Wartung aller Konstanten, da diese zentral innerhalb eines Blocks angelegt werden. Dies wurde für alle Konstanten innerhalb des PEQ-Moduls realisiert.
- MAT2 verfügt über eine verbesserte grafische Übergabe der Benutzereingaben innerhalb der *Property Page* in dB-Lab. Wenn möglich wurden Auswahllisten oder Checkboxen für die Festlegung der Benutzereingaben realisiert. Textbasierte Eingaben lassen sich dank der erweiterten Parameterdefinition des MAT2-Formats einfacher halten. Der Nutzer muss keine Wortmatrizen mehr erstellen; Pfade können wahlweise via Text oder Dateibrowser übergeben werden. Abbildung 19 illustriert die Verbesserung.
- *OnRun* ist der Funktionsblock für die eigentliche Datenverarbeitung. Für das PEQ-Modul wurden die in Abschnitt 3.4.2 beschriebenen Funktionsblöcke in diesen Block portiert bzw. implementiert.

| ✓ xB_WAV TC\ProgramData\Kippet\BnD\Scripts5\vi | Paste | | Standard Inputs | | | Past |
|--|-------|----------------------------|-----------------------------|-------------------------------------|---|------|
| xT_WAV C:\ProgramData\Kippel\RnD\Scripts5\vi | | | (WAV) Reference Signal | C:\ProgramData\Klippel\RnD\Scripts5 | | - |
| xR_VEC | Edit | | (WAV) Test Signal | C:\ProgramData\Klippel\RnD\Scripts5 | | Liea |
| xT_VEC | | | (VEC) Reference Signal | | | |
| ✓ GR () ⇒ Advanced | | | (VEC) Test Signal | | | |
| | | Reference Gain (0 default) | 0 | | | |
| | | Input Mode | wave file 🔹 | 7 | | |
| < > | | Wave Export? | | | | |
| | | 8 | Advanced Settings | | | |
| | | | Target SPL Reference for 60 | | | |
| | | | Mode for Equalization G | by level | | |
| | | | Bypass Equalization Gain | | | |
| | | | Correlation Treshold in p | 75 | | |
| | | | Known Delay of Test to | | - | |
| | | Ir | put Mode | | | |

Abbildung 19: Vergleich der Eingabemasken. links: MAT1, rechts: MAT2.

4.1.2 Implementierung von Zeit- und Frequenzcursor

In der MAT1-Version des Moduls bestand keine Möglichkeit bereits berechnete Daten in die grafische Ausgabe von dB-Lab zu laden ohne das Hauptskript nochmals auszuführen. Eine Abfrage von Spektren zu einem bestimmten Zeitpunkt oder der Erregungsmuster für eine bestimmte Frequenzgruppe war somit nur umständlich möglich. MAT2 hingegen unterstützt *Events*. Events sind Benutzerinteraktionen, die verschiedene Funktionen aufrufen können. Somit können Cursor innerhalb der grafischen Ausgabe implementiert werden, die vom Nutzer bewegt werden können. Jede Bewegung eines Cursors ist ein Event, das die zugehörige Funktion aktiviert.

Für das PEQ-Modul wurden ein Zeitcursor und ein Frequenzcursor implementiert. Der Zeitcursor kann in allen Fenstern mit Zeitachsen als Abszisse entlang dieser verschoben werden. Die Verschiebung ruft eine Funktion auf, welche die FFT-Spektren von Referenz-, Test- und Differenzsignal, sowie die zugehörigen Erregungsmuster und Lautheitsverteilungen für den jeweiligen Zeitpunkt darstellt. Der Frequenzcursor kann in allen Fenstern mit Frequenzachsen als Abszisse entlang dieser verschoben werden. Die zugehörige Funktion aktualisiert die Darstellung der Zeitverläufe der Erregungspegel der jeweiligen Frequenzgruppe.

Die Cursors wurden innerhalb des Funktionsblocks OnEvent des MAT2-Formats implementiert.

4.1.3 Optimierung der Migrationsfähigkeit

Eine Kernanforderung an Module innerhalb von dB-Lab ist, dass Datenbanken bzw. Elemente aus Datenbanken ohne Datenverlust zwischen verschiedenen Nutzern bzw. Maschinen migriert werden können.

Die MAT1-Version des PEQ-Moduls erfüllte diese Anforderung nicht. Sämtliche berechnete Modelldaten wurden auf dem Hauptspeicher des Computers hinterlegt und von dort für die grafische Ausgabe in dB-Lab ausgelesen. U.a. ließ der Einsatz von absoluten Dateipfaden die Migration der Datenbank bzw. der Operation zwischen verschiedenen Maschinen nicht zu. Die Migration zwischen verschiedenen Benutzerkonten auf der selben Maschine funktionierte nur eingeschränkt.

Das Automation Interface unterstützt das Einbinden von Anhängen an Operationen, sog. *Attachments*. Mittels dieser Funktionalität wurde implementiert, dass sämtliche Modelldaten, die zur Darstellung oder Weiterverarbeitung auf einem anderen Computer nötig sind, direkt an die Operation gebunden sind und nicht mehr auf der Festplatte hinterlegt werden müssen. Die Operation kann somit wahlweise separat oder mit der jeweiligen Datenbank migriert werden.

4.2 Anpassungen des Filterbankmodells

Die benutzte Implementierung des Filterbankmodells wurde von Straube hinsichtlich der psychoakustischen Qualifikation ausreichend evaluiert (vgl. 2013, S. 66 ff.). Entsprechend wird an dieser Stelle darauf verzichtet die Prüfung der Auditory Filter, Erregungsmuster und Maskierungssimulation zu wiederholen.

Jedoch wurde die Implementierung als eigene Funktionsbibliothek angelegt und für den Einsatz für verschiedene Analysemethoden die Skalierung der Eingangssignale und die Hochpassfilterung entfernt. Diese sind für den jeweiligen Zweck separat zu implementieren. Eine Prüfung ergab hierbei Verbesserungsbedarf im PEQ-Modul.

4.2.1 Skalierung der Eingangssignale

Zunächst konnte festgestellt werden, dass die Skalierung der Eingangssignale für das Filterbankmodell nicht implementiert war. Folglich war der simulierte Schalldruckpegel für das Filterbankmodell ausschließlich von der Aussteuerung der Audiodateien abhängig und die Skalierer G_L und G_P für diesen Teil des perzeptiven Evaluationsmoduls nichtig und nicht vom Nutzer zu beeinflussen. Nach Gl. 9 konnte somit lediglich ein Spitzenpegel von 94 dB SPL simuliert werden, da die Speicherung digitaler Audiodateien keinen höheren Amplituden als 1 erlaubt. Dies wurde durch die Implementierung der Skalierer für das Filterbankmodell behoben.

Zunächst werden die synchronisierten Zeitsignale nach Gl. 13 und Gl. 14 auf den Bezugsschalldruck p_0 normiert:

$$p_R = \frac{y_R}{p_0}$$
 mit $p_0 = 2 \cdot 10^{-5}$ Pa und GI. 13

$$p_T = \frac{y_T}{p_0}$$
Gl. 14

Anschließend werden die Schalldruckverläufe mit den Skalierern auf den Zielschalldruckpegel gebracht. Gl. 16 berücksichtigt dabei für das Testsignal auch den Equalization Gain:

$$p_R^* = p_R \cdot 10^{\left(\frac{G_P}{20}\right)}$$
Gl. 15

$$p_T^* = p_T \cdot 10^{\left(\frac{G_P + G_L}{20}\right)}$$
 Gl. 16

4.2.2 Hochpassfilter

Mit der im PEQ-Modul zunächst implementierten Version des Hochpassfilters zur Vermeidung tieffrequenter Störungen folgt Straube dem Vorschlag der ITU zur Verwendung zwei kaskadierter Butterworth-Hochpassfilter zweiter Ordnung mit einer Grenzfrequenz $f_g = 20$ Hz (vgl. 2013, S. 27). Die Übertragungsfunktion ist in Gl. 17 mit den entsprechenden Filterkoeffizienten für eine Abtastfrequenz $f_s = 48$ kHz angegeben:

$$H_{HP,PEAQ} = \frac{(z-1)^2}{z^2 + a_{01}z + a_{02}} \cdot \frac{(z-1)^2}{z^2 + a_{11}z + a_{12}}$$
Gl. 17
mit $a_{01} = -1,99517, a_{02} = 0,995174, a_{11} = -1,99799, a_{12} = 0,997998.$

Das kaskadierte IIR-Filter wurde mit Scilab implementiert, jedoch erlaubte die verwendete Methode nicht das Speichern der Filterzustände am Ende eines Signalblocks. Da die Zeitsignale y_R und y_T blockweise eingelesen werden führte dies in der Folge dazu, dass das Filter für jeden Block erneut einschwingen musste. Entsprechend waren für stationäre Signale im angeschlossenen Rauigkeitsmodell teils starke Artefakte zu sehen. Der Wechsel zu einer anderen Methode der Implementierung erlaubt das Speichern der Filterzustände am Ende eines Signalblocks, die als Anfangszustände für die Filterung des anschließenden Signalblocks benutzt werden. Auf diese Weise konnten die Artefakte eleminiert werden.

Weiterhin ist die Anforderung an das Modul, dass die Eingangssignale auch in anderen Abtastraten als der von PEAQ verwendeten 48 kHz verarbeitet werden können. Entprechend musste auch das Hochpassfilter angepasst werden. Da nicht ersichtlich ist welche Kriterien bei der Entwicklung des PEAQ-Standards die Kaskadierung von zwei Filtern zweiter Ordnung gefordert hat (vgl. ITU, 2001, S. 47), wurde für das PEQ-Modul alternativ ein Butterworth-Hochpassfilter vierter Ordnung implementiert. Die Systemgleichung im z-Bereich und die zugehörigen Koeffizienten für $f_s = 48$ kHz sind Gl. 18 zu entnehmen:

$$H_{HP,PEQ} = \frac{a_1 z^4 + a_2 z^3 + a_3 z^2 + a_4 z + a_5}{z^4 + b_1 z^3 + b_2 z^2 + b_3 z + b_4}$$
Gl. 18
mit $a_1 = 0.9965853$, $a_2 = -3.9863411$, $a_3 = 5.9795116$,
 $a_4 = -3.9863411$, $a_5 = 0.9965853$
und $b_1 = -3.9931589$, $b_2 = 5.9795$, $b_3 = -3.9795233$, $b_4 = 0.9931822$.

Die Implementierung passt die Koeffizienten des Filters dynamisch der Abtastrate der übergebenen Eingangssignale x_R und x_T an. Dabei wurde an Hand des vorgegeben Filters der PEAQ-Filterbank geprüft, dass die Abweichungen der neuen Implementierung hinsichtlich des Amplitudenfrequenzgangs und des Phasengangs für $f_s = 48$ kHz vernachlässigbar sind.





Abbildung 20: Vergleich der Hochpassfilter

Blau / PEAQ: Filter des PEAQ-Verfahrens (vgl. ITU, 2001, S. 47). Rot: Butterworth-Filter vierter Ordnung, Hochpass.

Die blauen Kurven zeigen das Hochpassfilter der PEAQ-Filterbank, während die roten Kurven das neu implementierte Filter darstellen. Die Abweichungen sind für die gegebene Anwendung vernachlässigbar klein, wenngleich die neue Implementierung unterhalb der Grenzfrequenz eine minimal verbesserte Dämpfung aufweist und der Phasengang außerhalb des Durchlassbereiches etwas abweicht.

4.3 Rauigkeitsmodell

Die Implementierung des Rauigkeitsmodells geht aus einer Studienarbeit von Kunz hervor und basiert auf dem Modell von Daniel und Weber (vgl. 2014, S. 27 ff.).

Offenbar lag der Quellcode in einer veränderten Version vor, da das Rauigkeitsmodell zunächst nicht funktionierte bzw. die Scilab-Instanz abstürzen ließ. Daher werden zunächst die eingeführten Änderungen in der Implementierung beschrieben und nachfolgend der Vergleich zu den empirischen Daten von Zwicker gezogen.

4.3.1 Änderungen in der Implementierung

Die Ursache für das Abstürzen der Scilab-Instanz ist auf Fehler in der Blockverarbeitung zurückzuführen. Zwar schreibt Kunz dass die Implementierung komplett blockweise funktioniert (vgl. 2014, S. 29), jedoch wird in der zu Beginn der Arbeit vorliegenden Version des Quellcodes immer der vollständige Datensatz an die jeweiligen Funktionsblöcke des Rauigkeitsmodells übergeben. Wenngleich die Funktionsblöcke intern blockweise Arbeiten führt das vollständige Laden der Hüllkurven für lange Eingangssignale zu Problemen, da der

von Scilab allozierbare Arbeitsspeicher nicht ausreicht. Weiterhin lagen innerhalb einzelner Funktionsblöcke Indexfehler vor, die durch unterschiedliche Dimensionen der berechneten Matrizen zu Abbrüchen führten. Die Fehler wurden entfernt und eine konsequente Blockverarbeitung eingeführt. Die Blockverarbeitung teilt die Ausgangssignale der Filterbank nach der Bildung der Hüllkurven in Blöcke zu einer Länge von je t = 200 ms und liest diese nacheinander für die Funktionsblöcke des Rauigkeitsmodells ein. Für eine Abtastrate von $f_s = 48$ kHz der Eingangssignale werden die Signale bis zu dieser Stufe in der Filterbank auf $f_{s,red\#1} = 1500$ Hz heruntergetastet. Entsprechend haben die Blöcke für die Rauigkeitsberechnung eine Länge von N = 300 Samples.

Zusätzlich wurde eine Differenzberechnung zur grafischen Ausgabe hinzugefügt. Die Rauigkeit des Referenzsignals wird von der Rauigkeit des Testsignals blockweise abgezogen um die hervorgerufene Änderung durch die im Testsignal vorhandenen Verzerrungen übersichtlicher darstellen zu können.

Weiterhin wurde der Einfluss der Abtastrate auf das Rauigkeitsmodell untersucht. Hierfür wurden für gängige Abtastraten zwischen 44,1 kHz und 96 kHz Referenzschalle erzeugt und analysiert. Diese folgen der Vorgabe von Fastl & Zwicker und bestehen aus einem Sinuston der Trägerfrequenz $f_c = 1$ kHz, der mit einer Modulationsfrequenz $f_m = 70$ Hz bei einer Modulationstiefe von m = 1 amplitudenmoduliert wird (vgl. 2007, S. 257). Dabei war festzustellen, dass mit steigender Abtastrate die berechnete Rauigkeit stieg. Folglich wurde der Kalibrierfaktor *cal* frequenzabhängig implementiert. Die Werte für den Kalibierfaktor mit den zugehörigen Abtastraten können Tabelle 1 entnommen werden.

| Abtastrate f_s | Kalibrierfaktor cal | | | |
|------------------|---------------------|--|--|--|
| 44,1 kHz | 0,5786 | | | |
| 48 kHz | 0,5362 | | | |
| 88,2 kHz | 0,4357 | | | |
| 96 kHz | 0,4246 | | | |
| | | | | |

Tabelle 1: Kalibrierfaktoren des Rauigkeitsmodells bei variabler Abtastrate.

Ausgehend von den eingeführten Änderungen der Implementierung wird das Rauigkeitsmodell an Hand der von Fastl & Zwicker empirisch ermittelten Daten nachfolgend evaluiert.

4.3.2 Variation des Modulationsindexes bei fester Trägerfrequenz

Um den Einfluss der Modulationstiefe bei amplitudenmodulierten Signalen zu überprüfen wurden nach Gl. 19 Stimuli generiert (vgl. Kunz, 2014, S. 9):

$$x(t) = (1 + m \cdot \sin(2\pi \cdot f_m \cdot t)) \cdot \sin(2\pi \cdot f_c \cdot t)$$
 Gl. 19

Dabei wurde als Modulationsfrequenz $f_m = 70$ Hz und als Trägerfrequenz $f_c = 1$ kHz gewählt. Der Modulationsindex *m* wurde schrittweise von m = 0,1 bis m = 1 in einer Schrittweite von $\Delta m = 0,1$ erhöht. Die Stimuli wurden mit einer Abtastrate von $f_s = 48$ kHz, einer Wortbreite von 16 Bit und einer Länge von t = 4 s mit Scilab erzeugt. Zur Analyse wurde die Stimuli auf einen Schalldruckpegel $L_p = 60$ dB SPL skaliert. Die empirischen Daten liegen zum Vergleich nicht analytisch vor und wurden daher aus der Quelle digitalisiert (vgl. Fastl & Zwicker, 2007, S. 258). Abbildung 21 stellt die berechneten Daten (rot) den von Fastl & Zwicker erhobenen Daten (blau) und der entsprechenden Approxiamtion (blau, unterbrochen) gegenüber:



Abbildung 21: Rauigkeit bei variabler Modulationstiefe.

$$f_m = 70 \text{ Hz}, f_c = 1 \text{ kHz}, L_p = 60 \text{ dB}$$

Die Darstellung zeigt deutliche Abweichungen der Modellberechnung zu den erhobenen Daten, wenngleich die Steigungen ähnlich sind und alle Kurve beim Referenzschall übereinstimmen und 1 asper erreichen. Die Differenz beträgt am Punkt der größten Abweichung rund 0,2 asper für einen mittleren Modulationsindex von $m \approx 0,6$ und weicht somit für diesen Punkt um ca. 45 % von den Daten nach Fastl & Zwicker ab. Für kleinere und größere Modulationsindizes fällt die Differenz wesentlich geringer aus.

Daniel & Weber stellen an das Rauigkeitsmodell die Anforderung, dass die Abweichungen der berechneten Werte zu den empirisch erhobenen Daten innerhalb der *Just Noticeable Difference* (JND) liegen sollen (vgl. 1997, S. 114). Diese wird für den Einfluss des Modulationsindex mit einer Größe von $\Delta R/R = 17\%$ angegeben (vgl. Fastl & Zwicker, 2007, S. 260). Die durchgeführte Evaluation zeigt daher, dass diese Anforderung von der aktuellen Implementierung des Rauigkeitsmodell mit den vorgenommenen Änderungen verfehlt wird.

4.3.3 Variation der Modulationsfrequenz bei festen Trägerfrequenzen und konstantem Modulationsindex

In weiteren Versuchen haben Fastl & Zwicker den Einfluss der Modulationsfrequenz bei verschiedenen Trägerfrequenzen mit einem festen Modulationsindex von m = 1 untersucht und die Ergebnisse der Studie grafisch veröffentlicht (vgl. ebd., S. 259).

Auch die Abbildung dieses Zusammenhangs durch das Rauigkeitsmodell lässt sich an Hand der erhobenen Daten evaluieren. Hierzu wurden wiederum nach Gl. 19 Stimuli in Scilab generiert.

Jedoch wurde für diesen Teil der Evaluation der Modulationsindex m = 1 gesetzt und für die Trägerfrequenzen $f_c = \{125 \text{ Hz}, 250 \text{ Hz}, 500 \text{ Hz}, 1 \text{ kHz}, 2 \text{ kHz}, 4 \text{ kHz}, 8 \text{ kHz}\}$ die Modulationsfrequenz f_m variiert. Die Schrittweite Δf_m betrug 10 Hz. Der relevante Bereich für die Modulationsfrequenz erstreckt sich von 20 Hz bis 300 Hz. Für die Trägerfrequenzen 250 Hz und 500 Hz wurde die Modulationsfrequenz bis $f_c/2$ erhöht. Für $f_c = 125$ Hz wurde f_m zwischen 20 Hz und 100 Hz variiert um einen ähnlichen Bereich wie Zwicker & Fastl zu betrachten. Zur Analyse wurden die Stimuli gemäß der Vorgabe auf einen Schalldruckpegel von $L_p = 60$ dB SPL skaliert.

Die berechneten Daten sind den digitalisierten Daten in Abbildung 22 gegenübergestellt. Die unterbrochenen Kurven stellen die digitalisierten Daten von Fastl & Zwicker dar, die durchgezogenen repräsentieren die Modellberechnungen. Der Referenzpunkt ist durch ein Kreuz markiert. Die zugehörigen Trägerfrequenzen sind der Legende zu entnehmen:



Rauigkeit bei variabler Modulationsfrequenz

Abbildung 22: Rauigkeit bei variabler Modulationsfrequenz für mehrere Trägerfrequenzen.

$m = 1, L_p = 60 \text{ dB}.$

Die Übereinstimmung der Modellberechnungen mit den empirischen Daten fällt je nach Trägerfrequenz deutlich unterschiedlich aus. Durch die Kalibrierung stimmt das Modell für $f_c = 1$ kHz im Referenzpunkt überein, jedoch berechnet es für Modulationsfrequenzen unter 70 Hz zu große und für höhere Modulationsfrequenzen zu geringe Rauigkeiten. Die tendenziellen Verläufe für $f_c = 500$ Hz und $f_c = 1$ kHz weisen hohe Ähnlichkeit zu den empirischen Daten auf. Hingegen stimmen für die tieferen Trägerfrequenzen weder die Lage der Maxima noch die berechneten Größen der Rauigkeit überein. Die Abweichung beträgt hier bis zu 0,35 asper. Für Trägerfrequenzen oberhalb von 1 kHz stimmt die Lage der Maxima bzgl. Modulationsfrequenz hinreichend überein, jedoch werden generell zu große Rauigkeiten berechnet. Eine extreme Abweichung ist für die Trägerfrequenz 4 kHz zu finden und beträgt rund 0,5 asper bzw. 80% bei einer Modulationsfrequenz von 70 Hz. Hinsichtlich der von Daniel & Weber genannten Unterschiedsschwelle von $\Delta R/R = 17\%$ (vgl. 1997, S. 114) ist daher auch für die Abbildung der Änderung der Modulationsfrequenz auf die berechnete Rauigkeit bei verschiedenen Trägerfrequenzen festzuhalten, dass die Anforderung deutlich verfehlt wird. Weiterhin konnten die gute Beurteilung des implementierten Modells durch Kunz (vgl. 2014, S. 32 ff.) nicht nachvollzogen werden. Das implementierte Rauigkeitsmodell ist somit nur bedingt praxistauglich.

Die Ursachen für die unzureichende Abbildung der Rauigkeit können vielfältig sein. Da das Modell zum Großteil von der korrekten Bestimmung des Modulationsindexes aus der Hüllkurve des jeweiligen Filterkanals abhängt kann die Gestaltung der Filterbank deutlichen Einfluss haben: Beispielsweise nutzen Daniel & Weber eine Filterbank mit 47 Kanälen der Breite 0,5 Bark (vgl. 1997, S. 117), während die evaluierte Implementierung 40 Kanäle der Breite 0,707 Bark benutzt. Weiterhin wird die Korrelation zur Gewichtung des berechneten Modulationsindex des Kanals *i* im Vorschlag von Daniel & Weber mit den Kanälen *i*-2 und i+2 durchgeführt (vgl. ebd.), während die Implementierung von Kunz die Kanäle i-1 und i+1 benutzt.

Die exakte Ursache ist nicht Thema dieser Arbeit und konnte mit den durchgeführten Versuchen nicht ermittelt werden. Weiterhin wurde nach den bisherigen Ergebnissen der Evaluations des implemntierten Rauigkeitsmodells auf die Untersuchung des Einfluss der Frequenzmodulation verzichtet, da keine gute Übereinstimmung mit empirischen Daten im jetzigen Entwicklungsstand zu erwarten ist.

Neben der Optimierung der Berechnungsgenauigkeit können überdies Verbesserungen bei der Rechenzeit eingeführt werden. Durch Auswahl einer FFT-gerechten Blockgröße kann die Berechnungszeit je Block verkürzt werden. Für Eingangssignale mit $f_s = 48$ kHz wäre an Stelle der 300 Samples Blockgröße eine Länge von 256 Samples denkbar. Die ausgeführten Hin- und Rücktransformationen würde in dem Falle als FFT und nicht wie bisher als DFT ausgeführt. Speziell für lange Eingangssignale ist ein signifikanter Performancegewinn zu erwarten. Die Frequenzauflösung verschlechtertet sich im genannten Falle um rund 15%; den Einfluss dieser geringeren Auflösung gilt es in einer weiteren Entwicklung zu prüfen.

4.4 Berechnung der Schärfe

Die Berechnung der MOV Schärfe wurde von Liebig implementiert aber nicht weiter evaluiert (vgl. 2012, S. 41). Die Schärfe wird in der aktuellen Implementierung als Ausgabe des FFT-Modells berechnet (siehe Abbildung 14). Sie folgt der von Fastl & Zwicker vorgegeben Vorschrift zur Berechnung in Gl. 20 mit Anpassungen an das FFT-Modell (vgl. 2007, S. 242):

$$S = 0.11 \frac{\int_0^{24Bark} N'g(z)z \, dz}{\int_0^{24Bark} N'dz} acum$$
Gl. 20

Dabei wird die spezifische Lautheit N' über den Frequenzgruppen integriert und im Zähler der Gleichung mit der frequenzabhängigen Wichtungsfunktion g(z) gewichtet. Diese Wichtungsfunktion stellt den erhöhten Einfluss von spektralen Anteilen oberhalb von ca. 3 kHz auf die Empfindung der Schärfe eines Klangs nach und wird nicht analytisch vorgegeben. Entsprechend wird die Funktion angepasst auf das FFT-Model approximiert; die Quelle der Approximationsgleichung konnte nicht nachvollzogen werden.

Die Evaluation des implementierten Berechnungsmodells erfolgt an Hand der von Fastl & Zwicker erhobenen Daten. Diese lagen nicht numerisch vor und wurden daher aus den gegebenen Diagrammen der Veröffentlichung digitalisiert (vgl. Fastl & Zwicker, 2007, S. 240). Die Umrechnung von der Barkskala in Hertz folgt dem Vorschlag von Völk gemäß Gl. 21 (vgl. 2015, S. 4):

$$\frac{f(z)}{Hz} \square 873, 47 \left[\left(\frac{32,12}{32,12 - \frac{z}{Bark}} \right)^{2,5} - 1 \right]^{\frac{1}{1,18}}$$
GI. 21

mit $0 \le z \le 24,86$ *Bark*

Drei Gruppen von Stimuli wurden für die Evaluation benutzt. Die erste Gruppe beinhaltete Rauschen bei einer festen unteren Grenzfrequenz von $f_{gu} = 0,2$ kHz und einer variablen oberen Grenzfrequenz f_{go} . Die zweite Gruppe hatte eine feste obere Grenzfrequenz $f_{go} = 10$ kHz und eine variable untere Grenzfrequenz f_{gu} . Als variable Grenzfrequenzen wurden 0,5 kHz, 1 kHz, 2 kHz, 5 kHz und 8 kHz gewählt. Die dritte Gruppe beinhaltete Rauschen in der kritischen Breite Δf der jeweiligen Frequenzgruppe mit der Mittenfrequenz f_m . Als Mittenfrequenzen wurden wiederum 0,5 kHz, 1 kHz, 2 kHz, 5 kHz und 8 kHz gewählt. Die kritische Breite wurde gemäß Fastl & Zwicker nach Gl. 22 approximiert (vgl. 2007, S. 159):

$$\Delta f \Box 0, 2 \cdot f_m \qquad \qquad \text{Gl. 22}$$

mit $f_m \ge 500 \text{ Hz}$

Die Stimuli wurden im Frequenzbereich in Scilab entworfen und in den Zeitbereich transformiert, bei einer Abtastrate $f_s = 48$ kHz, einer Wortbreite von 16 Bit und einer Länge von t = 5 s. Für die Analyse mit dem PEQ-Modul wurden die Stimuli auf einen Lautstärkepegel von $L_N = 60$ phon skaliert.

Die Modelldaten sind den digitalisierten Daten in Abbildung 23 gegenübergestellt:



Abbildung 23: Schärfe. Vergleich zwischen Modell und Daten nach Fastl & Zwicker.

Der Vergleich ergibt tendenziell gute Übereinstimmungen der berechneten Daten zu den empirisch erhobenen Daten.

Für Rauschsignale mit der Bandbreite Δf wird die Rauigkeit für Mittenfrequenzen $f_m \le 3$ kHz zu gering bewertet. Oberhalb von 3 kHz berechnet das Modell etwas mehr Schärfe als empirisch ermittelt. Der Referenzpunkt ist in der Darstellung durch ein schwarzes Kreuz markiert. Hier beträgt die Abweichung zwischen dem Modell und der Vorgabe ca. 0,1 acum. Für Rauschsignale mit fester unterer Grenzfrequenz von $f_{gu} = 0,2$ kHz berechnet das Modell im Vergleich zu den empirischen Daten etwas zu geringe Werte. Die Schärfe der Rauschsignale mit fester oberer Grenzfrequenz von $f_{go} = 10$ kHz wird für große Bandbreiten (z.B. $f_{gu} = 0,5$ kHz) und sehr kleine Bandbreiten (z.B. $f_{gu} = 8$ kHz) vom Modell überbewertet.

Eine Reihe von möglichen Ursachen können diese Abweichungen bedingen. Neben der Auflösung der Digitalisierung der empirischen Daten kann zunächst die Approximation der Wichtungsfunktion g(z) aus Gl. 20 zu Abweichungen führen. Weiterhin kann ein Unterschied in den Bandbreiten der verwendeten Stimuli begründet sein: Während für die kritische Bandbreite Δf dieser Evaluation die Approximation nach Gl. 22 erfolgte, könnten für die Erhebung der Daten Stimuli mit den jeweils exakt berechneten Bandbreiten der jeweiligen Frequenzgruppe benutzt worden sein.

Zuletzt ist die Erzeugung der Stimuli zu berücksichtigen. Unter der Annahme das zur Erhebung der empirischen Daten gefiltertes weißes Rauschen genutzt worden ist, werden die resultierenden Flanken im gefilterten Signal das Ergebnis beeinflussen. Zur Veranschaulichung sind in Abbildung 24 die Spektren des zur Evaluation genutzten Schmalbandrauschens (blau) und eines gefilterten weißen Rauschens (rot) für die Eckfrequenzen $f_{gu} = 8$ kHz und $f_{go} = 10$ kHz abgebildet. Für letzteres wurde ein Bandpassfilter vierter Ordnung nach Butterworth benutzt. Die FFT wurde mit einer Länge von $n_{fft} = 2^{14}$ Samples durchgeführt und entspricht der vom PEQ-Modul genutzten FFT-Länge.



Abbildung 24: Vergleich der Amplitudenspektren zweier Rauschsignale.

Die vom Modell berechneten Schärfen für die dargestellten Rauschsignale sind Tabelle 2 zu entnehmen und weichen immerhin um rund 0,32 acum voneinander ab. Entsprechend ist ein Einfluss durch die Auswahl des Filters und der zugehörigen Flanken gegeben, welcher ursächlich für eine Abweichung zwischen den Daten nach Fastl & Zwicker und der Modellberechnung sein kann.

| Signal | Schärfe [acum] |
|-------------------------------|----------------|
| Weißes Rauschen, gefiltert | 8,092 |
| Schmalbandrauschen | 8,409 |

Tabelle 2: Berechnete Schärfe S.

4.5 Distortion To Mask Ratio

Das Maskierungsverhältnis bzw. die MOV Distortion To Mask Ratio wurde für die Anwendung an Audiocodecs entwickelt um die Hörbarkeit der durch den Codec verursachten Verzerrungen zu quantifizieren. Die Berechnung erfolgt gemäß Gl. 23 und stellt den Abstand zwischen der Maskierungsschwelle ("*Masking Pattern*") des Referenzsignals und dem im FFT-Modell berechneten Fehlersignal ("*Distortion Pattern*") als logarithmisches Maß dar (vgl. Liebig, 2012, S. 37).

$$DMR(b) = 10\log_{10}\left(\frac{1}{Z}\sum_{k=0}^{Z-1}\frac{DP(k,b)}{MP(k,b)}\right)dB$$
Gl. 23

mit b = Blocknummer, k = Frequenzgruppe

und Z = Anzahl der Frequenzgruppen

Die erfolgte Mittelung über die Frequenzgruppen schränkt die Aussagekraft dieser MOV hinsichtlich der Anwendung für nichtlineare Verzerrungen ein. Während für Audiocodecs wenig bis keine Interaktion zwischen verschiedenen spektralen Bereichen zu erwarten ist, bestehen bei Lautsprechern mit nichtlinearen Verzerrungen häufig starke Wechselwirkungen. Folglich kann eine Mittelung über alle Frequenzgruppen nur als Anhaltspunkt für die Hörbarkeit nichtlinearer Verzerrungen bei Lautsprechern dienen.

Zur Verdeutlichung stellt Abbildung 25 ein Maskierungspattern (blau) sowie ein Verzerrungspattern (rot) für einen gegebenen Zeitpunkt dar. Es handelt sich um die Auralisation elektronischer Tanzmusik eines mittels LPM und LSI vermessenen Woofers mit moderaten Kraftfaktorverzerrungen und starken Steifigkeits- und Induktivitätsverzerrungen. Das Referenzsignal ist das linear verzerrte Signal und das Testsignal die Summe aus linearen und nichtlinearen Verzerrungen. Das berechnete *DMR* für diesen Zeitpunkt beträgt -3,8 dB. Der Wert impliziert zunächst, dass die Verzerrungen vollständig maskiert sein sollten. Jedoch zeigt die Grafik, dass das Verzerrungssignal im Bereich zwischen 100 und 200 Hz deutliche harmonische Verzerrungen enthält, die oberhalb der Maskierungschwelle liegen und auf die nichtlineare Steifigkeitskennlinie des Treibers zurückzuführen sind. Die Verzerrungen sind gegensätzlich zum berechneten *DMR*-Wert deutlich hörbar.



Abbildung 25: Masking Pattern und Distortion Pattern. DMR = -3,8 dB.

Auralisation eines Woofers mit geringen Bl(x)-Verzerrungen, starken $K_{ms}(x)$ - und $L_e(x)$ -Verzerrungen, Stimulus: elektr. Tanzmusik.

Zusätzlich sei der Bereich um 2 kHz erwähnt: Hier dominieren Intermodulationsverzerrungen das Verzerrungspattern, deren Ursache die nichtlineare Induktivitätskennlinie des Lautsprechers ist. Diese liegen zwar rund 10 dB unterhalb der Maskierungsschwelle, können jedoch noch wahrgenommen werden. Drastischer wird dies in Abbildung 26:



Abbildung 26: Masking Pattern und Distortion Pattern. DMR = -11,9 dB.

Auralisation eines Woofers mit starken Bl(x)-Verzerrungen, moderaten $K_{ms}(x)$ - und geringen $L_e(x)$ -Verzerrungen, Stimulus: elektr. Tanzmusik.

Hierbei wurde der gleiche Stimulus mit einem anderen Woofer mit starken Kraftfaktorverzerrungen auralisiert. Das Verzerrungssignal enthält für den gegebenen Zeitpunkt über weite Teile der Frequenzgruppen Intermodulationsverzerrungen, die deutlich wahrnehmbar sind. Der berechnete *DMR*-Wert hingegen beträgt -11,9 dB und impliziert somit vollständige Maskierung der Verzerrungen.

Auch der ausgewählte Modus zur Pegelkompensation hat einen signifikanten Einfluss auf die berechneten Werte. Je nachdem, ob der Equalization Gain auf Basis des RMS-Werts der Eingangssignale oder ihrer Lautheiten berechnet wird, kann die Abweichung der berechneten Werte des *DMR* bei entsprechend starken Verzerrungen mehrere dB betragen.

Die Beispiele illustrieren, dass das Maskierungsverhältnis nur als Anhaltspunkt für die perzeptive Bewertung nichtlinearer Verzerrungen genutzt werden kann. Die Mittelung der *DMR*-Werte über alle Blöcke zur Berechnung der Total Distortion To Mask Ratio führt dabei zu keiner objektiven Verbesserung: So hat Liebig in einem Hörversuch herausgestellt, dass eine Detektionsschwelle für einen gegebenen Test durchaus einem Wert von $DMR_{tot} = -20$ dB entsprechen kann (vgl. 2012, S. 79). Eine mögliche Erklärung für die Ungenauigkeit des Verfahrens besteht darin, dass es sich bei dem Fehlersignal des FFT-Models nicht wie bei den Erregungsmustern der Referenz- und Testsignale um eine interne Repräsentation des Hörvermögens handelt. Für die perzeptive Bewertung nichtlinearer Verzerrungen gilt es daher ein aussagekräftigeres Maß zu finden. Zunächst muss das *DMR* jedoch als Abschätzung für die Hörbarkeit von Verzerrungen herangezogen werden.

4.6 Discoloration, Treble Stressing, General Bass Enhancement

Ähnlich wie das Maskierungsverhältnis sind auch die relativen Maße Discoloration, Treble Stressing und General Bass Enhancement direkt von dem ausgewählten Modus zur Berechnung des Equalization Gain betroffen. Durch die Berechnung der Maße aus den spezifischen Lautheitsverteilungen der frequenz- und zeitverschmierten Eingangssignale im FFT-Modell, welche wiederum nach der Anwendung des Equalization Gain erfolgt, besteht ein direkter Einfluss. Es konnten Differenzen von mehreren Prozent für Spitzenwerte festgestellt werden.

Der relative Bezug der Maße durch die Angabe in Prozent eignet sich gut als deskriptives Vokabular um den groben Einfluss nichtlinearer Verzerrungen gegenüber technisch nichtversierten Personen darzustellen. Gleichwohl besteht eine inhaltliche Dopplung zwischen der absoluten Schärfe und dem Maß Treble Stressing. Die Betrachtung der jeweils absolut oder relativ angegebenen Werte kann nichtversierte Anwender durcheinanderbringen.

4.7 Anwendung auf ausgewählte Nichtlinearitäten

Die Charakteristiken der jeweiligen nichtlinearen Verzerrungen lassen vermuten, dass je nach dominierender Verzerrungsart bestimmte MOVs des PEQ-Moduls bzw. bestimmte Kombinationen dieser signifikante Änderungen aufweisen. Zur Untersuchung dieser These wurden mehrere Versuche durchgeführt, die folgend beschrieben werden.

Für reguläre nichtlineare Verzerrungen wurden ein Titel aus dem Genre elektronischer Tanzmusik auf den in Abschnitt 4.5 verwendeten Treibern mit dem SIM-AUR-Modul auralisiert. Die Treiber weisen jeweils dominierende Bl(x)-Verzerrungen oder $K_{ms}(x)$ - und $L_e(x)$ -Verzerrungen auf. Für die Simulation des Großsignalbereichs wurde in diesem Fall jede der genannten Nichtlinearitäten exklusiv auralisiert und die resultierenden Verzerrungen mit $S_{dis} = 12$ dB hochskaliert. Zur Analyse im PEQ-Modul diente das linear verzerrte Signal als Referenzsignal und die Summe aus linear und nichtlinear verzerrtem Signal als Testsignal. Die Berechnung des Equalization Gain erfolgte im Loudness Modus. Zur Erzeugung irregulärer nichtlinearer Verzerrungen wurden zwei weitere Titel des selben Genres mit einem Breitbandlautsprecher wiedergegeben und davon Mikrofonaufnahmen angefertigt. Abbildung 27 zeigt den Versuchsaufbau beispielhaft. Dabei wurden folgende irreguläre Nichtlinearitäten bzw. Defekte erzeugt: *Hard Limiting* beschreibt das Anschlagen der Membran an einem vorgehaltenem Stift, *Loose Particle* wurde durch Salzkörner auf der Membran realisiert und *parasitäre Vibrationen* mit schwereren Kleinteilen auf der Membran erzeugt. Als Referenzsignal diente in diesem Fall die Aufnahme ohne Defekte, während als Testsignale die jeweiligen verzerrten Aufnahmen benutzt wurden. Weder die Verstärkung noch die räumliche Anordnung des Testaufbaus wurde während des Versuchs verändert. Die Berechnung des Equalization Gain erfolgt ebenfalls im Loudness Modus.



Abbildung 27: Versuchsaufbau für irreguläre Nichtlinearitäten.

Die berechneten MOVs wurden hinsichtlich ihrer Änderung untersucht und die Ergebnisse als starke Zu- oder Abnahmen, moderate Zu- oder Abnahmen oder nicht nennenswerte Veränderungen klassifiziert. Tabelle 3 stellt die Daten dar:

| Kategorie | Nichtlinearität | V | BE | TS | S | R |
|-------------------|---------------------|----------|---------|---------|----------|---------|
| Reguläre | Bl(x) | ++ | ++ | - | - | ++ |
| Nichtlinearitäten | K _{ms} (x) | ++ | ++ | 0 | - | 0 |
| | $L_{e}(x)$ | ++ | 0 | 0 | 0 | ++ |
| Irreguläre | Hard Limiting | ++, imp. | -, imp. | 0 | +, imp. | +, imp. |
| Nichtlinearitäten | Loose Particle | +, imp. | 0 | +, imp. | +, imp. | 0 |
| | Parasitäre | ++, imp. | -, imp. | 0 | ++, imp. | +, imp. |
| | Vibrationen | | | | | |

Tabelle 3: MOVs und nichtlineare Verzerrungen.

++: starke Zunahme, +: moderate Zunahme, 0: keine nennenswerte Änderung, -: moderate Abnahme, imp.: impulshafte Änderungen.

Für die regulären nichtlinearen Verzerrungen lassen sich Muster feststellen. Allen gemein sind eine starke berechnete Verfärbung V und nur geringe Änderungen der Schärfe S. Die durch den nichtlinearen Kraftfaktor Bl(x) erzeugten harmonischen und Intermodulationsverzerrungen bewirken starke Zunahmen bei der Rauigkeit und der Bassbetonung *BE*. Diese Kombination findet sich nicht für die weiteren Nichtlinearitäten.

Die Steifigkeitsverzerrungen führen vor allem zur starken Zunahme der Bassbetonung *BE*, da vorrangig harmonische Verzerrungen zu den Basstönen der Stimuli erzeugt werden. Die Änderungen der übrigen MOVs sind hierfür gering oder vernachlässigbar. Die nichtlineare Induktivität erzeugt starke Intermodulationsverzerrungen, welche die berechnete Rauigkeit stark zunehmen lassen. Die übrigen MOVs ändern sich nicht nennenswert.

Den irregulären Nichtlinearitäten ist ein impulshafter Charakter der Änderungen gemein. Für den Fall *Hard Limiting* entstehen folglich nur dann sprunghafte Änderungen, wenn die Auslenkung *x* der Membran diese an die vorgehaltene Begrenzung anschlagen lässt. Durch die Begrenzung entstanden im Versuch moderate Änderungen der Rauigkeit und Schärfe. Die durch die Begrenzung verminderte Auslenkung vermindert die Basswiedergabe, was sich in einer geringen Abnahme der Modellausgabe Bass Enhancement zeigt. Lose Partikel auf der Membran führen zu moderaten Zunahmen der Verfärbung *V* und der Höhenbetonung *TS*. Ebenfalls steigt die Schärfe moderat an. Die Auswirkungen parasitärer Vibrationen sind stark abhängig von den vibrierenden Teilen. Die Tabelle spiegelt den Fall eines Plastikteils am Membranrand wieder. Die in Abbildung 27 dargestellte Maschinenschraube mit Mutter spricht andere MOVs an.

Die Versuche sind zwar nicht repräsentativ, jedoch ist festzuhalten, dass für reguläre nichtlineare Verzerrungen eindeutige Muster erkennbar sind. Das PEQ-Modul eignet sich daher für die Evaluation dieser Verzerrungen. Die Ergebnisse korrelieren mit den Überlegungen zur psychoakustischen Relevanz aus Abschnitt 2.3.2. Für irreguläre Nichtlinearitäten erscheinen die Änderungen der MOVs weniger systematisch. Dies ist vor allem durch die vielfältigen Erscheinungsformen dieser Verzerrungen bedingt und kein Defizit des Moduls.

4.8 Ausblick

Weitere Optimierungen sind v.a. hinsichtlich der Rechenzeit notwendig. Derzeit benötigt das PEQ-Modul für die Verarbeitung eines Stimulus der Länge t_{sig} in Sekunden ca. 4,17 $\cdot t_{sig}$ Sekunden Rechenzeit¹. Der Großteil davon entfällt auf die Ausführung der perzeptiven FFTund Filterbankmodelle. Da das Filterbankmodell ohnehin ausgeführt werden muss um die Rauigkeit berechnen zu können, kann in einer künftigen Version die Berechnung der weiteren MOVs auch daran angeschlossen werden. Dann kann das FFT-Modell entfernt werden was zu einer drastischen Verkürzung der Rechenzeit führen wird.

Hinsichtlich der Model Output Variables gilt es verlässlichere Größen zur Klassifizierung der Hörbarkeit von Verzerrungen zu finden. Wie gezeigt wurde sind die MOVs Distortion To Mask Ratio und Total Distortion To Mask Ratio nicht hinreichend genau, um zielsicher die Hörbarkeit bzw. Maskierung von Verzerrungen zu zeigen. Ausgehend von dem Vorschlag der exklusiven Benutzung des Filterbankmodells wäre ein Maskierungsverhältnis je Frequenzgruppe bzw. Filterbankkanal möglich. Die psychoakustischen Zusammenhänge sind bekannt und es kann ausgehend vom Referenzsignal je Filterbankkanal eine Hüllkurve des Maskierungsverlaufs berechnet werden, die mit einer zu berechnenden Hüllkurve des Differenz- bzw. Fehlersignals zu vergleichen ist. Ausgehend von der größten Überschreitung des Maskierungsverlaufs und dem zugehörigen Zeitindex können somit schnell kritische Zeitpunkte für die Hörbarkeit der Verzerrungen identifiziert werden. Alternativ kann der Zeitverlauf der bisher implementierten Distortion To Mask Ratio auch mit

¹ Konfiguration des Bezugscomputers: 4-Kern-Prozessor Intel i7 (4x 2,6 GHz), 6. Generation aus 2016, 16 GB Arbeitsspeicher, Solid State Hauptspeicher, Windows 10, dB-Lab 209.134. Berechnung des Equalization Gain im Level Modus.

einem zeitlichen Schwellwert versehen werden. Ein Zeitpunkt in einem Signal kann somit als kritisch identifiziert werden, wenn das Maskierungsverhältnis länger als der zu spezifizierende Schwellwert positiv ist.

Weiterhin ist der Audioexport derzeit nur bei der vollständigen Ausführung des Skripts möglich. Die zeitsynchronisierten Audiodaten könnten in einer späteren Implementierung an den Knoten der Operation innerhalb der Datenbank gebunden werden. Somit ist nach Ausführung des Skripts ein Export der Daten für einen Hörversuch möglich. Denkbar wäre auch ein benutzerdefinierter Export einer bestimmten Sektion der Audiodaten mittels weiterer Zeitcursors und eines Funktionsaufrufs in der Property Page der Operation. Der Nachteil ist, dass für lange Eingangssignale auch die Datenbank entsprechend groß wird. Die Größe der Datenbank nimmt in der aktuellen Implementierung für ca. 6 min analysiertes Audiomaterial bereits rund 700 MB Speicherplatz in Anspruch, bei einer Abtastrate von $f_s = 44,1$ kHz und 16 Bit Wortbreite. Ein zusätzliches Speichern der synchronisierten Zeitsignale steigert den nötigen Festplattenspeicher.

5 Konzept zur Suche kritischer Stimuli

In diesem Kapitel wird ein Konzept zur Suche kritischer Signale vorgestellt. Der Fokus dieses Abschnitts liegt auf der Implementierung einer Automationsroutine, die für eine Sammlung von digitalem Audiomaterial und den Daten beliebig vieler DuTs alle gewünschten Kombinationen auralisiert und perzeptiv evaluiert. Das Ziel ist mittels der Automationsroutine für beliebige Treiber kritische Stimuli finden zu können, welche die nichtlinearen Verzerrungen besonders gut hörbar machen und somit auch für einen Hörversuch geeignet sind. Die Anwendung ist für die Verarbeitung von komplexen Stimuli gedacht, wie z.B. Musik oder Sprache.

5.1 Prinzip

Ein Stimulus ist nur dann für die Bewertung nichtlinearer Verzerrungen interessant, wenn seine Wiedergabe die relevanten Verzerrungen aktivieren kann. Im Falle der eingangs vorgestellten dominanten Kraftfaktor-, Steifigkeits- und Induktivitätsverzerrungen geschieht dies bei einer ausreichend hohen Auslenkung *x*. Um diese zu erreichen muss neben dem richtigen Verstärkungsfaktor auch die spektrale Beschaffenheit des Stimulus anwendungsgerecht sein.

Beispielsweise besteht für elektrodynamische Lausprecher für die Tief- und Mittentonwiedergabe der Anspruch, dass ein Mindestmaß an Tieftonanteilen im Stimulus enthalten ist. Ist dieser nicht vorhanden lässt sich mit ausreichend Verstärkung zwar die gewünschte Auslenkung erreichen, jedoch führt der Temperaturanstieg der Schwingspule dann zur Zerstörung des Treibers.



Abbildung 28: Differenztemperaturen (links) und Auslenkung (rechts) für einen Woofer.

Verwendete Stimuli: 100 Hz (blau) mit 18 dBV Verstärkung, 1000 Hz (rot) mit 21 dBV Verstärkung.

Abbildung 28 verdeutlicht den Zusammenhang: Ein Woofer wurde hierfür mit jeweils einem Sinuston bei 100 Hz (blau) und 1000 Hz (rot) auralisiert. Die Differenz der Verstärkung beträgt 3 dBV. Während die Spitzenauslenkung für den Basston rund 3,7 mm beträgt, wird für den Stimmton bei 1 kHz nur knapp 0,1 mm erreicht. Die Differenztemperaturen ΔT_{ν} zeigen die Erwärmung der Schwingspule im Vergleich zum Ruhezustand über der Länge des reproduzierten Signals. Deutlich ersichtlich ist die stärkere Erwärmung bei der Wiedergabe des Stimmtones. Die geringere Auslenkung erzeugt hier eine geringere Konvektionskühlung. Nach 30 Sekunden beträgt der Abstand der Schwingspulentemperaturen bereits rund 60 K.

Für höhere Verstärkungen können somit schnell Temperaturen erreicht werden, die zur Zerstörung der Schwingspule führen können.

Entsprechend darf bei der Aussteuerung eines DuTs zur Ermittlung der nötigen Verstärkung hinsichtlich der Anregung nichtlinearer Verzerrungen sowohl der Temperaturanstieg der Schwingspule als auch der Magnetanordnung nicht vernachlässigt werden.

Ist der notwendige Verstärkungsfaktor für eine Kombination aus Stimulus und Treiber bekannt, kann die Auralisation durchgeführt werden. Für die Betrachtung der nichtlinearen Verzerrungen auf Treiberebene dürfen dabei keine Einflüsse aus der Einbausituation in Gehäuse oder Raumeinflüsse in die Auralisation einfließen, es sei denn dies ist von explizitem Interesse.

Um in der perzeptiven Analyse ausschließlich die Einflüsse der nichtlinearen Verzerrungen betrachten zu können, muss die Auralisation des Stimulus im Kleinsignalbereich mit der Auralisation im Großsignalbereich pegelkompensiert und zeitsynchron verglichen werden. Mit diesen Bedingungen enthält das Differenzsignal die regulären und irregulären nichtlinearen Verzerrungen. (vgl. Liebig, 2012, S. 24)

Ausgehend von diesen Überlegungen wurde für das Klippel R&D-System eine Automationsroutine im MAT2-Format implementiert. Diese verknüpft verschiedene Module des R&D-Systems mittels des Klippel Automation Interface um zeiteffizient Audiomaterial auf beliebig vielen Treibern auralisieren und perzeptiv analysieren zu können.

Die Vorteile eines solchen Verfahrens sind vielfältig; die wichtigsten Punkten sind:

- Das Verfahren läuft nach der Konfiguration voll automatisiert und erfordert keine Überwachung.
- Neben Treiber- und Audiodaten ist nur die Software dB-Lab und ein Computer notwendig.
- Speziell präparierte Messräume sind nicht notwendig.
- Es entfallen notwendige Kühl- und Ruhephasen der DuTs, die bei der tatsächlichen Wiedergabe an der Leistungsgrenze notwendig wären.

Als notwendige Voraussetzungen müssen gemessene oder synthetisch generierte Daten der DuTs vorliegen.

Nachfolgend werden die Funktionsblöcke der Routine referenziert zum R&D-System beschrieben.

5.2 Ablauf und Funktionsblöcke

Die Routine wird als MAT2-Operation in dB-Lab aufgerufen. Der Nutzer muss der Operation mindestens den Pfad zu den Treiber- und Audiodateien übergeben. Weiterhin ist eine Auswahl der zu simulierenden Nichtlinearitäten zu treffen. Die Routine legt für jeden Treiber Operationsvorlagen an, von denen aus anschließend die tatsächlichen Simulationen und Auralisationen durchgeführt werden. Der Ablauf erfolgt dabei in einer zweifach verschachtelten *for*-Schleife. Der Index der äußeren Schleife läuft über alle Stimuli, während der Index der inneren Schleife jeweils über alle Treiberdateien läuft. Exemplarisch gibt es für *i* Stimuli und *j* Treiber somit für jede *i*-te äußere Schleife *j* innere Schleifen zu berechnen.



Abbildung 29 fasst den Ablauf unterteilt in die wichtigsten Funktionsblöcke übersichtlich zusammen.

Abbildung 29: Funktioneller Ablauf der Automationsroutine.

Aussteuerung

Zunächst ist die korrekte Aussteuerung für jede Kombination von Stimuli und Treibern zu ermitteln. Dazu wird das SIM-AUR Modul im thermischen Simulationsmodus genutzt. In mehreren Iterationen wird die Eingangsverstärkung G_{aur} angepasst, bis das Simulationsergebnis die Aussteuerungsgrenzen erreicht. Die Nutzung eines iterativen Verfahrens ist durch die nur schwer abschätzbare Dynamik der jeweiligen nichtlinearen Verzerrungskomponenten bedingt.

Die von Schulze implementierte Zeitraffertechnik kommt nicht zum Einsatz, da die Abweichungen für transiente Signale wie z.B. Musik hinsichtlich der thermischen Simulation erheblich sind, wenn die Anzahl der Stützstellen sehr gering ist. Diese Bedingung wäre jedoch notwendig um einen signifikanten Einfluss auf die benötigte Rechenzeit zu haben. Als Stützstellen werden dabei Zeitpunkte im Anregungssignal verstanden an denen die präzise Berechnung der Zustandsvariablen erfolgt. Die Stützstellen sind zum Teil äquidistant verteilt. Weitere Stützstellen werden nach einer ersten Lösung des Modells an Zeitpunkten gesetzt, an denen sich die Energie des Anregungssignals erheblich ändert. Entsprechend leidet die Genauigkeit der Berechnung bei einer geringen Anzahl an Stützstellen. (vgl. Schulze, 2014, S. 31 ff.)

Für die erste Iteration wird ausgehend von einer Startspannung U_{init} in V_{RMS} und dem RMS-Pegel L_{RMS} des jeweiligen Stimulus ein erster Verstärkungsfaktor $G_{aur,init}$ nach Gl. 24 abgeschätzt.

$$G_{aur,init} = 20 \cdot \log_{10} \left(\frac{U_{init} \mathbf{V}_{\text{RMS}}}{1 \mathbf{V}_{\text{RMS}}} \right) - L_{RMS} \mathbf{dB}$$
 Gl. 24

Entsprechend werden die Werte des Zeitsignals des Stimulus als Spannung interpretiert. Der Nutzer kann die vorgegebene Startspannung dem Treiber anpassen um die nötige Anzahl von Iterationen zu minimieren, beispielsweise mit den Ergebnissen des SIM2-Moduls.

Nach der ersten Iteration der thermischen Simulation wird das Erreichen der Aussteuerungsgrenzen überprüft. Dabei werden die Spitzenauslenkung x_{max} und die Änderung der Schwingspulentemperatur ΔT_V mit abgestufter Priorität betrachtet. Die höhere Priorität wird der Spitzenauslenkung beigemessen. Diese ist bedingt durch den Gültigkeitsbereich der

gemessenen bzw. simulierten Treiberdaten des LSI-Moduls oder des SIM2-Moduls, welche der Routine eingangs übergeben werden. Innerhalb einer jeden Treiberdatei wird diese durch die Größe x_{PSE} festgelegt. Die mittels der thermischen Simulation zu erreichende Auslenkung x_{max} wird mit der Größe x_{PSE} und dem Faktoren min_x und max_x innerhalb eines Bereichs nach Gl. 25 definiert:

$$\min_{x} \cdot x_{PSE} \le x_{max} \le \max_{x} \cdot x_{PSE}$$
Gl. 25

Die Faktoren min_x und max_x sind vom Nutzer anpassbar und bestimmen somit unmittelbar die Aussteuerung. Bei Verfehlung der Zielauslenkung in der ersten Iteration wird in den weiteren Iterationen eine Korrektur des Verstärkungsfaktors G_{aur} nach Gl. 26 vorgenommen:

$$G_{aur} = G_{aur} + 20 \cdot \log_{10} \left(\frac{x_{PSE}}{x_{max}} \right) dB$$
 Gl. 26

In jeder Iteration wird mit nachgeordneter Priorität die Änderung der Schwingspulentemperatur ΔT_v erfasst. Überschreitet diese die kritische Grenze $\Delta T_{V,krit}$ wird für die weiteren Iterationen die Regelung des Verstärkungsfaktors G_{aur} nicht mehr an der Auslenkung, sondern an der Änderung der Schwingspulentemperatur bemessen. Analog zu Gl. 26 erfolgt die Regelung dann nach Gl. 27:

$$G_{aur} = G_{aur} + 10 \cdot \log\left(\frac{\Delta T_{V,krit}}{\Delta T_V}\right)$$
Gl. 27

Die kritische Temperatur $\Delta T_{v,krit}$ kann vom Benutzer angepasst werden; die Zielvorgabe für die temperaturabhängige Pegelreduktion ist mit

$$0, 9 \cdot \Delta T_{V,krit} \le \max\{\Delta T_V\} \le 1, 1 \cdot \Delta T_{V,krit}$$
Gl. 28

zunächst festgelegt worden. Als kritische Grenze wurde eine maximale Änderung der Schwingspulentemperatur von $\Delta T_{V,krit} = 120$ K benutzt; dabei handelt es sich um einen Erfahrungswert. Sowohl für die auslenkungsabhängige als auch die temperaturabhängige Regelung wurde der Mindestbetrag einer notwendigen Pegelanpassung in Versuchen auf ±0,5 dB festgelegt.

Die Prüfung der thermischen Simulation hinsichtlich der erreichten Spitzenauslenkung und er der Differenz der Schwingspulentemperatur trägt dem in Abschnitt 5.1 beschriebenem Sachverhalt Rechnung. Wurde der Datensatz für den jeweiligen Treiber mit dem LSI-Modul erzeugt und der Messbereich bis an die Schutzparameter ausgereizt, so lässt sich mit der Automationsroutine der Betrieb bei hoher Last des Treibers für viele verschiedene Stimuli schnell und realitätsnah simulieren.

Die Implementierung ermöglicht das Erreichen des anwendungsgerechten Verstärkungsfaktors in den meisten Fällen innerhalb von maximal acht Iterationen. Wird die Startspannung günstig gewählt, so werden die gewünschten Grenzwerte zumeist innerhalb von maximal vier Iterationen erreicht. Die final ermittelten Werte für die Spitzenauslenkung x_{max} , die Differenztemperatur der Schwingspule ΔT_V und der Verstärkungsfaktor G_{aur} werden neben der jeweiligen Kombination aus Treiber und Stimulus, dem RMS-Pegel L_{RMS} und der Länge des Stimulus in einer Übersichtstabelle in dB-Lab dargestellt und gespeichert. Weiterhin wird dargestellt ob die Aussteuerungsgrenze von der Spitzenauslenkung oder der Differenztemperatur der Schwingspule bedingt ist.

Auralisation

Die Auralisation wird ebenfalls mit dem SIM-AUR-Modul realisiert. Abweichend zum Aussteuerungsblock wird das Modul hierfür im Auralisationsmodus genutzt. Für jede Kombination von Stimulus und Treiber wird mit den jeweils ermittelten Verstärkungsfaktoren und den ausgewählten Nichtlinearitäten die Lautsprecherwiedergabe auralisiert.

Im Rahmen des Forschungsinteresses dieser Arbeit sind etwaige Raumeinflüsse o.ä. zu minimieren. Entsprechend muss für die Auralisation wie auch für die thermische Simulation eine reflexionsarme Umgebung gewählt werden. Die Auralisation erfolgt daher mittels der Modellierung eines reflexionsarmen Halbraums. Weiterhin sind sowohl verschiedene Gehäusetypen und -prinzipien als auch das außeraxialeVerhalten des DUTS nicht von Interesse. Für Auralisation und thermische Simulation werden die DuTs innerhalb der Routine daher für den Einbau in einer Schallwand simuliert für einen Hörabstand von 1m.

Die digitale Repräsentation der auralisierten Schalldrucksignale liegt nach der Auralisation getrennt als linear und nichtlinear verzerrtes Signal vor. Der Verstärkungsfaktor S_{dis} skaliert die nichtlinearen Verzerrungen wie in Abschnitt 3.3.4 beschrieben. Dies bedeutet für die Auralisation, das für die Skalierung der nichtlinearen Verzerrungen $S_{dis} = 0$ dB bzw. $S_{dis} = 1$ angenommen werden muss, um dem tatsächlich vermessenen Lautsprecher zu entsprechen. Die Signale werden als Schalldrucksignale berechnet. Das linear verzerrte Signal entspricht hier dem Verhalten des Lautsprechers im Kleinsignalbereich und ist somit das Referenzsignal p_R für die anschließende perzeptive Analyse. Der Großsignalbereich ist demzufolge durch die Summe der linear und nichtlinear verzerrten Signale gegeben und dient als Testsignal p_T :

$$p_R = p_{lin} Gl. 29$$

$$p_T = p_{lin} + S_{dis} \cdot p_{nlin} = p_{lin} + 1 \cdot p_{nlin}$$
Gl. 30

Die auralisierten Signale werden innerhalb der Routine an das PEQ-Modul übergeben und ebenfalls als Audiodateien auf der Festplatte gespeichert.

Perzeptive Evaluation

Die perzeptive Evaluation jeder auralisierten Kombination wird mit dem PEQ-Modul ausgeführt. Der Verarbeitungsschritt zur Zeitsynchronisierung der Eingangssignale des Moduls kann übersprungen werden, da die Signale bereits zeitsynchron aus der Auralisation übernommen werden können.

Die Skalierung der Signale hat den Zielschalldruckpegel von 85 dB und orientiert sich an dem Vorschlag von Maier für Abhörpegel in Tonproduktionsumgebungen (vgl. Maier, 2008, S. 308). Dieser Bezugspegel ist für kritische Bewertungen von Klängen während der Audioproduktion vorgesehen und wurde daher für die pezeptive Evaluation ausgewählt. Abweichend von Maier wurde der Pegel jedoch ohne zusätzliche A-Bewertung benutzt.

Die Berechnung des Equalization Gains G_L erfolgt im Level Mode und orientiert sich dementsprehend an den RMS-Pegeln von Referenz- und Testsignal. Grundsätzlich stellt der Loudness Mode durch den Lautheitsbezug die bessere Variante hinhsichtlich der gehörgerechten Pegelkompensation dar. Jedoch ist die Rechenzeit für den Loudness Mode länger als für den Level Mode, weshalb letztere gewählt wurde. Die entstehenden Ungenauigkeiten sind für die Anwendung zur Suche kritischer Stimuli mit den gegebenen Einschränkungen aus Abschnitt 4.5 und 4.6 als vernachlässgbar klein anzusehen.

Zusätzlich zu den Ausgaben der vorangegangen Blöcke werden dem Nutzer innerhalb der Übersichtstabelle folgende MOVs des PEQ-Moduls angezeigt:

- DMR_{tot} und der Maximalwert des Maskierungsverhältnis
- Extrema der Rauigkeitsänderung zwischen Referenz- und Testsignal
- Extrema der Schärfeänderung zwischen Referenz- und Testsignal

Alle detaillierten Ergebnisse jeder Teilopereation stehen innerhalb dB-Lab in den jeweiligen Operationen zur Ansicht bereit. Die Datenbank ist dabei nach der Struktur "Datenbank\Stimulus\DuT\Operation" aufgebaut.

5.3 Anwendungsbeispiel

Für die Durchführung eines Onlinehörversuchs wurde eine Auswahl von 60 Musiktiteln verschiedener Genres mit drei verschiedenen Treibern mit der Automationsroutine untersucht. Das Ziel bestand in der Identifizierung der für den jeweiligen Treiber kritischen Stimuli. Die Untersuchung wurde zunächst mit allen dominanten Nichtlinearitäten aktiviert durchgeführt. Die im Anschluss als kritisch identifizierten Stimuli wurden auch für jeweils separat aktive Bl(x)-, $K_{ms}(x)$ und $L_e(x)$ -Verzerrungen verarbeitet.

5.3.1 Auswahl der Stimuli

Die Auswahl der Stimuli erfolgte manuell nach folgenden Gesichtspunkten:

- Angemessener Bassanteil, klar identifizierbare Stimmen und Melodieinstrumente im Mitteltonbereich. Intermodulationsverzerrungen sind dann besonders hörbar, wenn Basstöne klar verfolgbare Stimmtöne bei der Lautsprecherwiedergabe amplitudenmodulieren.
- Moderate Verzerrungen. Stimuli die bereits hohe Verzerrungsanteile enthalten maskieren die durch den Lautsprecher hervorgerufenen Verzerrungen, z.B. stark verzerrte E-Gitarrenmusik.
- Akustische und synthetische Klänge sollen in der Auswahl enthalten sein.
- Eine Breite Aufstellung an Genres enthält verschiedene genretypische Dynamikumfänge und bildet damit unterschiedlichste Lastfälle der DuTs ab.

Auf dieser Basis wurden aus einer privaten Musikbibliothek 60 Titel ausgewählt. Die vollständige Liste kann Anhang 1 entnommen werden. Die Auswahl ist exemplarisch zu verstehen und daher sind auch die als kritisch identifizierten Titel nur im Kontext dieser begrenzten Auswahl zu betrachten. Eine Erweiterung der Auswahl führt entsprechend unweigerlich zu anderen Ergebnissen.

Die Titel lagen im Stereoformat vor und wurden für die Verarbeitung monofon summiert, da auch die Auralisation monaural erfolgt.

5.3.2 Auswahl der Treiber

Die Auswahl der Treiber erfolgte mittels Messungen mit dem LPM- und LSI-Modulen um einerseits verschiedene Verzerrungskomponenten besonders herauszustellen und anderseits typische Anwendungsfälle darzustellen. Letztlich wurden zwei Konuslautsprecher und ein Miniaturlautsprecher benutzt um die Anwendung der Automationsroutine darzustellen. Die Besonderheiten der Treiber werden nachfolgend dargestellt.

Speaker#001

Der Lautsprecher mit der Kennung Speaker#001 ist ein Konuslautsprecher mit ca. 5 Zoll Korbdurchmesser und einer Nennimpedanz von 2 Ω .



Abbildung 30: Front- und Rückansicht, Speaker#001

Der Amplitudenfrequenzgang ist in Abbildung 31 dargestellt und lässt auf primäre Anwendung des Treibers als Tiefmittentöner schließen, der Breitbandbetrieb ist jedoch nicht ausgeschlossen.



Abbildung 31: Amplitudenfrequenzgang, Speaker#001

 $U_l = 3, 5 \text{ V}, \text{Sweep } f_l = 30...16000 \text{ Hz}, \text{ simuliert mit SIM2}, \text{ aktive Nichtlinearitäten: } Bl(x), K_{ms}(x), L_e(x).$

Der Treiber ist für einen Hörversuch besonders interessant, da eine stark nichtlineare Bl(x)-Kennlinie vorliegt, welche den Anteil der nichtlinearen Verzerrungen im Ausgangssignal dominiert:



Abbildung 32: Bl(x), Speaker#001. Vermessen mit LSI.

Entsprechend sind Intermodulationsverzerrungen über das gesamte Audioband zu verzeichnen. Abbildung 33 stellt die kumulierten Modulationsverzerrungen gemäß IEC 60268 dar. Der charakteristisch konstante Verlauf für Bl(x)-Verzerrungen bei Stimmtonsweeps wie in Abschnitt 2.3.2 beschrieben ist hier gegeben. Der leichte Anstieg zu hohen Frequenzen ist auf eine moderate $L_e(x)$ -Nichtlinearität zurückzuführen, deren Einfluss für diesen Treiber aber sehr gering ist. Der Abfall zu den tiefen Frequenzen ist durch die eingeschränkte Tieftonwiedergabe bedingt. Der erzeugte Gleichanteil für die dargestellte Anregung ist vernachlässigbar klein; der Treiber ist vollständig stabil.



Abbildung 33: Kumulierte Modulationsverzerrung, Speaker#001.

 U_l , $U_2 = 2,8$ V, Sweep $f_l = 200...10000$ Hz, $f_2 = 20$ Hz. Simuliert mit SIM2, aktive Nichtlinearitäten: Bl(x), $K_{ms}(x)$, $L_e(x)$.

Der vollständige Messreport zu diesem Lautsprecher ist dem digitalen Anhang zu entnehmen.
Speaker#002

Speaker#002 ist ebenfalls ein Konuslautsprecher mit ca. 5 Zoll Korbdurchmesser. Die Nennimpedanz beträgt 6,8 Ω .



Abbildung 34: Front- und Rückansicht, Speaker#002

Dieser Lautsprecher weist gegenüber dem ersten DuT eine verbesserte Tieftonwiedergabe auf. Im Vergleich ist die Höhenwiedergabe bei 10 kHz bereits um rund 8 dB vermindert. Der Treiber ist somit eingeschränkt für breitbandige Signale geeignet. Abbildung 35 stellt den Amplitudenfrequenzgang dar:



Abbildung 35: Amplitudenfrequenzgang, Speaker#002.

 $U_1 = 6.5 \text{ V}$, Sweep $f_1 = 30...16000 \text{ Hz}$. Simuliert mit SIM2, aktive Nichtlinearitäten: Bl(x), $K_{ms}(x)$, $L_{\theta}(x)$.

Der Lautsprecher ist als Testobjekt relevant, da sowohl stark nichtlineare Induktivitätsverzerrungen als auch stark nichtlineare Steifigkeitsverzerrungen vorliegen. Die Kraftfaktorverzerrungen fallen im Vergleich zu Speaker#001 wesentlich geringer aus, da der Lautsprecher mit der in Abschnitt 2.3.2 beschriebenen Überhangkonfiguration versehen ist. Die für die untersuchten Anregungen auftretenden Gleichanteile der Auslenkung sind vernachlässigbar gering; der Treiber ist stabil.

Abbildung 36 zeigt die $L_e(x)$ -Kennlinie und Abbildung 37 stellt analog zu Abbildung 33 die kumulierten Modulationsverzerrungen des Lautsprechers dar. Deutlich erkennbar ist der über der Frequenz ansteigende Anteil der Modulationsverzerrungen, welcher charakteristisch für dominante Induktivitätsverzerrungen bei Stimmtonsweeps ist.



Abbildung 37: Kumulierte Modulationsverzerrungen, Speaker#002.

 U_1 , $U_2 = 5,5$ V, Sweep $f_1 = 200...10000$ Hz, $f_2 = 20$ Hz. Simuliert mit SIM2, aktive Nichtlinearitäten: Bl(x), $K_{ms}(x)$, $L_e(x)$.

Die Steifigkeitskennlinie ist auf der nächsten Seite zusammen mit den verursachten harmonischen Verzerrungen dargestellt.



Abbildung 39: Relative harmonische Verzerrungen, Speaker#002.

Nur harmonische Verzerrungen zweiter und dritter Ordnung dargestellt. $U_1 = 6,5$ V, Sweep $f_1 = 30...16000$ Hz. Simuliert mit SIM2, aktive Nichtlinearitäten: Bl(x), $K_{ms}(x)$, $L_e(x)$.

Der vollständige Messreport liegt als digitaler Anhang vor.

Microspeaker

Der Datensatz für den *Microspeaker* stammt aus der Beispielsammlung der Klippel GmbH. Die Klasse der Microspeaker beinhaltet Miniaturlausprecher die typischer Weise in Mobiltelefonen und kleinsten Multimediaabspielgeräten eingesetzt werden. Das Wirkprinzip ist identisch, jedoch ergeben sich durch die geringen Baugrößen Änderungen in der Geometrie. Beispielsweise kann zu Gunsten des zur Verfügung stehenden Bauraums auf eine separate Zentrierspinne verzichtet werden. Die zugehörige Funktion wird dann von der Aufhängung und Membran erfüllt. In der Folge entstehen starke Asymmetrien einzelner Kennlinien. Überdies sind nur kleine Auslenkungen möglich. Weiterhin schränkt die geringe Membranfläche die Tieftonwiedergabe und den erreichbaren Schalldruckpegel ein. Die geringen Massen führen zu hohen Resonanzfrequenzen. Von den in dieser Arbeit betrachteten Nichtlinearitäten ist die $L_e(x)$ -Verzerrung für diesen Microspeaker nicht besonders relevant.

Der gemessene Treiber ist neben der veränderten Bauart als Testobjekt interessant, da sowohl die Bl(x)-Kennlinie als auch die $K_{ms}(x)$ -Kennlinie stark nichtlinear und asymmetrisch sind. Ihre Verläufe sind in Abbildung 40 und Abbildung 41 und der zugehörige Amplitudenfrequenzgang in Abbildung 42 dargestellt:



Abbildung 40: Bl(x), Microspeaker. Vermessen mit LSI.



Abbildung 41: $K_{ms}(x)$, Microspeaker. Vermessen mit LSI.



Abbildung 42: Amplitudenfrequenzgang, Microspeaker.

 $U_l = 1,7 \text{ V}$, Sweep $f_l = 30...16000 \text{ Hz}$. Simuliert mit SIM2, aktive Nichtlinearitäten: Bl(x), $K_{ms}(x)$, $L_e(x)$.

Die Resonanzfrequenz des Treibers beträgt $f_s = 595$ Hz. Unterhalb der Resonanzfrequenz treten starke harmonische Verzerrungen auf Grund der nichtlinearen Steifigkeitskennlinie auf:



Abbildung 43: Relative harmonische Verzerrungen, Microspeaker.

Nur harmonische Verzerrungen zweiter und dritter Ordnung dargestellt. $U_I = 1,7$ V, Sweep $f_I = 30...16000$ Hz. Simuliert mit SIM2, aktive Nichtlinearitäten: Bl(x), $K_{ms}(x)$, $L_e(x)$.

Die kumulierten Modulationsverzerrungen in Abbildung 44 werden durch den nichtlinearen Kraftfaktor dominiert und weisen entsprechend einen weitestgehend konstanten Verlauf oberhalb von 2 kHz auf. Zu tiefen Frequenzen ist ein Anstieg zu verzeichnen, der bedingt durch zusätzliche Intermodulationen mit den harmonischen Verzerrungen der nichtlinearen Steifigkeit entsteht.



Abbildung 44: Kumulierte Modulationsverzerrungen, Microspeaker.

 U_l , $U_2 = 1,1$ V, Sweep $f_l = 600...10000$ Hz, $f_2 = 100$ Hz. Simuliert mit SIM2, aktive Nichtlinearitäten: Bl(x), K_{ms}(x), L_e(x).

Der Treiber ist für die untersuchten Anwendungen stabil. Der vollständige Messbericht liegt als digitaler Anhang vor.

Um eine anwendungsgerechte Untersuchung zu ermöglichen, wurden die Musikbeispiele für den Microspeaker mit einem Hochpassfilter zweiter Ordnung nach Butterworth mit der Grenzfrequenz $f_g = 250$ Hz gefiltert.

5.3.3 Benötigte Rechenzeit

Zur Verdeutlichung der benötigten Rechenzeit sind nachfolgend für das gesamte verarbeitete Audiomaterial der Auswahl aus Anhang 1 auf allen drei Treibern einige Kennwerte angeführt.

Die Gesamtspieldauer des Programmmaterials betrug ca. 4,25 h. Die durchschnittliche Länge je Stimulus betrug rund 255 s. Tabelle 4 stellt die durchschnittliche benötigte Rechenzeit der Teiloperationen aus Abschnitt 5.2 als Funktion der individuellen Signallänge t_{sig} dar:

| Teiloperation | \emptyset Rechenzeit $f(t_{sig})$ |
|-----------------------------|-------------------------------------|
| Berechnung G _{aur} | $1,12 \cdot t_{sig}$ |
| Auralisation | $0,45 \cdot t_{sig}$ |
| Perzeptive Evaluation | $4,17 \cdot t_{sig}$ |

Tabelle 4: Durchschnittliche Rechenzeit als Funktion der Signallänge.

Wie bereits erläutert ist der Zeitaufwand zur Berechnung des Verstärkungsfaktors G_{aur} hauptsächlich von der gewählten Startspannung U_{init} und der daraus resultierenden Anzahl an Iterationen zur Erreichung der Aussteuerungsgrenzen abhängig. Die Auralisation erfolgt durchschnittlich in der Hälfte der tatsächlichen Signallänge. Die Perzeptive Evaluation ist der langsamste Teilschritt und benötigt die rund vierfache Signallänge t_{sig} .

Die Gesamtrechendauer für die Verarbeitung der 60 Musiktitel mit den drei ausgewählten Treibern lag bei rund 73 h. Dabei waren alle in dieser Arbeit betrachteten Nichtlinearitäten aktiviert, mit Ausnahme der Le(i)-Verzerrung. Diese war zum Zeitpunkt der Durchführung nicht stabil im SIM-AUR-Modul implementiert.

5.3.4 Auswertung der Daten

Die berechneten Daten des Anwendungsbeispiels wurden für jedes DuT absteigend nach der Größe der Total Distortion To Mask Ratio DMR_{tot} sortiert, um für jeden Treiber die kritischen Stimuli der Auswahl identifizieren zu können. Dieser Ansatz wurde gewählt, da sich keine direkten Korrelationen zwischen den Maxima der Schärfe- und Rauigkeitsänderungen und den Maskierungsmaßen feststellen ließ. Eine Sortierung nach dem Maximum des DMR-Verlaufs DMR_{max} erwies sich nicht als sinnvoll, da die Maxima zumeist nur für einen Block erreicht wurden. Die Maskierungsverhältnisse werden im FFT-Modell berechnet und bei einer Blocklänge von 2048 Samples ist ein Block bei einer Abtastrate von $f_s = 44,1$ kHz rund 0,046 Sekunden lang. Dies erschien für die Verwendung in einem Hörversuch zu kurz.

In Tabelle 5, Tabelle 6 und Tabelle 7 sind die Ergebnisse mit der geringsten Maskierung für die jeweiligen Kombinationen aus Stimulus und Treiber auszugsweise dargestellt. Als geringe Maskierung verstehen sich Werte, die nahe 0 dB oder darüber liegen. Die vollständigen Namen der Künstler und Titel können Anhang 1 entnommen werden. Diese Ergebnisse wurden genutzt um Testsignale für den in Abschnitt 6 beschriebenen Hörversuch zu generieren.

| Stim | ulus (Index Künstler - Titel) | DMR _{tot} | DMR_{max} |
|------|-------------------------------------|--------------------|-------------|
| 9 | Boundzound - So Long | -6,7 dB | 14,3 dB |
| 56 | Skrillex - Summit | -8,1 dB | 7,4 dB |
| 38 | Kelis - Trick me | -12,1 dB | -1,7 dB |
| 44 | Massive Attack - Splitting the Atom | -12,6 dB | 8,3 dB |
| 54 | Rudimental - Hide | -14,2 dB | 3,7 dB |

Tabelle 5: Ergebnisse Speaker#001. Aktiv: Bl(x), $K_{ms}(x)$, $L_e(x)$.

| Tabelle 6: | Ergebnisse | Speaker#002. | Aktiv: | Bl(x). | $K_{\dots}(x)$. | L(x) |
|------------|------------|--------------|-----------|----------------|------------------|---------------------------|
| rubene o. | Ligeomose | Speaker#002. | I IIII V. | $Di(\Lambda),$ | $m_{ms}(n)$ | $\mathbf{L}_{e}(\Lambda)$ |

| Stim | ulus (Index Künstler - Titel) | DMR _{tot} | DMR_{max} |
|------|-------------------------------------|--------------------|-------------|
| 56 | Skrillex - Summit | -11,2 dB | 5 dB |
| 9 | Boundzound - So Long | -12,5 dB | 7,8 dB |
| 44 | Massive Attack - Splitting the Atom | -15,9 dB | 4 dB |
| 38 | Kelis - Trick me | -17,4 dB | -3,7 dB |
| 29 | Glashaus - Haltet die Welt an | -17,9 dB | 1,8 dB |
| | | | |

| Stimulus (Index Künstler - Titel) | | DMR _{tot} | DMR _{max} |
|-----------------------------------|----------------------------------|--------------------|--------------------|
| 4 | Bassida - You make me wanna stay | -14,1 dB | 5,5 dB |
| 37 | Kelis - Stick Up | -18 dB | -0,5 dB |
| 5 | Bassida - Be mine | -18,2 dB | 1,9 dB |
| 1 | Avicii - Addicted To You | -20,2 dB | 8 dB |
| 43 | Marcus Miller - Free | -20,4 dB | 0,1 dB |
| | | | |

Tabelle 7: Ergebnisse Microspeaker. Aktiv: Bl(x), $K_{ms}(x)$, $L_{e}(x)$.

Die Anregungssignale wurden mit einem Butterworth-Hochpassfilter zweiter Ordnung bei $f_g = 250$ Hz gefiltert.

Ausgehend von dem Gesamtmaskierungsverhältnis und dem zugehörigen Maximum wurden für Speaker#001 und Speaker#002 die Titel 9 und 56 für die Verwendung in einem Hörversuch gewählt. Für den Microspeaker wurden Titel 4 und 5 gewählt, da nach Abhören des auralisierten Signals und Betrachtung der Zeitverläufe des *DMR* Titel 5 eine bessere Eignung für einen Hörversuch als Titel 37 aufwies. Dieser Umstand begründet sich in der eingeschränkten Aussagekraft der Maskierungsmaße, die schon in Abschnitt 4.5 besprochen wurde.

Weiterhin wurde für alle drei Treiber Titel 1 gewählt, um zwischen den sehr unterschiedlichen Treibern anhand eines gemeinsamen Stimulus Vergleiche ziehen zu können. Die Analyse ergab für Speaker#001 $DMR_{tot} = -18,3$ dB und $DMR_{max} = -1,3$ dB und für Speaker#002 $DMR_{tot} = -24,6$ dB und $DMR_{max} = -3,7$ dB. Ausgehend von der Betrachtung der Zeitverläufe des DMR und den Daten von Liebig scheint dieser Stimulus daher noch hinreichend geeignet, da eine Detektion bis zu einem Wert von $DMR_{tot} = -20$ dB erfolgen kann (vgl. 2012, S. 79).

Die ausgewählten Stimuli wurden für die Verwendung im Hörversuch anschließend für jeweils separat aktive Bl(x)-, $K_{ms}(x)$ - und $L_e(x)$ -Verzerrungen mit der Automationsroutine auralisiert und analysiert.

5.4 Ausblick

Die vorgenommene Implementierung muss weiter optimiert werden. Die Rechenzeit schwankt stark abhängig von der Länge und Komplexität der Stimuli, der Anzahl der aktivierten Nichtlinearitäten und der Hardwarekonfiguration des benutzten Computers. Der Großteil der Rechenzeit entfällt dabei auf die perzeptive Evaluation durch das PEQ-Modul. Der bestehende Optimierungsbedarf zu diesem Modul wurde bereits in Abschnitt 4.8 besprochen. Wenn der benötigte Zeitaufwand des PEQ-Moduls in einer zukünftigen Version halbiert werden, kann sinkt die Gesamtrechendauer der Automationsroutine für das angegebene Beispiel auf ca. 47 h. Weitere Verbesserungen ließen sich durch eine adaptive Version der Zeitraffertechnik des SIM-AUR-Moduls erzielen. Wenn die in Abschnitt 5.2 umrissene Verteilung der Stützstellen nicht-äquidistant auf Basis einer vorhergehenden Analyse des Anregungssignals erfolgt kann eine mehrfache Lösung des Lautsprechermodells u.U. vermieden werden und die Genauigkeit der Methode erhöht werden.

Zudem kann die Automationsroutine derzeit nicht unterbrochen werden. Dies stellt für große zu verarbeitende Datenmengen ein Problem dar, da der betreffende Computer nicht abgeschaltet werden kann bevor die Berechnungen beendet wurden. Zwar wurde eine Backup-Funktionalität implementiert, die alle berechneten Ergebnisse nach jeder *j*-ten inneren

Schleife auf dem Hauptspeicher sichert, jedoch geht bei einem Programmabbruch der Inhalt der Übersichtstabelle innerhalb von dB-Lab verloren und kann nicht wiederhergestellt werden. Eine zukünftige Version sollte daher zunächst in getrennten Operationen alle Verstärkungsfaktoren berechnen und die jeweiligen Auralisationen erzeugen, bevor eine weitere Operation unter Nutzung des *Batch Run* Modus von dB-Lab die perzeptive Evaluation erledigt. Eine finale Operation kann dann die Ergebnisse zusammentragen und darstellen. Ein Vorteil dieser Änderung wäre die durch den Batch Run Modus gegebene Möglichkeit des Pausierens. Der gewichtigste Nachteil ist, dass die Routine dann zu bestimmten Schlüsselzeitpunkten Benutzereingaben erfordert und nicht vollständig ohne Überwachung ablaufen kann. Eine Wahlmöglichkeit zwischen dem bisher implementierten Ablauf und der zuletzt vorgeschlagenen Änderung ist denkbar.

Eine weitere Einschränkung stellt die Handhabung der Treiberdaten dar. Im Moment müssen die gemessenen oder simulierten Daten aus den LPM-, LSI- oder SIM2-Operationen in eine Textdatei exportiert werden. In einer zukünftigen Version müssen die betreffenden Daten direkt aus den Quelloperationen eingelesen werden können. Die Integration innerhalb des R&D-Systems und dB-Lab wird somit erhöht und der fehleranfällige Umweg über Textdateien vermieden.

Die vorgeschlagenen Optimierungen wurden im Verlauf dieser Arbeit nicht weiter verfolgt da verschiedene Aktualisierungen des Klippel Automation Interface und des SIM-AUR Moduls eine grundlegend neue Implementierung der Routine erfordert hätten. Dies war innerhalb des gegebenen zeitlichen Rahmens nicht möglich.

Zusätzlich zu der vorgestellten Routine ist der Aufbau einer Datenbank aus verarbeiteten Musikstücken und Treibern wünschenswert. Innerhalb der Datenbank sollten dabei Signaleigenschaften der Stimuli wie der Verlauf des Effektivwertes oder die spektrale Zusammensetzung und ihre zeitliche Änderung erfasst werden. Ebenfalls müssen wichtige Treiberparameter der linearen und nichtlinearen Zustände wie z.B. der Betrag des Kraftfaktors in Ruhelage und Verlauf der Bl(x)-Kennlinie erfasst werden. Gemeinsam mit stets zu aktualisierenden Erkenntnissen aus der perzeptiven Analyse kann ein trainierter Klassifikationsalgorithmus somit für ein neues DuT wahrscheinlich kritische Stimuli aus der Datenbank vorschlagen. Die Vorschläge können dann mit der aktualisierten Automationsroutine verarbeitet werden.

6 Erweiterung des Hörversuchs

Der bisherige Hörversuch auf der Internetpräsenz der Klippel GmbH befasste sich mit der Detektionsschwelle nichtlinearer Verzerrungen für verschiedene Kombinationen aus Treibern und Stimuli. Darüber hinaus findet keine Untersuchung des Einflusses der Verzerrungen auf die Hörwahrnehmung oder den Klangeindruck statt. Das folgende Kapitel beschäftigt sich daher mit der Entwicklung und ersten Umsetzung einer Erweiterung des bestehenden Hörversuchs.

6.1 Forschungsinteresse

Die Hörbarkeit nichtlinearer Verzerrungen für simple Testsignale, wie z. B. einzelner Sinustöne oder Zweitonsignale, ist gegenüber komplexen Signalen häufig schon für geringe Intensitäten gegeben. Liebig führt das auf den quasi-stationären Charakter der Signale und den daraus resultierenden stationären Zeitverläufen der MOVs zurück (vgl. 2012, S. 79), in denen bereits kleine Variationen detektiert werden können. Der bisherige Hörversuch dient der Ermittlung dieser Detektionsschwelle und funktioniert prinzipiell auch für komplexe Signale. Letzteres ist für die Untersuchung von Lautsprechern relevant, da diese zumeist für die Wiedergabe komplexer Signale genutzt werden.

Komplexe Signale weisen jedoch selten einen stationären Charakter auf, entsprechend sind die Zeitverläufe der psychoakustischen MOVs ebenso unstet. Folglich liegen die Wahrnehmungsschwellen für komplexe Signale über denen von simplen Signalen. Dies bedeutet, dass es wahrscheinlich ist, dass für eine Auralisation eines komplexen Stimulus mit gleicher Verzerrungsintensität wie der Auralisation eines simplen Stimulus, die nichtlinearen Verzerrungen nicht detektiert werden können. Das wird bereits in Ansätzen in den Daten aus Liebigs nicht repräsentativem Hörversuch ersichtlich (vgl. 2012, S. 78 f.). Als gleiche Verzerrungsintensität versteht sich hierbei, dass der gleiche Treiber mit den gleichen aktiven Nichtlinearitäten und gleichen Grenzen der Aussteuerung zu nutzen ist. Müller machte die Detektionswahrscheinlichkeit davon abhängig, ob die spektrale Zusammensetzung des Musiksignals entsprechende Maskierungseffekte aktiviert (vgl. 2008, S. 1144). Zusätzlich stellte er heraus, dass die Hörbarkeit von Verzerrungen nicht zwangsläufig Aussagen hinsichtlich der Beeinträchtigung der Klangqualität zulässt (vgl. ebd., S. 1145).

Es bestehen nur Annahmen zum prinzipiellen Einfluss bestimmter nichtlinearer Verzerrungen auf die empfundene Klangqualität. Diese wurden aber bisher nicht quantitativ oder qualitativ verifiziert.

6.1.1 Konzept des Begriffs Klangqualität

Der Begriff Klangqualität ist vielfältig besetzt. Im Rahmen dieser Arbeit ist es nicht möglich diesen Begriff umfassend zu untersuchen und mit den in Abschnitt 6.3 vorgestellten Testverfahren abzubilden. Wie Fastl & Zwicker feststellten, ist es in Versuchen nur selten möglich Variablen wie ästhetische Vorlieben der Teilnehmer oder kognitive Effekte zu erfassen, die ihrerseits die Klangqualität beeinflussen können (vgl. 2007, S. 328).

Im Zusammenhang mit der Wiedergabe von Signalen durch elektrodynamische Lautsprecher ist der Begriff Klangqualität daher als Qualität der Wiedergabe zu verstehen. Diese sei maßgeblich von den Eigenschaften des Wiedergabesystems bestimmt und wird im Folgenden mit der Größe *SQ* beschrieben.

In Abbildung 45 wird zur Verdeutlichung der angenommene Verlauf der empfundenen Klangqualität für ein unterschiedlich nichtlinear verzerrtes Testsignal über der Skalierung der Verzerrung schematisch abgebildet. Zusätzlich ist die Detektionsrate dargestellt. Im Bereich weit links der verzeichneten Detektionsschwelle sind die jeweiligen nichtlinearen Verzerrungen nicht hörbar; weit rechts sind diese hingegen deutlich hörbar. Die eingezeichnete Detektionsschwelle kann natürlich für verschiedene Verzerrungsarten an verschiedenen Punkten liegen und ist hier zur Vereinfachung der Darstellung normiert worden. Die Darstellung illustriert die je nach Verzerrungstyp unterschiedlichen Einflüsse auf die Änderung der Klangqualität durch die verschiedenen Steigungen der Kurven. Grundsätzlich ist anzunehmen, dass ein gerichteter Zusammenhang zwischen der Intensität der nichtlinearen Verzerrungen und der Änderung der Klangqualität besteht. Der Begriff Verzerrung impliziert dabei, dass eine Erhöhung des Pegels der nichtlinearen Verzerrungen nichtlinearen Klangqualität führt. Die Überlegungen zu den Symptomen nichtlinearer Steifigkeit $K_{ms}(x)$ ermöglichen jedoch, dass für einen bestimmten Bereich um die Hörschwelle zunächst auch eine Zunahme der empfundenen Klangqualität möglich sein kann.

Der Einfluss nichtlinearer Verzerrungen kann somit als psychometrische Funktion der Änderung der Klangqualität SQ über der Änderung der Skalierung S_{dis} der nichtlinearen Verzerrungen verstanden werden.





Schematische Darstellung. Durchgehende Linie: $K_{ms}(x)$ -Verzerrung; gestrichelt: Le(x)-Verzerrung; gepunktet: Rub & Buzz.

Eine Untersuchung des Zusammenhangs ist vor allem für Hersteller von Lautsprechern interessant. Mit der Kenntnis des Einflusses nichtlinearer Verzerrungen auf die empfundene Klangqualität lassen sich anwendungsgerecht Zielvorgaben formulieren. Eine solche Zielvorgabe kann z.B. für einen Treiber für Unterhaltungselektronik bedeuten, dass der Anteil der nichtlinearen Steifigkeitsverzerrungen sich in einem Bereich nahe der potenziellen Hörschwelle der späteren Nutzergruppe bewegt, in dem die empfundene Klangqualität durch die Verzerrungen steigt. Analog zur vorherigen Darstellung illustriert Abbildung 46 dieses Beispiel. Denkbar wäre, dass ausgehend von der Auralisation eines Produktionsmusters oder eines digitalen Prototypen Daten erhoben werden, um die Zielperformance für die spätere Produktion festlegen zu können. Die Relevanz besteht in der Optimierung des Verhältnisses von Produktionskosten und Performance der Treiber.



Detektionsschweite Al

Abbildung 46: Beispiel zur Zielperformance.

Schematische Darstellung. Durchgehende Linie: $K_{ms}(x)$ -Verzerrung; schraffierter Bereich: Zielperformance des zu produzierenden Treibers.

Um gezielt den Einfluss der nichtlinearen Verzerrungen auf die Klangqualität des DuTs zu evaluieren bieten sich Paarvergleiche an. Dabei ist sicherzustellen, dass der Unterschied zwischen Referenz- und Testsignalen stets ausschließlich durch die nichtlinearen Verzerrungen gegeben ist. Dies verhält sich analog zu den in Kapitel 4 und 5 durchgeführten Analysen hinsichtlich der Eignung des PEQ-Moduls für die Analyse nichtlinearer Verzerrungen und der Identifikation kritischer Stimuli.

6.1.2 Reduktion multivariater Abhängigkeiten

Die tatsächliche Änderung der Klangqualität kann von einer Vielzahl von Variablen bedingt werden. Zu den wichtigsten zählen

- der Treiber (lineare Parameter, dominante Nichtlinearitäten),
- der Stimulus (Komplexität),
- die Intensität der nichtlinearen Verzerrungen (S_{dis}) ,
- die Abhörbedingungen (Geräte, Lautstärke, Raum, Störgeräusche),
- und die Hörerfahrung und -gewohnheiten des Hörers.

Die Merkmalsausprägungen von Treiber, Stimulus und den nichtlinearen Verzerrungen sind eines Hörversuchs manipulierbar. Die beim Entwurf Hörerfahrung und die Abhörbedingungen in einem Onlinehörversuch sind hingegen nur mit geringer Genauigkeit Weitere Einflüsse wie musikalische Präferenzen der Teilnehmer oder messbar. Trainingseffekte über mehrere Durchläufe eines Hörversuchs sind nur schlecht oder gar nicht abzubilden. Daher wurde die Erfassung aller Beziehungen und Wechselwirkungen zwischen den Variablen hinsichtlich des Einflusses auf die Klangqualität für einen einfachen Onlinehörversuch als nicht realisierbar eingestuft.

Das Interesse muss folglich auf aussagekräftige und überprüfbare Bereiche beschränkt werden. Eine genaue statistische Rekonstruktion der abgebildeten Verläufe mag zwar möglich

sein, hat aber nur mäßige Relevanz für die Definitionen von Produktionszielen. Zudem erfordert diese Rekonstruktion sehr umfangreiche Stichproben, wodurch die Gesamtversuchsdauer steigt.

Zunächst können Extrembereiche sehr weit oberhalb oder unterhalb der Detektionsschwelle AT vernachlässigt werden, da bei extremen Verzerrungen ohnehin eine geringe Klangqualität zu erwarten ist und für sehr niedrige Verzerrungspegel wahrscheinlich keine Detektion erfolgt. Hinsichtlich der Untersuchung von Lautsprechern erscheinen der Bereich um die Detektionsschwelle AT eines Probanden P_i und der Bereich nahe $S_{dis} = 0$ dB interessant. Die Bewertung der Klangqualität eines Treibers durch eine ausreichend große Stichprobe für diese Bereiche lässt sowohl Rückschlüsse über die Beeinflussung der Klangqualität durch den Treiber als auch Schlussfolgerungen über den tendenziellen Einfluss einer nichtlinearen Verzerrungen auf die empfundene Klangqualität für eine Grundgesamtheit zu, der die Stichprobe entstammt.

Mit einer geringen Anzahl von Messpunkten je Versuchsteilnehmer lassen sich die tatsächlichen Verläufe der in Abbildung 45 schematisch dargestellten Kurven jedoch nicht ermittelten. Vielmehr können Gradienten bestimmt werden, welche die Kurvenform approximieren. Ein solcher Gradient ΔSQ ist als Funktion anhängig vom Treiber DuT, Stimulus *Sig*, Nichtlinearität *Nl*_i und der Skalierung der Verzerrung *S*_{dis}:

$$\Delta SQ = f\left(S_{dis}, DuT, Sig, Nl_{j}\right)$$
Gl. 31

Folglich sind die Gradienten an den genannten Punkten

$$\Delta SQ_{P_i,AT} = f\left(AT\left[P_i\right], DuT, Sig, Nl_j\right)$$
Gl. 32

$$\Delta SQ_{P,DuT} = f(0, DuT, Sig, Nl_i)$$
Gl. 33

Mit Gl. 31 erfolgt eine Festlegung auf Parameter, die mit den vorgestellten Methoden voriger Kapitel eingestellt und kontrolliert werden können. Die eingangs als unsicher klassifizierten Variablen werden ausgeklammert.



Abbildung 47: Beispielgrafik zu Gradienten ΔSQ .

Dicke Pfeile: Gradienten $\Delta SQ_{P_r,AT}$ und $\Delta SQ_{P_r,DuT}$. Schraffiert: zu vernachlässigende Bereiche.

Für die gewählten Bereiche müssen demzufolge die Gradienten ΔSQ_{P_iAT} und ΔSQ_{P_iDuT} ermittelt werden. Abbildung 47 stellt die Gradienten und die nicht relevanten Bereiche schematisch am Beispiel der $K_{ms}(x)$ -Verzerrung dar.

Analog zum Beispiel aus Abbildung 45 ist mit Gl. 31 auch möglich, dass verschiedene nichtlineare Verzerrungsarten verschiedene Gradienten aufweisen können:



Abbildung 48: Verschiedene Gradienten ΔSQ .

Rot: $\Delta SQ_{P_i DuT}$ für $K_{ms}(x)$ -Verzerrungen, Grün: $\Delta SQ_{P_i DuT}$ für $L_e(x)$ -Verzerrungen, Blau: $\Delta SQ_{P_i DuT}$ für Rub & Buzz-Verzerrungen. Detektionsschwellen zur Vereinfachung normiert dargestellt.

Die Kontributionen einzelner nichtlinearer Verzerrungen relativ zu anderen kann durch das Verhältnis zwischen den jeweiligen Gradienten abgeschätzt werden. Nach Gl. 35 werden die ermittelten Werte zueinander in Beziehung gesetzt:

$$\frac{\Delta SQ_{j}}{\Delta SQ_{j+1}} = \frac{f\left(S_{dis}, DuT, Sig, Nl_{j}\right)}{f\left(S_{dis}, DuT, Sig, Nl_{j+1}\right)}$$
Gl. 34

Es kann entsprechend geprüft werden ob sich feste Verhältnisse für verschiedene Nichtlinearitäten einstellen.

6.1.3 Forschungshypothesen

Auf Basis der Erläuterungen und der dargestellten und neu implementierten Methoden der vorangegangenen Kapitel ist es daher das Ziel der Erweiterung des Hörversuchs, die Abhängigkeit der empfundenen Klangqualität von der Intensität der nichtlinearen Verzerrungen zu quantifizieren. Das Forschungsinteresse ist in folgende Hypothesen zusammengefasst:

1. Die Hörbarkeit nichtlinearer Verzerrungen beeinflusst die empfundene Klangqualität.

Es besteht ein gerichteter Zusammenhang zwischen der Veränderung der Klangqualität und der Hörschwelle. Je größer der Abstand des Pegels der nichtlinearen Verzerrung S_{dis} eines

präsentierten Signals zur Detektionsschwelle ist, desto schlechter fällt die Bewertung der Klangqualität tendenziell aus.

Nullhypothese: Es besteht kein Zusammenhang zwischen der Hörbarkeit nichtlinearer Verzerrungen und der bewerteten Klangqualität.

2. Die Gradienten ΔSQ einzelner Nichtlinearitäten unterscheiden sich.

Die Beeinflussung der Klangqualität erfolgt für verschiedene Nichtlinearitäten unterschiedlich stark. Dies äußert sich in verschiedenen Steigungen der Bewertungskurven analog zu Abbildung 48.

Nullhypothese: Es besteht kein Unterschied zwischen den Gradienten ΔSQ verschiedener Nichtlinearitäten.

3. Die Gradienten ΔSQ verschiedener Verzerrungsarten stehen in bestimmten Verhältnissen zueinander.

Aus den Verhältnissen der Gradienten zueinander lassen sich Rangfolgen bilden. Somit können in der Lautsprecherentwicklung verschiedene Nichtlinearitäten mit geordneten Prioritäten behandelt werden.

Nullhypothese: Es bestehen keine bestimmten Verhältnisse zwischen den Gradienten ΔSQ . Eine Rangfolge ist nicht feststellbar.

6.2 Anforderungen an das Testverfahren

Ausgehend vom Forschungsinteresse und dem vorgestellten Konzept der Klangqualität wurden folgende Anforderungen an ein neues Testverfahren gestellt:

- Für verschiedene DuTs sollen verschiedene nichtlineare Verzerrungen untersucht werden. Die Gradienten ΔSQ für die Bereiche um $S_{dis} = 0$ dB und $S_{dis} = AT[P_i]$ müssen mit zu erhebenden Daten abgeschätzt werden können.
- Die Ergebnisse verschiedener Kombination aus Treibern und Stimuli sollen vergleichbar sein.
- Das Testverfahren muss ohne besondere technische und räumliche Vorrausetzungen durchzuführen sein.
- Ein Versuchsdurchlauf muss in der k
 ürzest m
 öglichen Zeit durchzuf
 ühren sein um Erm
 üdung der Teilnehmer zu vermeiden und das Absolvieren mehrerer Durchl
 äufe zu erm
 öglichen.
- Die zu erreichende Stichprobe muss möglichst groß sein.
- Das Anlegen eines Versuchs und die Bereitstellung der erhobenen Daten soll einfach sein.

Mit diesen Maßgaben wurde das nachfolgende beschriebene Testverfahren als Erweiterung des bestehenden Hörversuchs entwickelt.

6.3 Testverfahren zur Bewertung der Klangqualität

6.3.1 Beschreibung

Dem Teilnehmer werden ein Referenzsignal und drei Testsignale als Audiobeispiele präsentiert. Das Referenzsignal ist das linear verzerrte Signal. Die Testsignale beinhalten das linear verzerrte Signal und die jeweiligen nichtlinearen Verzerrungen bei drei verschiedenen Werten des Skalierers S_{dis} (siehe Gl. 30). Die Reihenfolge der präsentierten Testsignale ist randomisiert. Das Verfahren beschränkt sich zunächst auf drei Testsignale um die Dauer eines Versuchs möglichst kurz zu halten.

Der Teilnehmer ist aufgefordert die Testsignale mit dem Referenzsignal zu vergleichen und hinsichtlich der Klangqualität zu bewerten. Die Klangqualität ist dabei auf einer diskreten Skala mit neun Schritten anzugeben. Den Skalenstufen sind die in Tabelle 8 angegebenen qualitätsbeschreibenden Adjektive zugeordnet. Die neunstufige Skala wurde hier gegenüber den meist gebräuchlichen fünfstufigen Skalen bevorzugt, da eine bessere Auflösung der Ergebnisse zu erwarten ist. Die beschreibenden Adjektive wurden in Anlehnung an die kontinuierliche Skala der ITU zur Bewertung der mittleren Qualität von Audiosystemen ausgewählt (vgl. ITU, 2015, S. 9).

| Stufe | Beschreibung |
|-------|--------------|
| 9 | perfect |
| 8 | |
| 7 | good |
| 6 | |
| 5 | satisfying |
| 4 | |
| 3 | poor |
| 2 | |
| 1 | unacceptable |
| | |

Tabelle 8: Zuordnung von Skalenstufen und Qualitätsbeschreibungen.

Die Skala wird dabei zusätzlich von zwei Ankern in Form von Audiobeispielen definiert. Diese sollen als Bezugspunkte die Skala leichter verständlich machen. Sie sind den Stufen 2 und 8 zugeordnet und über eine Gruppe von Tests bzw. Kombinationen aus Treibern und Stimuli konstant. Der Teilnehmer muss sich die Anker anhören bevor er die erste Bewertung abgeben kann. Die Anker repräsentieren dabei gute (8) bzw. schlechte (2) Klangqualitäten. Der Anker mit schlechter Qualität soll eindeutig identifizierbare Verzerrungen beinhalten, die möglichst wenig harmonischen Bezug zum dargebotenen Musikmaterial haben. Entsprechend ist Rub & Buzz eine geeignete Verzerrungsart.

Ein weiterer Aspekt ist, dass durch den gezielten Einsatz entsprechender Anker der Forschungsgegenstand an die Stichprobe angepasst werden kann. Wird der Test z.B. nur mit Expertenhörern durchgeführt, kann der Abstand zwischen den Ankern geringer sein bzw. der schlechte Anker weniger stark verzerrt sein. Dadurch wird die Nutzung der Skala der zu erwartenden kritischeren Einstellung der Experten angepasst. Weiterhin ermöglichen global über mehrere Tests verwendete Anker Vergleiche zwischen mehreren Treibern und Verzerrungsarten.

Der Teilnehmer kann sich paarweise Referenz- und Testsignale anhören und relativ zu den Ankern bewerten. Dabei können Kombinationen wiederholt werden. Die finalen Bewertungen werden erst nach Bestätigung durch den Teilnehmer gespeichert.

Die Dauer aller Audiobeispiele beträgt zwischen 7 und 10 s um vorzeitige Ermüdung der Teilnehmer zu vermeiden. Die Beispiele sind jeweils mit Ein- und Ausblendungen von mindestens 0,005 s Länge versehen um Ein- bzw. Ausschaltknacken zu vermeiden. Beide Vorgaben orientieren sich wiederum an der ITU-Empfehlung BS.1234-3 (vgl. 2015, S. 9).

Da der einzige Unterschied zwischen Referenz- und Testsignalen in dem Vorhandensein und der Skalierung nichtlinearer Verzerrungen besteht, ist davon auszugehen, dass die vorgenommene Bewertung maßgeblich von diesen bestimmt wird.

6.3.2 Anschluss an den bestehenden Hörversuch

Um im vorgestellten Testverfahren verlässliche Daten zu erhalten ist sicherzustellen dass die präsentierten Verzerrungen in den Testsignalen vom Teilnehmer detektiert werden können. Entsprechend erfolgt die Auswahl der Testsignale basierend auf der individuellen Hörschwelle des Teilnehmers, welche vorher mit dem bestehenden Verfahren für die gewählte Kombination ermittelt wird (siehe Abschnitt 3.5). Die Hörschwelle wird mit Bezug zum Skalierer S_{dis} angegeben. Ausgehend vom Abstand zur realen Performance des DuTs bei $S_{dis} = 0$ dB werden für die drei Testsignale Werte für S_{dis} in vier Fällen nach folgenden Kriterien bestimmt:

- Ein Testsignal soll stets bei der Detektionsschwelle AT liegen.
- Ein Testsignal soll bei $S_{dis} = 0$ dB liegen, wenn nicht bereits je ein Testsignal mit $S_{dis} > 0$ dB und $S_{dis} < 0$ dB benutzt wird. Im letztgenannten Fall wird die Bewertung der Klangqualität für den Punkt $S_{dis} = 0$ dB linear interpoliert.
- Zwischen den S_{dis} -Werten zweier Testsignale muss mindestens ein Abstand von ΔS_{dis} bestehen.
- Liegt die Hörschwelle *AT* im Grenzbereich zweier Fälle müssen beide Fälle die Gleiche Auswahl erzeugen.

Die Fallunterscheidung erfolgt analytisch nach Gl. 35:

$$S_{dis} = \begin{cases} \left\{ AT; \ \frac{AT}{2}; \ 0 \right\} dB &, \text{ wenn } AT \leq -2 \cdot \Delta S_{dis} \end{cases}$$

$$GI. 35$$

$$S_{dis} = \begin{cases} \left\{ \frac{AT}{2} - \Delta S_{dis}; \ \frac{AT}{2}; \ \frac{AT}{2} + \Delta S_{dis} \right\} dB, \text{ wenn } -2 \cdot \Delta S_{dis} \leq AT \leq 0 \end{cases}$$

$$\{AT - \Delta S_{dis}; \ AT; \ AT + \Delta S_{dis} \} dB &, \text{ wenn } 0 \leq AT \leq \Delta S_{dis} \end{cases}$$

$$\{0; \ AT; \ AT + \Delta S_{dis} \} dB &, \text{ wenn } \Delta S_{dis} \leq AT \end{cases}$$

Die Größe des Abstandes ΔS_{dis} wurde nach Testläufen mit Mitarbeitern der Klippel GmbH auf $\Delta S_{dis} = 9$ dB festgelegt.

Wie in Abschnitt 6.3.4 beschrieben können die Werte von S_{dis} für die Testsignale ausschließlich die benutzten Werte der auralisierten Audiobeispiele annehmen. Entsprechend muss in Gl. 35 auf den Wert des nächsten verfügbaren Beispiels aufgerundet werden. Weiterhin bestehen Grenzen für die Extremwerte der verfügbaren Werte von S_{dis} . Sollte ein auszuwählender S_{dis} -Wert über dem maximal verfügbaren Wert liegen, so wird der Maximalwert an dessen Stelle benutzt. Analog wird der minimal verfügbare Wert benutzt, wenn die Hörschwelle den verfügbaren Bereich von S_{dis} ausreizt.

6.3.3 Robustheit

Die Rolle der Anker ist die Nutzung der Skala über eine Gruppe von Tests für alle Teilnehmer des Versuchs möglichst zu vereinheitlichen. Die Erfüllung dieser Funktion ist durch Untersuchungen der Streuung der Skalennutzung zu verifizieren. Für die zunächst angestrebte Betrachtung der Gradienten ΔSQ ist der Test ungeachtet der korrekten Funktion der Ankersignale trotzdem ausreichend robust.

Abbildung 49 illustriert die Robustheit des Verfahrens hinsichtlich der Auswertung der Gradienten an zwei fiktiven Messreihen der Teilnehmer P_1 und P_2 . Obwohl beide Teilnehmer die Skala unterschiedlich nutzen (P_1 : 9 bis 4, P_2 : 6 bis 1) sind die Gradienten sehr ähnlich. Es ist demzufolge davon auszugehen, dass auch bei Nicht-Verständnis der Anker durch den Teilnehmer valide Daten für die Ermittlung der Gradienten vorliegen können.



Abbildung 49: Beispiel zur unterschiedlichen Skalennutzung.

Dicke Pfeile: Gradienten $\Delta SQ_{P_i,AT}$.

Zudem wirkt sich die Benutzung der neunstufigen Skala positiv auf die Robustheit hinsichtlich der Bestimmung der Gradienten aus, da sich eine Abweichung um eine Skalenstufe bei einer Messwiederholung weniger stark auswirkt als bei einer fünfstufigen Skala.

6.3.4 Umsetzung mit Webinterface

Das Testverfahren wurde als Webseite implementiert. Somit eignet es sich zur Veröffentlichung im Internet und ermöglicht große Strichproben zu erreichen. Gleichzeitig sollte die Möglichkeit bestehen den Versuch offline auf lokalen Rechnern durchführen zu können. Hierfür wurde ein Installationspaket erzeugt, welches mit Microsoft Windows

kompatibel ist und auf einem lokalen Webserver die Funktionalität ohne Internetverbindung herstellt. Die folgenden Erläuterungen beschreiben grob den Ablauf zur Erstellung eines Tests und die dem Teilnehmer präsentierte Benutzeroberfläche.

Anlegen einer Testreihe

Der Versuchsleiter muss alle benötigten Audiobeispiele in einer festgelegten Ordnerstruktur ablegen, aus der eine neue Testreihe erzeugt wird. Die Ordnerstruktur wird in Abbildung 50 schematisch dargestellt:



Abbildung 50: Schema der Dateistruktur.

Innerhalb eines Stammverzeichnisses "Sounds" werden die Audiosignale der Anker abgelegt. Weiterhin wird dort für jeden Treiber ein eigener Unterordner angelegt. Innerhalb der Treiberordner wird für jedes auralisierte Signal wiederum ein Unterordner erzeugt. In diesem Unterordner werden das Referenz- und die Testsignale hinterlegt. Mittels der Datei- und Ordnernamen wird erkannt ob es sich bei einer Datei um das Referenz- oder ein Testsignal handelt. Im letztgenannten Fall wird aus dem Dateinamen der jeweilige Wert von S_{dis} ausgelesen. Mit den Ordnernamen erfolgt die Zuordnung der Signale zu den jeweiligen Verzerrungen.

Zur Automationsroutine aus Abschnitt 5.2 wurde eine Exportskript im MAT2-Format entwickelt welches mittels des Klippel Automation Interface direkt aus Datenbanken von dB-Lab die entsprechenden Ordner und Dateien mit den richtigen Bezeichnungen und den gewünschten Abstufungen von S_{dis} erzeugt. Diese müssen nur noch ins Testverzeichnis kopiert werden.

Benutzeroberfläche

Alle Seiten des Versuchs weisen den gleichen Aufbau auf: Zuoberst ist als Seitenkopf eine Darstellung abgebildet, die dem Corporate Design der Firma entspricht und weiterführende Links enthält. Darunter ist stets ein Satz Anweisungen dargestellt. Diese Anweisungen sind durch einen Rahmen hervorgehobenen und erläutern die Schritte, die ein Teilnehmer auf der jeweiligen Seite zu absolvieren hat. Die Benutzerführung erfolgt dabei strikt linear. Zu den Anweisungen werden zusätzlich die jeweiligen Schaltflächen interaktiv hervorgehoben, die im aktuellen Schritt zu benutzen sind. Schaltflächen die nach dem jeweils aktuellen Schritt benutzt werden müssen, sind zunächst ausgegraut und inaktiv. Alle Benutzereingaben sind unter der Anweisungsbox dargestellt.

Die Benutzeroberfläche untergliedert sich in vier Hauptseiten. Die Startseite stellt eine Begrüßung dar und erklärt kurz den Inhalt des Versuchs und den möglichen Nutzen für den Teilnehmer.

Die Testauswahl erfolgt auf der nachfolgenden Seite. Hier kann der Teilnehmer Treiber und Stimulus auswählen und mit einer Audiovorschau entscheiden, ob er den gewählten Test absolvieren möchte. Optional kann der Teilnehmer noch mitteilen, ob er Erfahrung mit diesem Hörversuch hat und welche Abhörgeräte benutzt werden. Letztere werden in folgende Kategorien unterschieden: Lautsprecher, Kopfhörer, Smartphonelautsprecher oder Notebooklautsprecher. Diese Auswahl wurde zusammengestellt um später eventuelle Korrelationen zwischen den benutzen Geräten und den Testergebnissen ziehen zu können. Eine spezifischere Abfrage erscheint nicht sinnvoll, da weitere Einflüsse wie akustische Gegebenheiten im Raum des Teilnehmers oder die einwandfreie Funktion der Abhörgeräte ohnehin nicht reliabel erfasst bzw. geprüft werden können.

Im Anschluss werden die jeweiligen Testabschnitte zur Ermittlung der Detektionsschwelle und der Bewertung der Klangqualität auf separaten Seiten dargestellt und durchgeführt.

Die Benutzeroberfläche wurde für möglichst viele Endgeräte optimiert, der Schwerpunkt zum Abschluss der Arbeit lag dabei vor allem auf der Benutzbarkeit mit Smartphones. Abbildung 51 stellt exemplarisch die Benutzeroberfläche der Erweiterung des Hörversuchs auf einem Smartphone dar.



Abbildung 51: Benutzeroberfläche auf Smartphones.

Good / Bad: Ankersignale; Stimulus: Referenzsignal; 1st / 2nd / 3rd: Testsignale. Die Weiterschaltung der Testsignale erfolgt mit den Schaltflächen Next / Previous. Der Teilnehmer wählt je Testsignal eine Bewertung auf der dargestellten Skala aus. Sind alle Bewertungen abgegeben wird der Test mit der Schaltfläche Finish beendet.

Die Benutzeroberfläche ist für mehrere Sprachen ausgelegt. Die jeweilige Sprache kann im Kopf der Seite ausgewählt werden. Derzeit ist nur eine englischsprachige Oberfläche verfügbar, spätere Versionen sollen Deutsch und Chinesisch beinhalten.

Anreize für den Teilnehmer

Nach jedem Testabschnitt werden dem Teilnehmer Zwischenergebnisse angezeigt und beschrieben. Diese Ausgaben dienen der Information und stellen gleichzeitig einen Anreiz dar weitere Durchläufe des Versuchs zu absolvieren. Die Zwischenergebnisse sind nachfolgend kurz beschrieben.



Abbildung 52: Verlaufsanzeige der Detektionsschwellenbestimmung.

Wenn die Detektionsschwelle für die gewählte Kombination aus Treiber und Stimulus ermittelt wurde, wird dem Teilnehmer der Verlauf seiner Antworten in Relation zur Skalierung der Verzerrung wie in Abbildung 52 dargestellt. Dabei ist ersichtlich bei welcher Skalierung die reale Performance des DuTs liegt. Ausgehend von der Lage der ermittelten Hörschwelle AT zu $S_{dis} = 0$ dB wird zusätzlich verbalisiert ob der Anteil der nichtlinearen Verzerrungen verstärkt oder vermindert wurde ("enhanced / attenuated distortion") und wie sich dies tendenziell auf die Hörbarkeit auswirkt ("more audible / less audible").

Eine weitere grafische Ausgabe zur Hörschwelle schließt sich an den zweiten Testabschnitt zur Bewertung der Klangqualität an. In einem Histogramm wie in Abbildung 53 wird dem Probanden die Verteilung der ermittelten Hörschwellen *AT* für die Stichprobe dargestellt. Die vom Teilnehmer erreichte Hörschwelle ist hervorgehoben. Der Teilnehmer kann in diesem Diagramm sehen wie viele weitere Teilnehmer den gewählten Test bereits absolviert haben und welche Hörschwellen für diese bestimmt wurden. Die Darstellung wird mit begleitenden Texten erläutert.



Abbildung 53: Histogramm der Hörschwellenverteilung.

Auch die vorgenommene Bewertung der Klangqualität durch den Teilnehmer wird grafisch aufbereitet. In einem Diagramm mit den Werten von S_{dis} als Abszisse und der Bewertungsskala als Ordinate werden die Bewertungen der drei Testsignale aufgetragen und ein möglicher Verlauf angedeutet. Per Mouseover kann der Teilnehmer erfahren, welches der drei Testsignale welchem Punkt in Abbildung 54 zugeordnet ist. Auch wird die bewertete oder interpolierte Klangqualität des DuTs dargestellt. Der Darstellung wurden erklärende Texte beigestellt.



Abbildung 54: Darstellung der Bewertung der Klangqualität.

Ausgehend von der ermittelten Bewertung der Klangqualität wird ein letztes Histogramm dargestellt. Dieses illustriert die Bewertung des Treibers durch den Teilnehmer in Relation zu den Bewertungen der restlichen Stichprobe des gewählten Tests. Das Histogramm hat analog zu Abbildung 53 die Anzahl der Bewertungen der jeweiligen Skalenstufe als Ordinate und die Bewertungsskala als Abszisse. Abbildung 55 zeigt ein Beispiel. Das Histogramm wird wiederum mit einem kurzen Text erläutert.



Abbildung 55: Histogramm der Bewertung der Klangqualität.

Zusätzlich zu den Darstellungen, welche die Ergebnisse des Teilnehmers einbeziehen, stehen auch jederzeit erklärende Texte und Illustrationen zu den relevanten Aspekten und Techniken des Verfahrens per Link in den Anweisungsboxen und dem Seitenkopf zur Verfügung.

Datenbank

Zur späteren Auswertung werden alle relevanten Daten in einer relationalen SQL-Datenbank gesichert. Für die Verarbeitung in statistischen Berechnungsprogrammen wie SPSS oder R ist eine Exportfunktion implementiert, welche die Daten in das kompatible CSV-Format speichert. Die Daten sind dabei nach Fällen geordnet und anonymisiert. Die für eine Auswertung relevanten Einträge sind auf der folgenden Seite in Tabelle 9 erläutert.

| Eintrag | Format | Beschreibung |
|----------------------------|-----------------------------------|--|
| userid | Hexadezimalzahl, Länge 128 Bit | Zufällig vergebene Fallnummer |
| outputDevice | Text | Verwendete Abhörgeräte |
| experience | Text | Bestehende Erfahrung mit dem Testverfahren |
| audibilitythresholdTestId | Dezimalzahl | Zähler aller bestimmten Detektionsschwellen |
| qualityratingTestId | Dezimalzahl | Zähler aller abgeschlossenen Bewertungen der Klangqualität |
| speaker | Text | Ausgewähltes DuTs für den jeweiligen Fall |
| distortion | Text | Ausgewählter Stimulus und aktive Nichtlinearitäten für den jeweiligen Fall |
| audibilitythresholdStarted | Datum | Datum und Uhrzeit der Detektionsschwellenbestimmung |
| audibilitythresholdResult | Dezimalzahl | Bestimmte Detektionsschwelle für den Fall |
| qualityratingStarted | Datum | Datum und Uhrzeit der Bewertung der Klangqualität |
| QRStep, QRValue, QRRated | Dezimalzahl | Zuordnung von Position des Testsignals zum jeweiligen Wert von S_{dis} und der Bewertung |
| realSpeakerInt | Dezimalzahl | Erfasste bzw. interpolierte Bewertung des DuTs |

Tabelle 9: Übersicht erfasster Daten.

Die Identifizierung der Teilnehmer erfolgt mittels Cookie im Browser des jeweiligen Benutzers. Ein Cookie ist eine Textinformation die z.B. Informationen über den Verlauf besuchter Webseiten enthält. So lang der Cookie auf dem Computer des Probanden nicht gelöscht wird, kann vom Webserver des Hörversuchs erkannt werden ob der Teilnehmer schon einmal einen Test absolviert hat. In diesem Falle wird einem neuen Datensatz die beim letzten Besuch der Webseite vergebene Identifizierung der Variable "userid" erneut zugewiesen. Sollte ein Teilnehmer keine Vorerfahrung angeben, so ist davon auszugehen, dass ein neuer Teilnehmer den Versuch absolviert und dabei einen Rechner nutzt, der schon einmal registriert wurde.

6.4 Anwendungsbeispiel

6.4.1 Verwendete Signale

Mit der in Abschnitt 5 vorgestellten Automationsroutine wurden für die drei ausgewählten Treiber aus der Musikauswahl in Anhang 1 mehrere kritische Stimuli identifiziert. Die identifizierten Titel aus Abschnitt 5.3.4 wurden mit der Automationsroutine für je separat aktive Bl(x)-, $K_{ms}(x)$ - und $L_e(x)$ -Verzerrungen und alle drei aktiven nichtlinearen Verzerrungen verarbeitet. Der auralisierte Bereich für S_{dis} betrug -42 dB $\leq S_{dis} \leq$ 24 dB mit einer Schrittweite von 3 dB.

Weiterhin wurden zwei Testsignale für alle dominanten nichtlinearen Verzerrungen auralisiert. Das erste ist ein Zweitonsignal mit den Frequenzen $f_1 = 30$ Hz und $f_2 = 800$ Hz bei gleichem Pegel. Wegen der tiefen unteren Frequenz f_1 wurde das Zweitonsignal nicht für den Microspeaker verwendet. Das zweite Signal ist ein Multitonsignal und enthält die nach Gl. 36 berechneten Frequenzen. Das Signal wurde für alle Treiber auralisiert.

$$f(i) = \frac{1500}{2^i}$$
 Hz, mit $i \in Z$ und $1 \le i \le 6$ Gl. 36

Die auralisierten Daten wurden in den Hörversuch eingepflegt. Entsprechend sind für Speaker#001 und Speaker#002 je 14 Tests wählbar und für den Microspeaker 13 Tests verfügbar.

In den auralisierten Daten wurden basierend auf den Zeitverläufen der Distortion To Mask Ratio Zeitpunkte identifiziert, die sich für die Verwendung im Hörversuch eignen. Die Zeitpunkte wurden in den geschnittenen Testsignalen zu je 7 s Spieldauer beinhaltet. Weiterhin wurden alle Klangbeispiele eines Tests auf identische RMS-Pegel normalisiert.

Die Ankersignale wurden mittels Mikrofonaufnahme einer Lautsprecherwiedergabe erzeugt. Der Aufbau glich dem aus Abbildung 27 (Seite 48). Für Referenz- und Testaufnahme wurde der Verstärkungsfaktor nicht verändert. Die irregulären Defekte wurden in der Testaufnahme durch das Auflegen der abgebildeten Maschinenschraube mit Mutter erzeugt. Der Anteil der Rub&Buzz-Verzerrung wurde nachträglich mit dem DIF-AUR-Modul um 3 dB reduziert. Der benutzte Titel ist Anhang 1 angegeben.

Die Audiobeispiele wurden für den durchgeführten Hörversuch ins MP3-Format mit einer konstanten Bitrate von 320 kbit/s encodiert, um Ladezeiten bei mobiler Konnektivität zu minimieren. Zudem wird die Wiedergabe von Audiodateien im PCM-Format nicht von allen Browsern unterstützt. Die durch die verlusthafte Codierung entstehenden Beeinträchtigungen der Signale werden als nicht relevant für den Hörversuch betrachtet. Alle Signale stehen als digitaler Anhang zur Verfügung.

6.4.2 Urheberrechtliche Bedenken

Die Mehrzahl der verwendeten Musikbeispiele unterliegt Leistungsschutz- bzw. Verwertungsrechten im Sinne des Urheberrechts. Handelt es sich bei den verwendeten Werken in einem Hörversuch nicht um gemeinfreie Werke, so ist die Verletzung von Urheberrechten zu prüfen. Als gemeinfrei werden dabei Werke betrachtet, deren sämtliche Schutzrechte durch Ablauf aller Schutzfristen erloschen sind (vgl. Matthé & Fischer, ohne Jahr, S. 4), oder Werke die unter einer freien Lizenz explizit zur Weiterverbreitung und Bearbeitung veröffentlicht worden sind.

Die im Sinne des Urheberrechts entstehenden Verwertungsrechte werden durch entsprechende Gesellschaften gewahrt. Für den deutschen Musikmarkt erledigt dies die *Gesellschaft für musikalische Aufführungs- und mechanische Vervielfältigungsrechte* (GEMA). Ihre Rolle ist vor allem die Erhebung und Abrechnung etwaiger Lizenzgebühren für die Verwertung musikalischer Werke.

Zu den im Rahmen dieses Hörversuchs wichtigsten Paragraphen des Urheberrechtsgesetzes zählen §14 und §19a. Nach §14 kann die Entstellung eines Werks vom Urheber verboten werden (vgl. Bundesministerium der Justiz, 2016, S. 11). In diesem Zusammenhang bleibt zu klären ob der Auralisationsprozess eine Entstellung des ursprünglichen Werkes darstellt. §19a umfasst das Recht auf orts- und zeitunabhängige Zugänglichmachung eines Werkes durch den Urheber (vgl. ebd., S. 12) und ist am ehesten auf die Verbreitung im Internet anzuwenden. Die Durchführung des Hörversuchs im Internet kann eine Verletzung dieses Rechts darstellen.

Jedoch werden Lehre und Forschung besondere Rechte eingeräumt: §52a räumt die Veröffentlichung kurzer Auszüge eines Werkes für Forschungszwecke ein (vgl. ebd., S. 25-26) und setzt somit §19a unter gewissen Umständen aus (vgl. Matthé & Fischer, ohne Jahr, S. 12). Zu diesen Umständen zählt die forschungsbedingte Veröffentlichung abseits kommerzieller Interessen (vgl. ebd. S. 13). Für Musikstücke werden dabei Auszüge bis 5 min Spieldauer als zulässig betrachtet (vgl. ebd. S. 14). Speziell für die Verbreitung im Internet ist der Zugang durch technische Beschränkungen dabei nur einem abgegrenzten Personenkreis zu ermöglichen (vgl. ebd.).

Hinsichtlich der Verwendung von nicht-gemeinfreien Musikstücken im Forschungsinteresse der vorliegenden Arbeit erscheint die Anwendung des §52a gerechtfertigt, da

- mit einer Spieldauer von max. 10 s nur äußerst kurze Auszüge der musikalischen Werke genutzt wurden,
- kein direktes kommerzielles Interesse besteht und
- der Zugang zum Hörversuch nur über einen in Suchmaschinen nicht-gelisteten Link ermöglicht wurde, der wiederum nur mit geschlossenen Interessengruppen, wie der Liste der Studenten der Audiokommunikation der Technischen Universität Berlin geteilt wurde.
- Weiterhin wurde darauf geachtet das zumindest Künstler und Titel ersichtlich sind.

Ungeachtet davon besteht weiterhin der Anspruch seitens der Urheber gemäß §14. Gleichwohl erheben die eben gemachten Darstellungen keinen Anspruch auf Vollständigkeit oder Rechtsverbindlichkeit.

6.4.3 Auszug aus den Ergebnissen

Mit dem beschriebenen Verfahren wurde am 19. Juni 2017 ein Testlauf gestartet. Der Zugangslink zum Hörversuch wurde in der Mailingliste der Studenten der Audiokommunikation der Technischen Universität Berlin geteilt und den Mitarbeitern der Klippel GmbH ausgehändigt. Exemplarisch wird folgend ein Auszug aus den Ergebnissen gegeben.

Bis zum 10. Juli 2017 nahmen 32 unterschiedliche Probanden am Hörversuch teil und absolvierten insgesamt 59 Tests.

Der überwiegende Teil der Tests wurde mit Kopfhörern absolviert. Tabelle 10 listet die Häufigkeiten der verwendeten Abhörgeräte auf.

| Abhörgeräte | Häufigkeit | Prozent | Kumulierte Prozente |
|-----------------------------|------------|---------|------------------------|
| Built-in Laptop Speakers | 4 | 6,8 | 6,8 |
| Headphones | 50 | 84,7 | 91,5 |
| Smartphone Speakers | 1 | 1,7 | 93,2 |
| Speakers | 4 | 6,8 | 100,0 |
| Gesamt | 59 | 100,0 | |

Tabelle 10: Häufigkeiten verwendeter Abhörgeräte.

Die Mehrheit der Benutzer wählte gemäß Tabelle 11 die Treiber Speaker#001 und den Microspeaker zur Durchführung der Tests aus.

| DuT | Häufigkeit | Prozent | Kumulierte Prozente |
|--------------|------------|---------|------------------------|
| Microspeaker | 23 | 39,0 | 39,0 |
| Speaker001 | 21 | 35,6 | 74,6 |
| Speaker002 | 15 | 25,4 | 100,0 |
| Gesamt | 59 | 100,0 | |

Tabelle 11: Häufigkeiten durchgeführter Test, sortiert nach Treibern.

Tabelle 12: Häufigkeiten durchgeführter Tests, DuT Microspeaker.

| Künstler - Titel - Verzerrung | Häufigkeit | Prozent | Kumulierte Prozente |
|---|------------|---------|------------------------|
| Avicii ft Adele - Addicted to you - all distortions | 6 | 26,1 | 26,1 |
| Avicii ft Adele - Addicted to you - $Bl(x)$ | 9 | 39,1 | 65,2 |
| Avicii ft Adele - Addicted to you - $K_{ms}(x)$ | 1 | 4,3 | 69,6 |
| Bassida - Be mine- all distortions | 1 | 4,3 | 73,9 |
| Bassida - You make me wanna stay - all distortions | 2 | 8,7 | 82,6 |
| Bassida - You make me wanna stay - $Bl(x)$ | 1 | 4,3 | 87,0 |
| Bassida - You make me wanna stay - $K_{ms}(x)$ | 3 | 13,0 | 100,0 |
| Gesamt | 23 | 100,0 | |

Die verhältnismäßig geringe Zahl an Teilnehmern führt dazu, dass nur wenige der 41 möglichen Tests durchgeführt wurden. Am Häufigsten wurde der Microspeaker mit Titel 1 und der Kraftfaktor-Nichtlinearität gewählt. Diese Testkombination wurde insgesamt neun Mal absolviert, wie Tabelle 12 zu entnehmen ist.

Weitere Ergebnisse werden an dieser Stelle der Arbeit nicht angeführt, da sie für die folgenden Abschnitte nicht von Bedeutung sind. Der erhobene Datensatz steht als digitaler Anhang zur Verfügung.

6.4.4 Auswertung des Testlaufs

Eine umfassende statistische Auswertung war nicht Teil der Aufgabenstellung und kann auf Grund der geringen Teilnehmerzahl auch nicht zu validen Ergebnissen führen. Vielmehr befasst sich dieser Abschnitt mit Problemen die im Testlauf ersichtlich wurden.

Die potenzielle Stichprobengröße beträgt rund 340 Personen. Diese setzt sich aus den rund 45 Mitarbeitern und Studenten der Klippel GmbH und den 295 gelisteten Empfängern der Mailingliste der Studenten des Fachbereichs Audiokommunikation zusammen. Mit 32 tatsächlichen Teilnehmern hat somit aus der potenziell erzielbaren Stichprobe nur rund ein Zehntel teilgenommen. Dem gegenüber steht eine Vielzahl von auswählbaren Tests. Diese beläuft sich auf 41 auswählbare Kombinationen von Treibern, Stimuli und Verzerrungen. Im Durchschnitt hat jeder der 32 Teilnehmer rund 1,8 Tests absolviert. Wie bereits Tabelle 12 zeigt, wurden einige Tests gar nicht oder von nur sehr wenigen Teilnehmern absolviert. Entsprechend ist eine statistische Auswertung mit dem erhobenen Datensatz noch nicht sinnvoll.

Weiterhin wurde bei der Verteilung des Zugangslinks um Rückmeldung der Teilnehmer gebeten, falls Probleme mit der Benutzung bestehen. Mehrere Teilnehmer nutzten diese Möglichkeit. Bei den erhaltenen Rückmeldungen wurde vor allem das Konzept bzw. die Rolle der Anker thematisiert. Es zeigte sich dabei mehrheitlich, dass es für Teilnehmer zunächst verwirrend sein kann, dass der Audioinhalt der Anker von dem der Referenz- und Testsignale abweicht.

Weitere Rückmeldungen befassten sich mit der Benutzeroberfläche. Diese wurde von Personen, die den Test nicht auf einem Smartphone durchgeführt haben, als nicht optimal beschrieben.

6.5 Ausblick

Für eine künftige Version des Hörversuchs können weitere Optimierungen vorgenommen werden. Zunächst muss die Anzahl der möglichen Tests der zu erwartenden Größe der Stichprobe angepasst werden. Denkbar wäre für bestimmte Zeiträume stets nur einen Treiber mit einer eingeschränkten Auswahl an Stimuli anzubieten. Gemeinsam mit einem stets angebotenen Referenztest können so mehrere Treiber aus verschiedenen Zeiträumen verglichen werden. Alternativ kann ein versteckter Anker in Form eines vierten zu bewertenden Testsignals unter Umständen geeignet sein um Vergleichbarkeit herzustellen.

Weiterhin ist das Konzept und die Notwendigkeit der Anker zur Definition der Bewertungsskala zu überprüfen. Es kann angenommen werden, dass die von Teilnehmern als verwirrend beschriebene Wirkung Einfluss auf zu erzielende Ergebnisse hat. Darüber hinaus wurde gezeigt, dass das Konzept der Gradienten ΔSQ theoretisch ausreichend robust gegenüber starken Streuungen in der Benutzung der Bewertungsskala ist. Für andere Analysen des Datensatzes kann jedoch eine Notwendigkeit der Ankersignale gegeben sein. Beispielsweise kann bei einer ausreichend großen Stichprobe neben den Gradienten auch der tatsächliche Verlauf der Änderung der Klangqualität mittels Regressionsanalysen ermittelt werden. In diesem Falle ist eine weite Streuung in der Benutzung der Bewertungsskala zu berücksichtigen.

Die Benutzeroberfläche muss neben der Optimierung für Smartphones auch in einer Version vorliegen, die komfortabel von anderen Endgeräten bedient werden kann, beispielsweise Laptops. Weiterhin ist die Mehrsprachigkeit bisher nur für die englische Sprache funktional. Weitere Sprachen müssen ergänzt werden.

Die Pegelkompensation der auralisierten Signale zueinander und zu den Ankersignalen innerhalb eines Tests ist zu optimieren. In der gegenwärtigen Version werden die Testsignale zueinander auf Basis des RMS-Pegels normalisiert. Eine Anpassung der Ankersignale erfolgt ebenfalls nur zwischen den Ankersignalen. In beiden Fällen ist eine Normalisierung der Signale zueinander auf Basis der spezifischen Lautheit der Signale zu bevorzugen. Darüber hinaus sollten auch Unterschiede innerhalb eines gewählten Tests zwischen Anker und Testsignalen kompensiert werden.

Hinsichtlich der Überprüfung der eingeführten Hypothesen ist ein erneuter Versuch durchzuführen. Idealer Weise sind die in diesem Ausblick enthaltenen Vorschläge dann bereits berücksichtigt. Mit einer reduzierten Anzahl an Tests und einer ausreichend großen Stichprobe können ausreichend Daten erhoben werden um die Hypothesen statistisch valide zu überprüfen bzw. die relevanten Zusammenhänge zu quantifizieren.

Zuletzt ist für die Überführung in ein Produkt der Klippel GmbH mit kommerziellem Interesse sicherzustellen, dass entweder sämtliche verwendeten Werke gemeinfrei sind oder entsprechende Lizenzrechte von der GEMA bzw. den Rechtsvertretern der Urheber erworben werden.

7 Zusammenfassung

Die vorliegende Masterarbeit befasste sich mit Methoden zur perzeptiven Evaluation nichtlinearer Verzerrungen bei der Lausprecherwiedergabe. Die Untersuchungen erfolgten untergliedert in drei Themenkomplexe: Der Evaluation von Teilen eines perzeptiven Evaluationsmoduls, der Entwicklung und Implementierung einer Routine zur Identifizierung kritischer Signale und der Entwicklung und Umsetzung eines neuen Hörversuchsverfahrens.

Das perzeptive Evaluationsmodul benutzt Implementierungen des PEAQ-Verfahrens der International Telecommunication Union zur rechnergestützten Analyse der Qualitätsbeeinträchtigungen bei der Lautsprecherwiedergabe. Dabei werden periphere Gehörmodelle eingesetzt um ein Referenz- und ein Testsignal in Erregungsmuster zu überführen und psychoakustisch relevante Modellausgaben zu berechnen.

Nach der zunächst notwendigen Aktualisierung des skriptbasierten Moduls konnten Defizite der Implementierung festgestellt werden. Derzeit erfordert die Verarbeitung lange Rechenzeiten, was in der Ausführung der teils redundanten Gehörmodelle begründet ist. Es kommt sowohl ein FFT-basiertes Modell als aus ein Filterbankmodell zum Einsatz um die Ausgaben zu berechnen. Darüber hinaus wurden Fehler bei der Skalierung und Filterung der Eingangssignale im Filterbankmodell identifiziert und behoben.

Die Evaluation der Berechnung der Modellausgaben liefert gemischte Ergebnisse: Das Rechenmodell nach Daniel & Weber zur Abschätzung der Rauigkeit der Signale war zunächst nicht funktional implementiert. Fehler in der Blockverarbeitung konnten behoben werden; die Modelberechnungen zeigen jedoch teils starke Abweichungen zu den empirischen Daten nach Fastl & Zwicker. Die Modellausgaben zur absoluten Schärfe hingegen zeigen eine gute Näherung an die empirischen Daten. Weiterhin konnte die mäßige Aussagekraft der in PEAQ vorgeschlagenen Maskierungsmaße *DMR* und *DMR*_{tot} hinsichtlich der Anwendung an Lautsprechern herausgestellt werden. Die relativen Maße Discoloration, Treble Stressing und Bass Enhancement liefern gute, deskriptive Informationen. Die Anwendung aller Modellausgaben für signifikant nichtlinear verzerrte Signale bilden die zu erwartenden Zusammenhänge aus der Psychoakustik ab. Dies gilt jedoch nur für reguläre Nichtlinearitäten.

Weitere Arbeiten am perzeptiven Evaluationsmodul können sich zunächst mit der Entfernung der Redundanzen der Gehörmodelle beschäftigen und die Abweichungen des Rauigkeitsmodells untersuchen und beheben. Darüber hinaus sind die Maskierungsmaße in für Lautsprecher sinnvolle Angaben zu transformieren. Diesbezüglich gilt es die Interaktionen zwischen Frequenzgruppen der Hörwahrnehmung abzubilden und zeitliche Schwellen für die Hörbarkeit signifikanter nichtlinearer Verzerrungen zu finden. Weiterhin können in Hörversuchen kritische Bereiche für die relativen Maße ermittelt werden.

Ein Konzept zur Suche kritischer Stimuli wurde entwickelt und vorgestellt. Basierend auf der thermischen Simulation der Zustandsvariablen bei der Lautsprecherwiedergabe können für beliebige Eingangssignale entsprechende Verstärkungsfaktoren ermittelt werden. Mit diesen kann die Wiedergabe auralisiert und anschließend nach perzeptiven Gesichtspunkten analysiert werden. Durch die Nutzung der linearen und nichtlinearen Parameter echter Treiber ist eine realitätsnahe Bewertung möglich.

Das vorgestellte Konzept wurde innerhalb des Klippel R&D-Systems als Automationsroutine implementiert und erfolgreich für drei Treiber und über 60 Stimuli getestet. Die Vorteile des Verfahrens liegen neben der Automatisierung im geringen technischen Aufwand der Durchführung. Schwächen liegen derzeit noch in der benötigten Rechenzeit, die durch die Defizite des perzeptiven Evaluationsmoduls bedingt sind. Weiterhin ist die Handhabung gemessener oder synthetisierter Treiberdaten zu verbessern.

Zukünftige Arbeiten können sich mit der Prozessoptimierung der Automationsroutine befassen. Insbesondere ist die weitere Integration in das R&D-System zu realisieren. Weiterhin ist der Aufbau einer Datenbank mit Daten verschiedener Stimuli und Treiber denkbar, die mit einem Klassifikationsalgorithmus eine Vorauswahl kritischer Stimuli treffen kann.

Eine Erweiterung des Hörversuchs wurde vorgeschlagen. Dazu wurde der Begriff Klangqualität als Qualität der Wiedergabe eines Lautsprechers definiert. Diese sei geprägt von den linearen und nichtlinearen Übertragungseigenschaften des Treibers. Speziell der Einfluss nichtlinearer Verzerrungen auf die Klangqualität ist von Interesse. Es wurde angenommen, dass ein gerichteter Zusammenhang zwischen der Hörbarkeit nichtlinearer Verzerrungen und der Bewertung der Klangqualität besteht. Dieser gerichtete Zusammenhang bestehe in Form einer psychometrischen Funktion, welche die Änderung der Klangqualität SQ über der Änderung Skalierung S_{dis} der nichtlinearen Verzerrungen abbildet.

Um die Funktion empirisch zu untersuchen wurden die Gradienten der Änderung der Klangqualität ΔSQ als Approximation vorgestellt. Zur Erfassung dieser Gradienten wurde ein Messverfahren vorgestellt, dass diese Gradienten ausgehend von der Hörschwelle eines Probanden mit drei Messpunkten an relevanten Stellen erfassen soll. Die Messpunkte werden durch Paarvergleiche von Klangbeispielen erhoben, bei denen ein Referenzsignal ausschließlich linear verzerrt ist und ein Testsignal lineare und nichtlineare Verzerrungen enthält. Drei unterschiedlich stark verzerrte Testsignale sind auf einer neunstufigen Skala zu bewerten. Die Skala wird sowohl verbal beschrieben, als auch mit festen Ankersignalen definiert. Der Proband kann einen Test aus einer Auswahl an Treibern, Stimuli und Verzerrungen dabei selbst wählen; mehrere Tests sollen durch konstante Anker vergleichbar sein.

Das Messverfahren wurde als Erweiterung an den bestehenden Onlinehörversuch der Klippel GmbH angeknüpft und erfolgreich getestet. Für den Testlauf wurden Audiobeispiele mit der Automationsroutine generiert und dem aktualisierten perzeptiven Modul ausgewertet. Der Fokus des Testlaufs lag nicht vorrangig auf der statistischen Validierung der hypothetischen Zusammenhänge, da dies nicht Teil der Aufgabenstellung gewesen ist. Vielmehr sollte die Benutzbarkeit des umgesetzten Verfahrens untersucht werden.

Es zeigte sich, dass für den Testlauf zu viele Testkombinationen wählbar waren; da von der potenziellen Stichprobe nur relative wenig Probanden akquiriert werden konnten, wurde einige Tests nicht absolviert. Die Teilnehmer wurden auch um Rückmeldung zum Verfahren gebeten. In eingetroffenen Rückmeldungen wurde vor allem das Konzept der konstanten Anker als problematisch thematisiert. Ein weiterer Kritikpunkt war die Smartphone-Orientierung der Benutzerführung.

Weitere Arbeiten können daher Versuche beinhalten die hypothetischen Zusammenhänge statistisch zu validieren. Darüber hinaus gilt es das Konzept der Anker zu prüfen und die Benutzeroberfläche für weitere Endgeräte zu optimieren.

Die Bearbeitung der gestellten Aufgaben verknüpfte Psychoakustik mit Elektroakustik und numerischer Simulation und lieferte überdies Ein- und Ausblicke in äußerst interessante Zusammenhänge der modernen Lautsprecherentwicklung.

LITERATURVERZEICHNIS

Aures, W. (1985). Ein Berechnungsverfahren der Rauhigkeit . Acustica (Vol. 58), 268 ff.

Bundesministerium der Justiz. (20. Dezember 2016). Urheberrechtsgesetz vom 9. September 1965 (BGBl. I S. 1273), das zuletzt durch Artikel 1 des Gesetzes vom 20. Dezember 2016 (BGBl. I S. 3037) geändert worden ist. Abgerufen am 5. Juli 2017 von https://www.gesetze-im-internet.de/urhg/UrhG.pdf

Daniel, P., & Weber, R. (1997). Psychoacoustical Roughness: Implementation of an Optimized Model. *Acustica* (Vol. 83), 113-123.

Dickreiter, M. (1997). *Handbuch der Tonstudiotechnik* (6. Ausg., Bd. 1). (S. f. Rundfunktechnik, Hrsg.) München: K.G. Saur.

Duisters, R. (25. März 2005). *The modeling of auditory roughness for signals with temporally asymmetric envelopes*. Eindhoven, Niederlande: Technische Universiteit Eindhoven.

Fastl, H., & Zwicker, E. (2007). *Psychoacoustics. Facts and Models.* Berlin Heidelberg: Springer Verlag.

Goertz, A. (2008). Lautsprecher. In S. Weinzierl (Hrsg.), *Handbuch der Audiotechnik* (S. 421 ff.). Berlin Heidelberg: Springer Verlag.

ITU. (2001). Recommendation BS.1387-1: Method for objective measurements of perceived audio quality. Genf: ITU.

ITU. (2015). Recommendation BS.1534-3: Method for the subjective assessment of intermediate quality level of audio systems. Genf: ITU.

Kaernbach, C. (1991). Simple adaptive Testing with the weighted up-down method. In *Perception & Psychophysics*. (S. 227-229). Bordeaux.

Klippel GmbH. (21. Januar 2015a). *AN 19 - Measurement of Nonlinear Thermal Parameters*. Abgerufen am 6. April 2017 von http://www.klippel.de/fileadmin/klippel/Files/Know_How/Application_Notes/AN_19_Nonlin ear_Thermal_Parameters_%28Convection%20Cooling%29.pdf

Klippel GmbH. (15. Juli 2014b). AN 67 - Auralization of signal differences. Abgerufen am 6. April 2017 von http://www.klippel.de/fileadmin/klippel/Files/Know_How/Application_Notes/AN_67%20aur alization%20of%20signal%20differences.pdf

Klippel GmbH. (24. Juli 2014c). *Difference Auralization*. Abgerufen am 2. Mai 2017 von http://www.klippel.de/fileadmin/klippel/Bilder/Our_Products/R-D_System/PDF/S22_Difference%20Auralization.pdf

Klippel GmbH. (25. Februar 2013a). *Hands-On Training 1: Linear Lumped Parameter Measurement*. Abgerufen am 23. März 2016 von https://www.klippel.de/training/attachments/training1/Training_1_Linear_Lumped_Parameter _Measurement_en.pdf

Klippel GmbH. (8. Juli 2014a). *Hands-On Training 3: Loudspeaker Nonlinearities*. Abgerufen am 24. März 2016 von https://www.klippel.de/training/attachments/training3/Training_3_Loudspeaker_Nonlinearitie s_en.pdf

Klippel GmbH. (24. November 2015b). *Large Signal Identification (LSI)*. Abgerufen am 2. Mai 2017 von http://www.klippel.de/fileadmin/klippel/Bilder/Our_Products/R-D_System/PDF/S1_LSI.pdf

Klippel GmbH. (13. August 2012). *Linear Parameter Measurement (LPM)*. Abgerufen am 20. Mai 2017 von http://www.klippel.de/fileadmin/klippel/Bilder/Our_Products/R-D_System/PDF/S2_LPM.pdf

Klippel GmbH. (2017). *Listening Test - How it works*. Abgerufen am 15. Mai 2017 von http://www.klippel.de/listeningtest/lt/?page=how

Klippel GmbH. (3. Februar 2016). *SIM - AUR Simulation/Auralization*. Abgerufen am 2. Mai 2017 von http://www.klippel.de/fileadmin/klippel/Bilder/Our_Products/R-D_System/PDF/S24_SIM-AUR_preliminary.pdf

Klippel GmbH. (13. August 2013b). *Simulation (SIM) Version 2*. Abgerufen am 2. Mai 2017 von http://www.klippel.de/fileadmin/klippel/Bilder/Our_Products/R-D_System/PDF/S3_SIM_Version_%202.pdf

Klippel, W. (7. - 9. März 2016). *Sound Quality of Audio Systems*. Dresden: Technische Universität Dresden.

Klippel, W. (2001). *Auralization – Subjective Evaluation of Speaker Distortion*. Abgerufen am 2. Mai 2017 von http://www.klippel.de/fileadmin/klippel/Files/Know_How/Literature/Papers/Auralization_01. pdf

Klippel, W. (2000). *Diagnosis and Remedy of Nonlinearities in Eletrodynamical Transducers*. Abgerufen am 7. April 2017 von http://www.klippel.de/fileadmin/klippel/Files/Know_How/Literature/Papers/Diagnosis_and_r emedy_of_Nonlinearities_00.pdf

Klippel, W. (2003). *Die Messung impulsiver Signalverzerrungen durch nichtlineare Systemidentifikation*. Abgerufen am 25. April 2017 von https://www.klippel.de/fileadmin/klippel/Files/Know_How/Literature/Papers/Messung_von_i mpulsiven_Signalverzerrungen_03.pdf

Klippel, W. (2006). *Loudspeaker Nonlinearities – Causes, Parameters, Symptoms*. Abgerufen am 3. April 2017 von https://www.klippel.de/fileadmin/klippel/Files/Know_How/Literature/Papers/Klippel_Nonlin earity_Poster.pdf

Klippel, W. (1990). Multidimensional Realtionship between Subjective Listening Impression and Objective Loudspeaker Parameters. In *Acustica Vol.* 70 (S. 45-54). Stuttgart: S. Hirzel Verlag.

Kunz, R. (2014). Untersuchung der Eigenschaften der Wahrnehmungsgröße Rauigkeit. Implementierung und Evaluation eines Berechnungsmodells. Dresden: TU Dresden.

Liebig, M. (2012). Subjektive und objektive Bewertung der Klangqualität im Automobil. Dresden: TU Dresden.

Möser, M. (2012). *Technische Akustik* (9., aktualisierte Ausg.). Berlin Heidelberg: Springer Vieweg.

Müller, S. (2008). Messtechnik. In S. Weinzierl (Hrsg.), *Handbuch der Audiotechnik* (S. 1087 ff.). Berlin Heidelberg: Springer Verlag.

Maier, P. (2008). Studioakustik. In S. Weinzierl (Hrsg.), *Handbuch der Audiotechnik* (S. 267 ff.). Berlin Heidelberg: Springer Verlag.

Martner et. al. (2000). Neues psychoakustisches Modell zur objektiven Bestimmung der Rauigkeit bei Verbrennungsmotoren. *MTZ Motortechnische Zeitschrift* (61), S. 678 ff.

Maschke, C., & Jakob, A. (2010). Psychoakustische Messtechnik. In M. Möser (Hrsg.), *Messtechnik der Akustik* (S. 599 ff.). Berlin Heidelberg: Springer Verlag.

Matthé, F., & Fischer, M. (ohne Jahr). *E-Learning Unterlagen zum Urheberrecht*. Abgerufen am 4. Juli 2017 von http://www.unipotsdam.de/emoon/fileadmin/user_upload/eingangsgrafiken/modul_5/Texte/emoon_Modul_U rheberrecht_Alle_Texte.pdf

Moore, B. C. (2007). Psychoacoustics. In T. D. Rossing (Hrsg.), Springer Handbook of Acoustics (S. 459 ff.). New York: Springer Verlag.

Oberdorfer, G. (1984). *Lehrbuch der Elektrotechnik* (Fünfte, unveränderte Ausg., Bd. 1.). München: Leibniz Verlag.

Schulze, R. (4. April 2013). Implementierung eines Hörmodelles zur Berechnung der Rauigkeit von Kühlschrank- und Wäschetrocknergeräuschen . Dresden, Deutschland: Technische Universität Dresden.

Schulze, R. (2014). *Thermische und nichtlineare Modellierung von Audiosystemen*. Dresden: TU Dresden.

Straube, F. (2013). *Rauigkeitsempfindung bei der Wahrnehmung von Lautsprecherverzerrungen*. Dresden: TU Dresden.

Thiede et. al. (1998). *PEAQ - der künftige ITU-Standard zur objektiven Messung der wahrgenommenen Audioqualität.* Berlin: TU Berlin.

Völk, F. (2015). DAGA Paper. Updated analytical expressions for critical bandwidth and critical-band rate. Nürnberg.

Zeller, P. (2012). Handbuch Fahrzeugakustik (2. Ausg.). Wiesbaden: Vieweg & Teubner.

ANHANG 1

Musikauswahl

| # | Künstler | Album | Titel |
|----|----------------------------------|--|-----------------------------|
| 1 | Avicii | True | Addicted to You |
| 2 | | Wake Me Up - Single | Wake Me Up |
| 3 | Bassida | Marmelade | Love Drug |
| 4 | | | You make me wanna stay |
| 5 | | | Be mine |
| 6 | Beck | Mellow Gold | Loser |
| 7 | Béla Fleck And The Flecktones | Left Of Cool | Sojourn Of Arjuna |
| 8 | Bombay Bicycle Club | So Long, See You Tomorrow | Come To |
| 9 | Boundzound | Boundzound | So Long |
| 10 | | | Marathon Mann |
| 11 | Brother Dege | Folk Songs Of The American Longhair | Old Angel Midnight |
| 12 | | | The Girl Who Wept Stones |
| 13 | C2C feat. Rita J. & Moongai | Tetra | Kings Season |
| 14 | Chris Stapleton | Traveller | The Devil Named Music |
| 15 | | | Was It 26 |
| 16 | | | More of You |
| 17 | | | Traveller |
| 18 | Christina Aguilera | Stripped | Walk Away |
| 19 | | | Get Mine, Get Yours |
| 20 | Coco Rosie | Tales of a GrassWidow | Villain |
| 21 | Eatmorecake | The Red Sky E.P. | Forever Dreaming |
| 22 | | | Smoke & Mirrors |
| 23 | | | All I Need |
| 24 | Esbjörn Svensson Trio | E.S.T. Plays Monk | I Mean You |
| 25 | Florence + The Machine | Ceremonials | Shake It Out |

| # | Künstler | Album | Titel |
|----|------------------------------------|-----------------------------------|-----------------------------|
| 26 | Florence + The Machine | Ceremonials | All This and Heaven Too |
| 27 | | Lungs | My Boy Builds Coffins |
| 28 | Gentleman And The Far East Band | The Cologne Session | Good Days |
| 29 | Glashaus | Drei | Haltet die Welt an |
| 30 | Grace Jones | Hurricane | I'm Crying (Mother's Tears) |
| 31 | Jamiroquai | Dynamite | Talulah |
| 32 | | | Hot Tequila Brown |
| 33 | | | Seven Days In Sunny June |
| 34 | Jan Pyroman | 4 Thumbs Up - Single | 4 Thumbs Up |
| 35 | Johnny Cash | American V: A Hundred Highways | God's Gonna Cut You Down |
| 36 | José González | In Our Nature | Down The Line |
| 37 | Kelis | Tasty | Stick Up |
| 38 | | | Trick Me |
| 39 | Little Dragon | Ritual Union | Crystalfilm |
| 40 | | | Brush the Heat |
| 41 | | | Ritual Union |
| 42 | Marcus Miller | Free | Jean Pierre |
| 43 | | | Free |
| 44 | Massive Attack | Heligoland | Splitting The Atom |
| 45 | | Mezzanine | Angel |
| 46 | | | Teardrop |
| 47 | Matt Andersen | Second Time Around | Lay Down With Me |
| 48 | Moby | Ambient | Piano & String |
| 49 | New York Ska-Jazz Ensemble | Step Forward | Step Forward |
| 50 | Rudimental feat. Angel Haze | Home | Hell Could Freeze |
| 51 | Rudimental feat. Becky Hill | | Powerless |
| 52 | Rudimental feat. Emeli | | Free |

| # | Künstler | Album | Titel |
|-------|--|------------|-----------------|
| | Sandé | | |
| 53 | Rudimental feat. MNEK & Sinead Harnett | Home | Baby |
| 54 | Rudimental feat. Sinead Harnett | | Hide |
| 55 | Sekuoia | Faces | Dissapear |
| 56 | Skrillex feat. Ellie Goulding | Bangarang | Summit |
| 57 | Suff Daddy | Suff Sells | Intro |
| 58 | The Beach Boys | Holland | Sail On, Sailor |
| 59 | The Low Flying Ducks | Almost | Almost |
| 60 | | | Howl |
| Anker | Christina Aguilera | Stripped | Impossible |
DIGITALER ANHANG