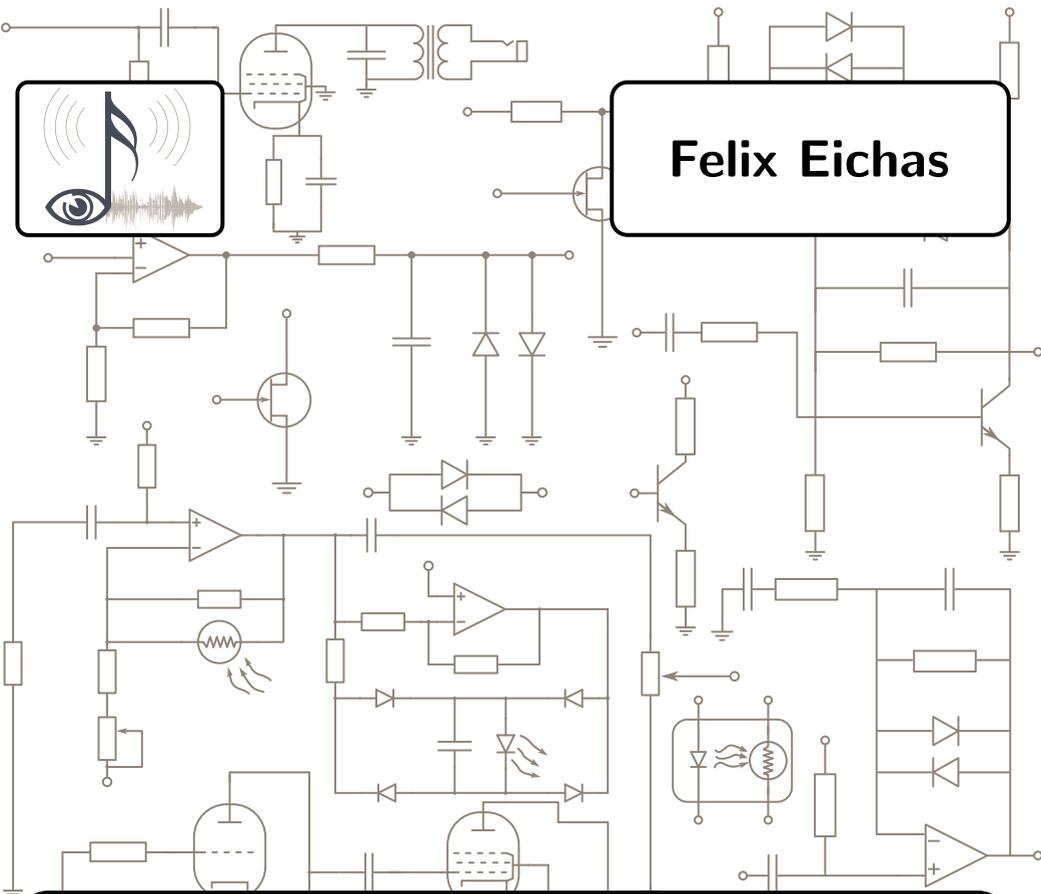
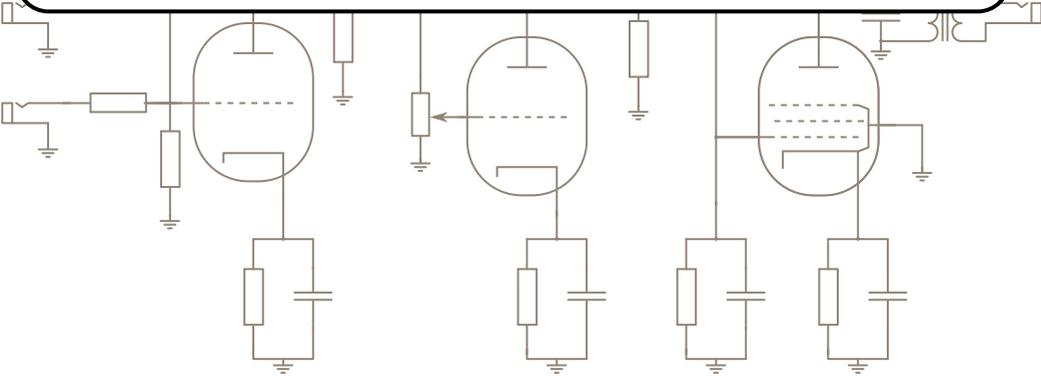




**Felix Eichas**



# **System Identification of Nonlinear Audio Circuits**





# System Identification of Nonlinear Audio Circuits

Systemidentifizierung nichtlinearer Audioschaltungen

Von der Fakultät für Elektrotechnik  
der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg  
zur Erlangung des akademischen Grades eines Doktor-Ingenieurs  
genehmigte

**DISSERTATION**  
vorgelegt von

**Felix Eichas**

**aus Osterode am Harz**

**Hamburg 2019**

Examiners: Prof. Dr.-Ing. Udo Zölzer  
(Helmut-Schmidt-Universität,  
Universität der Bundeswehr Hamburg)

Prof. Dr.Sc. (Tech.) Vesa Välimäki  
(Aalto University, Espoo, Finland)

Date of defense: 24.10.2019

---

# Contents

---

<b>Abstract</b>	<b>ix</b>
<b>Kurzfassung</b>	<b>xi</b>
<b>Lebenslauf</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Basic Principles of System Identification . . . . .	2
1.2 Nonlinear Audio Effects . . . . .	3
1.3 Overview of following Chapters . . . . .	5
<b>2 State of the Art</b>	<b>7</b>
2.1 White-Box Modeling . . . . .	7
2.2 Gray-Box Modeling . . . . .	9
2.3 Black-Box Modeling . . . . .	11
<b>3 Block-Oriented Models</b>	<b>13</b>
3.1 Linear Time-Invariant Blocks . . . . .	13
3.2 Nonlinear Blocks . . . . .	15
3.2.1 Piecewise-Defined Function: Hyperbolic Tangent . .	17
3.2.2 Anti-Aliasing . . . . .	20
3.3 Simple Block-Oriented Models . . . . .	25
<b>4 System Identification</b>	<b>27</b>
4.1 Parameter Optimization: Overview . . . . .	27
4.2 Parameter Optimization Methods . . . . .	29
4.2.1 Gauss-Newton Method . . . . .	29

4.2.2	Gradient-Descent Method . . . . .	30
4.2.3	Levenberg-Marquardt Method . . . . .	31
4.2.4	Finite Differences . . . . .	31
4.2.5	Defined Parameter Space . . . . .	32
4.2.6	Initial Parameter Vector . . . . .	33
4.3	Cost-Functions . . . . .	33
4.3.1	Time-domain Cost-Function . . . . .	34
4.3.2	Envelope Cost-Function . . . . .	34
4.3.3	Time-Frequency Cost-Function . . . . .	35
4.4	Identification of Linear Subsystems . . . . .	37
4.4.1	Equalization of Measurement Equipment . . . . .	39
4.5	Data Acquisition: Hardware Measurements . . . . .	40
4.5.1	Measurement Setup . . . . .	40
<b>5</b>	<b>Evaluation: Error Metrics</b>	<b>43</b>
5.1	Objective Metrics . . . . .	43
5.2	Perceptually Motivated Metrics . . . . .	44
5.2.1	Perceptual Evaluation of Audio Similarity (PEAS) . . . . .	45
5.3	Listening Tests . . . . .	47
5.3.1	ABX – ABC/HR . . . . .	48
5.3.2	MUSHRA . . . . .	48
<b>6</b>	<b>Dynamic Range Compression Systems</b>	<b>49</b>
6.1	Dynamic Range Compression . . . . .	49
6.1.1	Static Behavior . . . . .	50
6.1.2	Dynamic Behavior . . . . .	51
6.2	Case Study I: ‘Flatline Compressor’ Circuit Analysis . . . . .	52
6.2.1	Static Behavior . . . . .	55
6.2.2	Dynamic Behavior . . . . .	57
6.3	Case Study II: UREI 1176 LN . . . . .	60
6.3.1	Principle of Operation . . . . .	60
6.4	Digital Model . . . . .	63
6.5	Modeling . . . . .	68
6.5.1	Input Signals . . . . .	68
6.5.2	Initial Parameter Set . . . . .	69
6.5.3	Modeling Procedure . . . . .	70
6.5.4	Integrating User Controls: UREI 1176LN . . . . .	71
6.6	Results . . . . .	75
6.6.1	Objective Metrics . . . . .	75
6.6.2	UREI 1176LN: Listening Test . . . . .	76
6.7	Conclusion . . . . .	78

<b>7</b>	<b>Basic Distortion Circuits</b>	<b>79</b>
7.1	Circuit Analysis . . . . .	79
7.1.1	Diode Clipper . . . . .	79
7.1.2	Big Muff – Distortion Stage . . . . .	80
7.1.3	Tube Screamer . . . . .	80
7.2	Digital Model . . . . .	82
7.3	Modeling . . . . .	83
7.4	Results . . . . .	85
7.5	Conclusion . . . . .	89
<b>8</b>	<b>Distortion Guitar Effect Pedals</b>	<b>91</b>
8.1	Digital Model . . . . .	91
8.1.1	Linear Blocks . . . . .	92
8.2	Modeling . . . . .	93
8.2.1	Optimization of Nonlinear Parameters . . . . .	94
8.2.2	Optimization of Filter Parameters . . . . .	95
8.2.3	Optimization of All Parameters . . . . .	95
8.3	Results . . . . .	96
8.4	Conclusion . . . . .	99
<b>9</b>	<b>Guitar Amplifiers</b>	<b>101</b>
9.1	Digital Models . . . . .	101
9.1.1	Wiener – Hammerstein Model . . . . .	102
9.1.2	Cascaded Model Topology . . . . .	103
9.2	Parameter Optimization of the Wiener–Hammerstein Model	106
9.2.1	Adapting Second-Order IIR Filters . . . . .	108
9.2.2	Wiener–Hammerstein Model: Optimization of Non- linear Parameters . . . . .	111
9.3	Parameter Optimization of the Cascaded Model . . . . .	112
9.3.1	Cascaded Model: Filter Optimization . . . . .	112
9.3.2	Cascaded Model: Optimization of Nonlinear Param- eters . . . . .	113
9.4	Results . . . . .	114
9.4.1	Wiener–Hammerstein Model . . . . .	114
9.4.2	Cascaded Model . . . . .	115
9.4.3	Objective Scores: Cascaded Model . . . . .	117
9.4.4	Objective Scores: Wiener–Hammerstein Model . . . . .	119
9.4.5	Listening Test: Cascaded Model . . . . .	120
9.4.6	Listening Test: Wiener–Hammerstein Model . . . . .	123
9.5	Conclusion . . . . .	126
<b>10</b>	<b>Summary</b>	<b>129</b>
<b>11</b>	<b>List of selected Symbols</b>	<b>133</b>



---

## Abstract

---

Digital systems gain more and more popularity in today's music industry. Musicians and producers are using digital systems because of their advantages over analog electronics. They require less physical space, are cheaper to produce and are not prone to aging circuit components or temperature variations. Furthermore, they always produce the same output signal for a defined input sequence. However, musicians like vintage equipment. Old guitar amplifiers or legendary recording equipment are sold at very high prices. Therefore, it is desirable to create digital models of analog music electronics which can be used in modern digital environments.

This work presents an approach for recreating nonlinear audio circuits using system identification techniques. Measurements of the input- and output-signals from the analog reference devices are used to adjust a digital model treating the reference device as a 'black-box'. With this technique the schematic of the reference device does not need to be known and no circuit elements have to be measured to recreate the analog device.

An appropriate block-based model is chosen, depending on the type of reference system. Then the parameters of the digital model are adjusted with an optimization method according to the measured input- and output-signals. The performance of the optimized digital model is evaluated with objective scores and listening tests.

Two types of nonlinear reference systems are examined in this work. The first type of reference systems are dynamic range compressors like the 'MXR Dynacomp', the 'Aguilar TLC', or the 'UREI 1176LN'. A block-based model describing a generic dynamic range compression system is chosen and an automated routine is developed to adjust it. The adapted digital models are evaluated with objective scores and a listening test is performed for the UREI 1176LN studio compressor.

The second type of nonlinear systems are distortion systems like e.g. amplifiers for electric guitars. This work presents novel modeling approaches for different kinds of distortion systems from basic distortion circuits which

can be found in distortion pedals for guitars to (vintage) guitar amplifiers like the ‘Marshall JCM900’, or the ‘Fender Bassman’. The linear blocks of the digital model are measured and used in the model while the non-linear blocks are adapted with parameter optimization methods like the Levenberg–Marquardt method. The quality of the adjusted models is evaluated with objective scores and listening tests.

The adjusted digital models give convincing results and can be implemented as real-time digital versions of their analog counterparts. This enables the musician to save a snapshot of a certain sound and recall it anytime with a digital system like a VST plug-in or as a program on a dedicated hardware.

---

## Kurzfassung

---

Digitale Systeme werden immer beliebter in der heutigen Musikindustrie. Musiker und Produzenten nutzen digitale Systeme wegen ihrer Vorteile gegenüber analoger Elektronik. Sie brauchen weniger Platz, sind billiger herzustellen und sind nicht anfällig für Temperaturschwankungen oder die Auswirkungen alternder Schaltungskomponenten. Weiterhin reproduzieren sie immer das gleiche Ausgangssignal für eine definierte Eingangssequenz. Trotzdem mögen Musiker altes Equipment. Alte Gitarrenverstärker oder legendäres Studio-Equipment werden zu sehr hohen Preisen verkauft. Deswegen ist es wünschenswert digitale Modelle analoger Musikelektronik zu erzeugen, die in modernen, digitalen Umgebungen genutzt werden können.

In dieser Arbeit wird ein Ansatz zur modellierung nichtlinearer audio Schaltungen präsentiert, der die Methoden der Systemidentifizierung nutzt. Messungen der Ein- und Ausgangssignale an den analogen Referenzgeräten werden genutzt um das digitale Modell anzupassen. Damit wird das Referenzgerät als sogenannte 'black-box' behandelt. Dadurch werden weder der Schaltplan, noch Messungen von Schaltungskomponenten benötigt, um ein digitales Modell zu erzeugen.

Um ein digitales Modell zu erzeugen, muss ein geeignetes block-basiertes Modell gewählt werden, welches abhängig vom Typ des Referenzgerätes ausgewählt wird. Dann werden die Parameter des Modells mit Hilfe eines Optimierungsverfahrens angepasst. Die Performanz des optimierten digitalen Modells wird mit objektiven Zahlenwerten und Hörtests bewertet.

Zwei Typen von nichtlinearen Systemen werden in dieser Arbeit behandelt. Der erste Typ von Referenzsystemen sind Dynamikkompressoren wie etwa der 'MXR Dynacomp', der 'Aguilar TLC' oder der 'UREI 1176LN'. Ein block-basiertes Modell, welches einen generischen Dynamikkompresor beschreibt, wird ausgewählt und eine automatisierte Routine entwickelt um dessen Parameter zu optimieren. Das angepasste Modell wird mit objektiven Metriken bewertet und ein Hörtest für den UREI 1176LN Stu-

diokompressor wird durchgeführt.

Die zweite Art nichtlinearer Systeme sind Verzerrungssysteme wie z.B. Verstärker für elektrische Gitarren. In dieser Arbeit werden neue Modellierungsansätze für verschiedene Typen von Verzerrungssystemen von grundlegenden Verzerrungsschaltungen, wie man sie in Verzerrungspedalen für Gitarren findet, bis hin zu Gitarrenverstärkern, wie dem ‘Marshall JCM900’ oder dem ‘Fender Bassman’, präsentiert. Die linearen Blöcke der digitalen Modelle werden gemessen und direkt im Modell verwendet, während die nichtlinearen Modellblöcke mit Parameteroptimierungsverfahren wie dem Levenberg–Marquardt Verfahren optimiert werden. Die Qualität der angepassten Modelle wird durch objektive Metriken und Hörtests ermittelt.

Die optimierten Modelle liefern überzeugende Ergebnisse und können als echtzeitfähige Versionen ihrer analogen Gegenstücke implementiert werden. Dies ermöglicht dem Musiker eine Momentaufnahme von einem bestimmten Klang zu speichern und jederzeit in Form eines digitalen Systems wie etwa eines VST Plug-Ins oder als Programm auf einer dedizierten Hardware abzurufen.

---

## Lebenslauf

---

### 1. Persönliche Daten

---

Geboren am: *29.08.1986*

Geboren in: *Osterode am Harz*

### 2. Schule & Ausbildung

---

Allgemeine Hochschulreife *2006*

Wilhelm-Raabe Schule *Hannover*

### 3. Hochschulstudium

---

B.Sc. Elektrotechnik *2007-2011*

Technische Universität Hamburg *Hamburg*

M.Sc. Elektrotechnik (Nachrichtentechnik) *2011-2013*

Technische Universität Hamburg *Hamburg*

### 4. Berufstätigkeit

---

Wissenschaftlicher Mitarbeiter *Aug. 2013 - Sep. 2019*

Helmut-Schmidt-Universität *Hamburg*



---

## Introduction

---

The digitization of music electronics is gaining more and more importance in today's music industry. Digital products emulating analog audio systems are already in widespread use. Vintage audio equipment is very popular among musicians and producers and it is desirable to recreate popular analog devices for digital music production. Digitizing a specific analog circuit or device is called **virtual analog modeling** [1–79]. This enables musicians to use digitized versions of their favorite equipment without the need to acquire the physical device, which can reach absurdly high costs because of its age and scarcity value.

Modeling nonlinear audio systems is especially challenging. State-of-the-art circuit modeling techniques require knowledge about the schematic of the analog reference system as well as the characteristics of all nonlinear circuit elements, like e.g. diodes, transistors or vacuum tubes. Furthermore, the equations describing the nonlinear reference system have to be solved iteratively which can lead to a high computational load during run-time [22, 42, 43, 45, 47, 53, 55, 59–63, 67, 71, 76, 80–85].

This is the reason for investigating a **system identification** [86–92] approach in this work. Nonlinear analog systems are digitized using input/output measurements to adjust a digital model recreating the output of the physical device. This reference device can be treated as a ‘black-box’ or ‘gray-box’ where no knowledge about the schematic or characteristics of circuit elements is needed and a computationally efficient digital model can be used [91, 93–96].

Compared to its analog counterpart, the digital model has some advantages. For example, some components of an analog system can be prone

to temperature variations or aging and alter the characteristics of one or more circuit elements therefore changing the characteristic of the whole system. A digital system always produces the same output for a given input, independent of circuit elements or temperature fluctuations. Another advantage is the cost for a system. Each analog device always has to be assembled and physically distributed before it can be used. A digital solution could be downloaded and used right away.

## 1.1 Basic Principles of System Identification

The analog reference system needs to be analyzed and/or measured to create a virtual analog model. This process is called **system identification** and can be carried out in different ways. There exist three basic principles when using system identification in any modeling process [89, 91, 92]:

- **White-Box Modeling**

White-box modeling is used when comprehensive knowledge about the reference system exists. The model can be created from prior knowledge (e.g. the schematic of the reference system) and physical insight (e.g. measured characteristics of nonlinear circuit elements).

- **Gray-Box Modeling**

Gray-box modeling is used when some physical insight exists, but other characteristics still need to be determined. There are two main cases for this approach:

1. **Physical modeling**

If the model structure is physically motivated but some parameters of the model need to be determined by input/output measurements on the reference system.

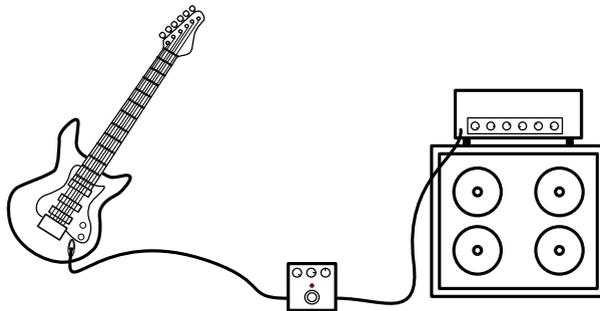
2. **Semi-physical modeling**

Physical insight is used to structure the modeling process. For example the choice of input signals or the model structure is based on physical observations but the adjustment of parameters is carried out in a black-box modeling way.

- **Black-Box Modeling**

No physical insight is used at all. A highly flexible model (or set of models) is adapted by using solely input/output measurements on the reference device.

There is no universal law which method should be used in which case. It strongly depends on the system which has to be modeled, the availability of information about the system, the use case of the model, and the available computational power. In this work a semi-physical gray-box



**Figure 1.1**

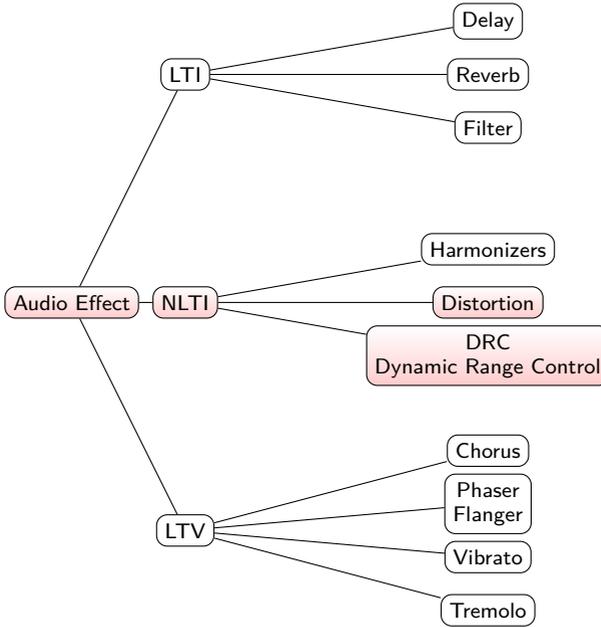
*Typical signal-chain for an electric guitar setup. The guitar (left) is connected to an effect device (middle), which is connected to the amplifier (right).*

approach is used because prior knowledge about the reference system is used to design the digital model and the input signals. Afterwards, the parameters of the model are adapted by an iterative optimization method. Pure black-box modeling is very rarely carried out because in most cases some physical insight about the reference system already exists and can be used to facilitate the identification.

## 1.2 Nonlinear Audio Effects

Many different kinds of audio effects can be used to alter the sound of an instrument like an electric guitar, which is the focus of this work. To apply the effect, the electric guitar is connected to a device which can be turned on by stepping on a foot switch enabling musicians to use their hands freely while operating the device. For this reason these types of effects are also called ‘stompboxes’. Guitar amplifiers are another type of analog device because they do not only amplify the sound of the guitar but also introduce filtering and harmonic distortion, creating their characteristic sound. A typical use-case of the signal-chain for an electric guitar setup is shown in Fig. 1.1.

Audio effects specific to altering the sound of an electric guitar can be categorized into three main groups [97, 98]: Linear time-invariant (LTI) effects like delay, reverb, and filters. Linear time-variant (LTV) effects like flanger, phaser, chorus, vibrato, or tremolo, and nonlinear time-invariant



**Figure 1.2**

*Categorization of typical audio effects into linear time-invariant (LTI), non-linear time-invariant (NLTI), and linear time-variant (LTV) effects.*

(NLTI) effects as depicted in Fig. 1.2. Linear effects will not be analyzed and modeled in this work.

Nonlinear guitar effects can be classified into three main groups: *distortion*, *dynamic range control* and *harmonizers*. Harmonizers generate one or several harmonics or subharmonics, typically (but not exclusively) one octave above or below the fundamental frequency of the input signal. Harmonizers will not be discussed in this work.

Dynamic range control systems like e.g. compressors reduce the dynamic range of a signal by attenuating larger signal levels according to a nonlinear characteristic [99]. Distortion systems like e.g. guitar amplifiers provide filtering and enrich the input signal by adding harmonic overtones [100, 101]. These two types of audio systems are analyzed in detail in this work and a modeling approach for each system is proposed.

Dynamic range control systems and distortion systems can be modeled with block-oriented models consisting of linear time-invariant blocks (filters) and nonlinear mapping functions which map the amplitude of their input signal to the amplitude of their output signal according to a non-

linear function. The circuit of a device is analyzed and a digital model consisting only of linear and nonlinear blocks is constructed to recreate the function principle of the reference devices circuit. Afterwards, the parameters of the constructed model are optimized to recreate the behavior of a specific reference device.

### 1.3 Overview of following Chapters

This section gives a brief overview of the following chapters. The state of the art is described in Chap. 2, with white-box and gray-box modeling approaches. Block-oriented models with linear and nonlinear blocks are presented in Chap. 3. The algorithms used to adapt the parameters of block-oriented models are explained in Chap. 4 as well as the used error functions or cost-functions and the method of data acquisition. The error metrics which were used to evaluate an adjusted digital model are presented in Chap. 5.

Chap. 6 describes the modeling of dynamic range compression systems by designing an appropriate model and adjusting it with measured data and iterative optimization. Two case studies have been performed before proposing a modeling procedure for dynamic range compression systems in general. An approach on integrating the user controls of a specific device is investigated and the results are evaluated with objective metrics and a listening test.

Chapters 7 – 9 describe the investigations toward modeling distortion circuits with a special emphasis on distortion circuits for electric guitars. In Chap. 7 relevant parts of popular circuits are analyzed and modeled with a simple block-oriented model. In the following chapter, this is extended to distortion pedals (Chap. 8). The Wiener–Hammerstein model is used, extending the model from the previous chapter by one linear block. Finally the modeling of whole guitar amplifiers is presented in Chap. 9 with a cascaded block-oriented model using two nonlinear blocks and the Wiener–Hammerstein model.

In Chap. 10 the work is summarized.



---

## State of the Art

---

In this chapter the existing methods to recreate analog audio circuits are compared and their advantages and disadvantages are illustrated. Three main modeling approaches exist: white-box modeling, gray-box modeling and black-box modeling.

### 2.1 White-Box Modeling

White-box modeling of musical distortion circuits is already well researched. One white-box approach is based on a state-space representation of the audio circuit. Several Ph.D. theses have been written about this topic [102–104]. Further investigations, case studies and refinements of the method have been made by Holters et al. [42–46, 55, 65, 69, 71, 76, 85, 105, 106]. All of these works have in common that a state-space system is constructed from the circuit diagram using some form of circuit analysis. The state-space representation is extended to be able to handle nonlinear circuit elements. The states are used to describe the energy storing circuit elements such as e.g. capacitors and inductors. This circuit-based approach achieves very convincing results and the digital model is perceptually nearly indistinguishable from the analog reference device, which can be confirmed through objective measurements and listening tests.

Another white-box modeling approach makes use of the wave-digital principle [107]. Thereby an electrical circuit is represented in terms of wave-digital filters. The method is based on a representation in terms of wave-variables (incident and reflected waves) instead of Kirchoff-variables

(voltage and current) as in the state-space modeling approach. This approach resulted in several Ph.D. thesis [108–110]. In [80], a real-time capable model of a transformer in a tube amplifier is developed. The emulation of operational amplifiers and diodes using wave digital filters is described in [53]. Standard wave digital filter techniques are expanded in [111] to be able to handle nonlinear elements with a case study of a common-cathode triode stage of a guitar amplifier. A new triode model is presented in [81].

Werner et al. also made huge contributions to the field of wave digital filters by introducing techniques to include complex circuit topologies and multiport linear elements in [83] and resolving the issue of including multiple and multiport nonlinearities in a wave digital filter structure in [82]. In [84] the circuit of a vintage guitar pre-amplifier with multiple nonlinearities and a complex topology was modeled with this technique giving convincing results.

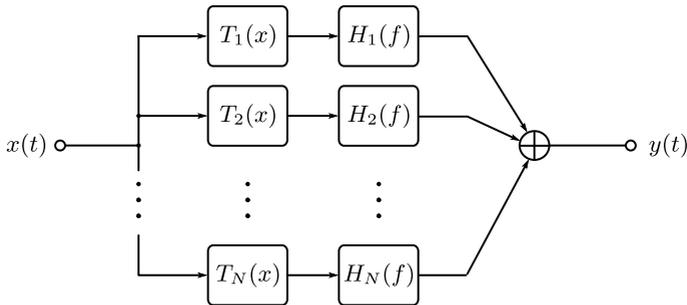
The third white-box modeling approach are port-Hamiltonian systems [59,112,113] which are a state-space description of power exchange between energy-storing parts, dissipative parts and external sources in a circuit. Port-Hamiltonian systems can not only be used to model analog circuits. They can also be used to model physical systems, like e.g. a nonlinear loudspeaker in a guitar amplifier [114]. In this case port-Hamiltonian systems are beneficial because they naturally preserve the power balance also in the nonlinear case.

All white-box approaches perform very well and they can reproduce all relevant characteristics of a reference device. If the sound of a specific analog device should be replicated with high accuracy, white-box modeling is preferable. But these approaches also have drawbacks. To create the digital model the circuit diagram has to be known, as well as the characteristic curves of every nonlinear circuit element, e.g. diodes, transistors, transformers or vacuum tubes. If no circuit diagram is obtainable, time-consuming reverse engineering of the circuit has to be performed, as described in [115]. Another drawback of the white-box modeling methods is the computational effort which arises due to nonlinear circuit elements. For every nonlinear circuit element at least one nonlinear equation has to be solved per time step. Depending on the nonlinear solver and the initial parameter set, this can drastically influence the computational load of the digital model. Although Holmes et al. described a method for improving the nonlinear solver [116], the computational load is still high, especially for complex circuits with multiple nonlinearities. Additionally Holters proposed a method to automatically decompose the large state-space matrices into smaller ones to be able to solve the system quicker [105]. Several other publications address the computational load of the white-box approaches and propose improvements to ensure real-time capability like [42, 49, 80, 85, 106, 113, 117]. A complete guitar amplifier circuit could be modeled with these approaches but without extensive optimization and

simplification the digital model would not be real-time capable.

## 2.2 Gray-Box Modeling

A big contribution to the identification of nonlinear audio systems was made by Novák [19, 20, 35–37, 64, 66, 118, 119] who applied several well-known system identification methods and used them in a gray-box modeling approach. His research included (simplified) Volterra series models as well as Wiener and Hammerstein models. The basic idea behind this approach is a nonlinear system decomposition in fundamental and harmonic signal components, which will be represented by the branches of the model. For example, in [119] a multi-branch Hammerstein model was used to model a distortion effect pedal. The model is depicted in Fig. 2.1. In each branch, the input signal  $x(t)$  is processed with a polynomial  $T_n(x)$  and the resulting output is filtered with the filter  $H_n(f)$ . The output of all branches is summed up afterwards and forms the output signal  $y(t)$ .



**Figure 2.1**

*Multi-branch Hammerstein model according to [119].*

The identification procedure is very elegant because only one exponentially swept sine wave is used to identify the system. Afterwards the filters for each branch of the model can directly be extracted from the resulting impulse response, which does not only include the systems response for the linear part but also the impulse responses for the generated harmonics [120].

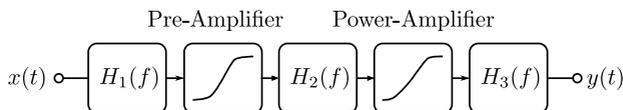
The nonlinear elements in each branch are the Chebyshev polynomials. This method yields nearly perfect results, but is restricted to the amplitude of the identification signal. Unfortunately this is a big limitation, since electric guitars have more than one output amplitude. The complexity of the model also grows with a more nonlinear reference system. If the reference device is highly nonlinear the model needs a lot of branches,

resulting in a high computational load. In [121] the complexity of the model is reduced by principal component analysis.

Block-oriented or block-based models are often used for the simulation of distortion systems which originate from the well-known Volterra series representation of nonlinear systems. In [122] an overview of simple block-oriented models and the corresponding equivalent Volterra kernels is shown. In [123] a Hammerstein model, which is a memoryless nonlinear block followed by a linear block (filter) is used to model a guitar amplifier. Primavera et al. proposed a measurement technique to classify whether the nonlinear system should rather be modeled by a Wiener or a Hammerstein model [124]. In [79] a parametric nonlinear transfer function for Wiener–Hammerstein models is investigated which is based on the Ebers–Moll equation. In [56] a uniformly distributed noise sequence is used to identify Hammerstein and Wiener systems. Furthermore, block-oriented models are also applied in the field of control theory for nonlinear system modeling [125].

In [126] a completely parametric Wiener–Hammerstein model has been used to model guitar distortion pedals. In [127] a Wiener model is used to model parts of famous distortion circuits and in [128] a cascaded block-oriented model is used to model guitar amplifiers using mapping functions based on polynomials and hyperbolic tangents.

Furthermore, gray-box modeling is used in commercial products. A block-oriented approach to model guitar amplifiers is presented in [129]. The model is constructed from linear time-invariant (LTI) blocks and nonlinear mapping functions, relating input to output amplitude by a nonlinear function (e.g. a hyperbolic tangent or a polynomial).



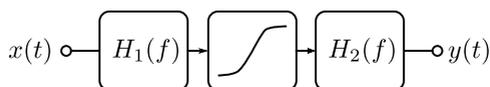
**Figure 2.2**

*Block-oriented guitar amplifier model with linear time-invariant blocks (filters) and nonlinear blocks (mapping functions).*

A model of a guitar amplifier can be constructed with this method by concatenating linear and nonlinear blocks, as shown in Fig. 2.2.  $H_1(f)$  to  $H_3(f)$  represent the LTI blocks and the nonlinear blocks are depicted as input–output mapping curves. The first nonlinear block of the model corresponds to the pre-amplifier nonlinearity and the second nonlinear block corresponds to the power-amplifier nonlinearity of the guitar amplifier.

The model has to be tuned manually by an experienced audio engineer to assure that it sounds the same as the analog reference device. This process can become quite tedious and time consuming requiring measurements of different circuit stages and extensive manual tuning.

While there is no automated method to tune a digital model of a guitar amplifier described in scientific publications, there exists a commercial product which is able to adjust its digital models parameters to recreate the measured sound of a guitar amplifier. It is described in a patent [130], which only vaguely specifies the technical details of the method. Nevertheless, it becomes clear that the used model is a Wiener-Hammerstein model, consisting of an LTI block in series with a memoryless nonlinear mapping function in series with another LTI block, as depicted in Fig. 2.3.



**Figure 2.3**

*Wiener-Hammerstein model, consisting of a memoryless nonlinear mapping function sandwiched between two filters (LTI).*

In [129] the Wiener-Hammerstein structure is described as ‘the fundamental paradigm of electric guitar tone’. The ‘Kemper profiling amplifier’, [130] uses this principle as the model structure and adjusts the internal model, informed by input-output measurements. The model is able to recreate the whole signal chain, from amplifier input to measured signal at the speaker cabinet.

The aim of this work is to analyze how well block-oriented models will perform when they are adjusted with an automated procedure. Therefore an automatic block-oriented gray-box modeling approach is systematically investigated. For this purpose different nonlinear circuits and systems are analyzed, the model is optimized with an automated procedure and the adjusted model is evaluated. The model for each type of reference system is built by recreating the general structure of the reference device. Afterwards the blocks of the model are adjusted by measuring them directly (if possible) and iteratively optimizing the parametric parts of the model.

## 2.3 Black-Box Modeling

Black-box modeling is performed when no information about the reference system is available and a generic model is adjusted using only input/output measurements. Neural networks can be seen as generic models for processing audio data and they have already been used to emulate nonlinear guitar

systems.

Early works with neural networks included the use of recurrent neural networks (RNNs) which use their internal memory to compute the current output. However, the results of first studies were unsatisfactory and exhibited relatively high error values [131, 132].

In recent years an advanced class of RNNs called long short term memory (LSTM) was used for modeling nonlinear audio systems. Several publications describe the use of LSTMs for guitar amplifier modeling with very good results [133–136].

Different network architectures are also used for distortion system modeling. The first is based on the WaveNet [137] convolutional neural network and was successfully used by Damskäg et al. for tube-amplifier modeling [138, 139]. The second uses end-to-end deep neural networks (DNNs) to model nonlinear audio systems [140].

Computing the output of a neural network in real-time is no trivial task. In [78, p. 158] an overview of computational times for different models is given, showing that not all of the presented models are real-time capable. Wright et al. also tackle the issue of real-time capability by using RNNs instead of the WaveNet architecture without deteriorating the results of the modeling procedure [136].

---

## Block-Oriented Models

---

A block-oriented model describes an arbitrary configuration of memoryless nonlinear blocks and linear time-invariant blocks. These models are already well-researched and have been used to digitally model nonlinear systems in several disciplines e.g. control theory [90]. This chapter introduces the basic components of block-oriented models: linear time-invariant blocks and memoryless nonlinear blocks. Countermeasures against aliasing are discussed and simple block-oriented models in the context of nonlinear audio systems are presented.

### 3.1 Linear Time-Invariant Blocks

Audio filters or linear time-invariant blocks have different topologies. The functionality of a block-oriented model will not be impaired by the choice of digital filter implementation. However, the filters which were used in this work are presented in the following.

Figure 3.1 illustrates an exemplary realization of two of many different filter topologies. The finite impulse response filter (FIR) (a) and the direct form 1 (DFI) (b). The digital input and output signals are denoted as  $x(n)$  and  $y(n)$  respectively and a block labeled with  $z^{-1}$  represents a delay of the signal by one sampling interval.

For the filter shown in Fig. 3.1 (a), only the input samples are used to calculate the filter output. The filter has a finite impulse response and is thus called finite impulse response (FIR) filter. This filter can be used to apply a measured impulse response which can be digitized and directly

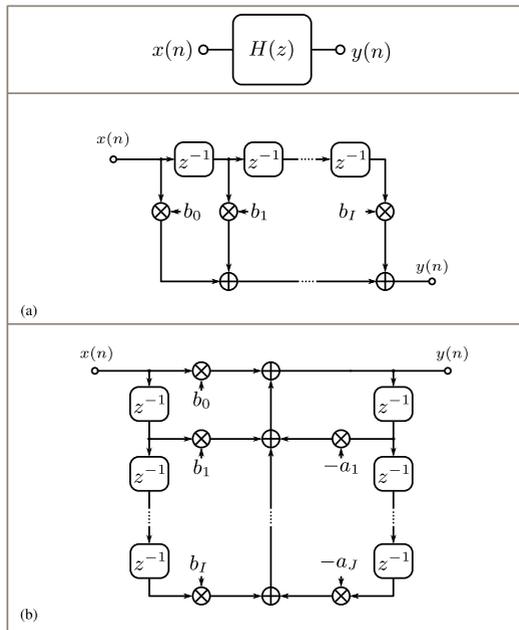
used as the coefficients of the filter.

If previous output samples are used to calculate the output of the filter, the impulse response could be infinitely long, depending on the value of the denominator coefficients, which leads to infinite impulse response (IIR) filters.

One example of an IIR filter is the direct form 1 (DF1) topology (b) which can be described by the function

$$y(n) = \sum_{i=0}^I b_i x(n-i) + \sum_{j=1}^J a_j y(n-j) . \quad (3.1)$$

The same function can be used for an FIR filter when all output coefficients are zero ( $\forall j, a_j = 0$ ).



**Figure 3.1**

*Different filter topologies. Finite impulse response filter (FIR) (a) and direct form 1 (DF1) (b).*

This leads to the transfer function

$$H(z) = \frac{Y(z)}{X(z)} = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_I z^{-I}}{1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_J z^{-J}} , \quad (3.2)$$

where  $I$  and  $J$  are the orders of numerator and denominator.

If the order of a DFI filter is high, the poles can be very close to the unit circle which leads to unstable behavior. To circumvent these stability issues it is common practice to either reduce the order of the used filters [141] or use second-order filters, also called biquads, and connect them in series or in parallel to achieve the desired frequency response [142, 143].

Filter designs often result in the coefficients for second order IIR filters [98, 144] and are useful because a combination of second-order sections which is tailored to the problem at-hand can be described by a few control parameters, like e.g. cutoff frequencies, gains or quality factors.

## 3.2 Nonlinear Blocks

A nonlinear block can be implemented as a memoryless mapping function, which relates its output amplitude  $y$  to its input amplitude  $x$  via a nonlinear function  $g(x)$ . A nonlinear function will add harmonic overtones to a sinusoidal signal which leads to a distorted output signal. The shape of the nonlinear function determines the spectral shape of the harmonics. This behavior is depicted in Fig. 3.2 for three different mapping functions. The topmost plot shows an arc tangent mapping function,

$$g(x) = \text{atan}(a \cdot x) \quad , \quad (3.3)$$

the upper middle plot shows a hyperbolic tangent,

$$g(x) = \tanh(a \cdot x) \quad , \quad (3.4)$$

the lower middle plot shows asymmetrical clipping with a diode characteristic from the Shockley diode equation [145],

$$g(x) = I_s \left( e^{\frac{x}{(n_i \cdot v)}} - 1 \right) \quad , \quad (3.5)$$

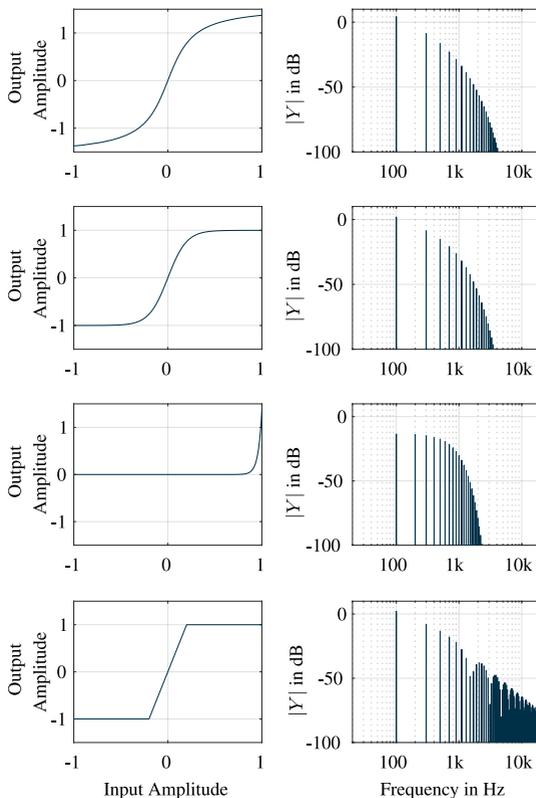
with  $I_s = 10^{-12}$  A as the saturation current of the diode,  $n_i = 1$  as its ideality factor and  $v = kT/q \approx 25.8563$  mV as the thermal voltage with  $T = 300$  K as the temperature in Kelvin. Furthermore,  $k$  is the Boltzmann constant and  $q$  is the elementary charge.

The lowermost plot shows hard clipping.

$$g(x) = \begin{cases} 1, & \text{for } (a \cdot x) > 1 \\ a \cdot x, & \text{for } -1 < (a \cdot x) < 1 \\ -1, & \text{for } (a \cdot x) < -1 \end{cases} \quad (3.6)$$

The gain-parameter was set to  $a = 5$  for all the above equations.

All mapping curves have a linear region around  $x = 0$  and a saturation region for  $|x| \approx 1/a$ . The transition from the linear region to the saturated



**Figure 3.2**

*Different mapping functions. The arc tangent (top), the hyperbolic tangent (upper middle), asymmetrical diode clipping (lower middle) and hard clipping (bottom). The mapping functions are depicted on the left and the spectrum of a 100 Hz sine wave after processing with each mapping function on the right.*

region determines the number and shape of the harmonics. The arc tangent and the hyperbolic tangent exhibit soft transitions from linear to saturated region. The mapping curves are shaped similarly and so are the harmonics they produce.

Due to the symmetric properties of all these mapping functions, only odd harmonics are created. Even harmonics can be created by asymmetrical clipping as shown in the lower middle plot of Fig. 3.2. There is harmonic content at even multiples of the fundamental frequency of  $f_0 = 100$  Hz, which is missing for all symmetric mapping curves.

The hard clipping curve has an abrupt transition between the linear and the saturated region. This creates a lot of harmonics which do not roll-off as fast as for the softer clipping functions, as can be seen in the lowermost spectrum of Fig. 3.2.

### 3.2.1 Piecewise-Defined Function: Hyperbolic Tangent

The properties of the aforementioned nonlinear mapping functions can be combined by using a piecewise-defined function based on e.g. the hyperbolic tangent. The proposed function consists of three hyperbolic tangents which are defined by two connection points for positive and negative input amplitudes. The middle part around  $x = 0$  is an unmodified hyperbolic tangent. Two parameters  $k_p$  and  $k_n$  denote the abscissa of the connection points for positive and negative amplitudes. Above the connection point  $k_p$  a scaled and shifted hyperbolic tangent is connected to the middle part. The same is done for negative input amplitudes below  $k_n$ . This results in the possibility to shape positive and negative half-wave of the input signal separately. For values between  $-k_n$  and  $k_p$  the unmodified hyperbolic tangent is used,

$$m(x) = \begin{cases} a_p \cdot \tanh(g_p \cdot (x - k_p)) + b_p & \text{if } x > k_p \\ \tanh(x) & \text{if } -k_n \leq x \leq k_p \\ a_n \cdot \tanh(g_n \cdot (x + k_n)) + b_n & \text{if } x < -k_n \end{cases} . \quad (3.7)$$

The shifted versions for positive and negative amplitudes are shifted by  $k_p$  or  $-k_n$  in  $x$ -direction and by  $b_p$  or  $b_n$  in  $y$ -direction. The slope of the shifted hyperbolic tangents is modified so that no discontinuities occur in the first derivative of  $m(x)$  resulting in the parameters,

$$a_p = \frac{1 - \tanh(k_p)^2}{g_p} , \quad (3.8)$$

$$b_p = \tanh(k_p) , \quad (3.9)$$

$$a_n = \frac{1 - \tanh(k_n)^2}{g_n} , \quad (3.10)$$

$$b_n = -\tanh(k_n) . \quad (3.11)$$

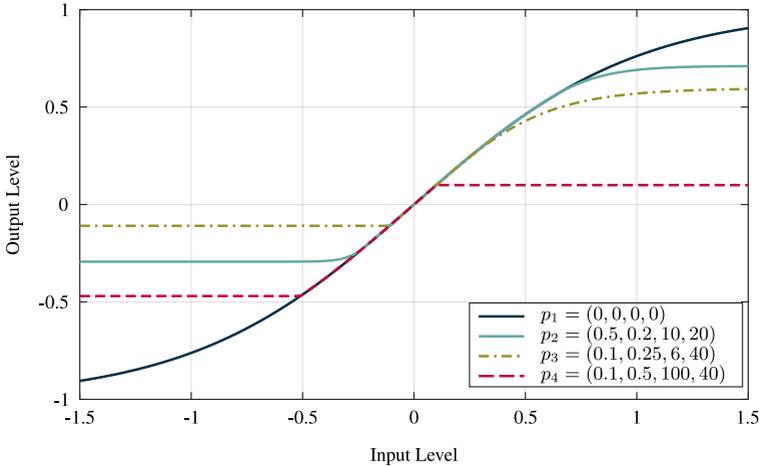
This allows a multitude of different mapping functions, because the second parameters for each half-wave  $g_p$  and  $g_n$  can be used to adjust how fast the mapping function saturates after the input amplitude exceeds the

connection points. When the parameters for positive and negative amplitudes have different values the whole function becomes asymmetrical which will result in even harmonics whose magnitude can be scaled by the difference in the parameters.

In Fig. 3.3  $m(x)$  is plotted for different parameter sets. The parameters are combined in a parameter vector  $\mathbf{p}_{\text{map}} = (k_p \ k_n \ G_p \ G_n)^T$ . The gains  $G_p$  and  $G_n$  are given in dB with

$$g_{\text{lin}} = 10^{(G_{\text{dB}}/20)}. \quad (3.12)$$

The curve corresponding to parameter set  $\mathbf{p}_1$  is for connection points  $k_p = k_n = 0$  and with gains of  $G_n = G_p = 0$  dB, which leads to an unmodified hyperbolic tangent with only odd harmonics (see Fig. 3.2). All other curves are configured to be asymmetrical so they will also produce even harmonics.

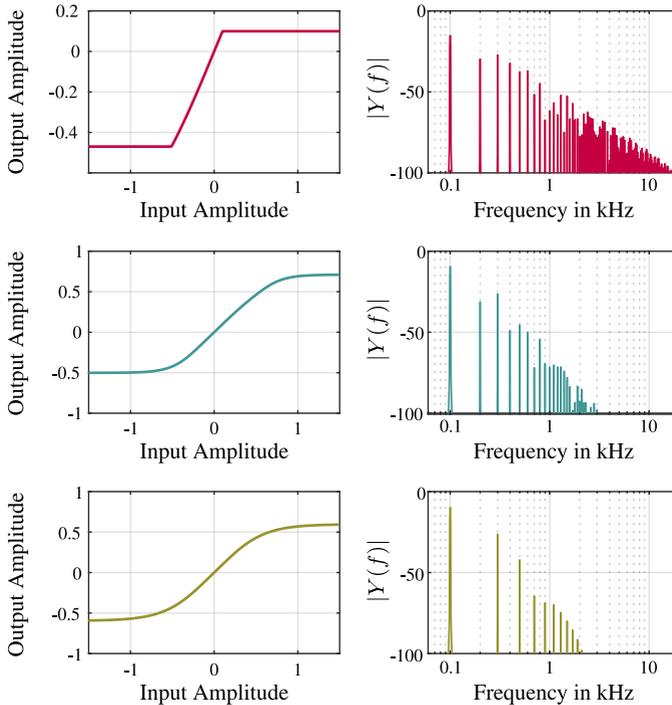


**Figure 3.3**

*Mapping function  $m(x)$  with different parameters ordered  $[k_p, k_n, G_p, G_n]$*

The flexibility of this mapping function is shown by Fig. 3.3. Four different curves with varying parameters are shown. The overall shape of the positive region, with positive input and output amplitudes, can be determined by the combination of the parameters  $k_p$  and  $G_p$  and the negative region with  $k_n$  and  $G_n$ , respectively. The lower the gain values are, the smoother the transition from linear to saturated region, creating a ‘smooth’ or ‘creamy’ distortion. For gain values above 40 dB the transition is sharp which leads to a ‘harsh’ or ‘rough’ sounding distortion. By

modifying the slope of the piecewise-defined function at the connection points, the harmonics produced by  $m(x)$  will not be as drastic as for the hard clipping function.



**Figure 3.4**

*Mapping function  $m(x)$  differently parametrized (left) with the corresponding output magnitude spectra for a sine wave with fundamental frequency of  $f_0 = 100$  Hz (right).*

The flexibility of the piecewise-defined function can be demonstrated by visualizing its output spectra for different parameter sets as shown by Fig. 3.4. The top row shows the parameter set  $p_4$  from Fig. 3.3

$$k_p = 0.1, \quad k_n = 0.5, \quad G_p = 100 \text{ dB}, \quad G_n = 40 \text{ dB} .$$

The connection points are asymmetrical and the gains are set to high values. The output spectrum for an input sine wave with fundamental frequency  $f_0 = 100$  Hz is shown on the right-hand side. A lot of harmonic overtones are produced not only at odd multiples of the fundamental fre-

quency but also for even harmonics. Due to the high gain values, the harmonics decay slowly with increasing frequency.

The second row shows a different parameter set for  $m(x)$ ,

$$k_p = 0.5, \quad k_n = 0.2, \quad G_p = 10 \text{ dB}, \quad G_n = 10 \text{ dB} \quad ,$$

which also has asymmetrical connection points but lower gain values. The produced harmonics decay quicker with rising frequency. Odd harmonics are produced as well as even harmonics but the even harmonics are not as pronounced as the odd ones.

The bottom row shows a parameter set with symmetric configuration and low gain values,

$$k_p = 0.1, \quad k_n = 0.1, \quad G_p = 6 \text{ dB}, \quad G_n = 6 \text{ dB} \quad .$$

The 20<sup>th</sup> harmonic at  $f = 2100 \text{ Hz}$  already has a very low magnitude of approximately  $-100 \text{ dB}$  because of the comparatively low gain values and due to the symmetric configuration only odd harmonics are produced by this parameter set.

### 3.2.2 Anti-Aliasing

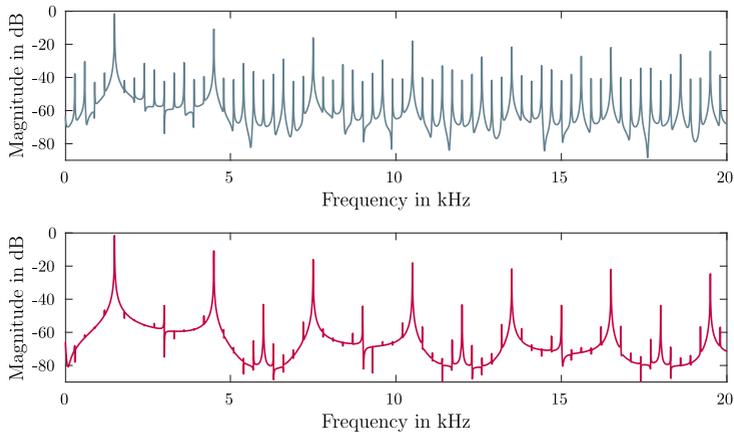
When digital signals are distorted, the harmonic components which go beyond the Nyquist frequency are mirrored back into the audible domain, creating unwanted artifacts. This effect can be observed in the upper plot of Fig. 3.5. The main peak at  $f_0 = 1500 \text{ Hz}$  and the harmonics at  $f = n \cdot f_0$  with  $n \in \mathbb{N}$  is visible but between those peaks, the aliasing frequencies can be seen as irregular peaks with a lower magnitude.

A usual countermeasure against this effect can be achieved by upsampling the digital signal with an oversampling factor of  $L = 8$  directly before the mapping function  $m(x)$  is applied and downsampling with the same factor directly after the mapping function. Different interpolation and decimation filters used for up- and downsampling are analyzed in [146]. The unwanted aliasing peaks are reduced as can be seen in the lower plot of Fig. 3.5.

Oversampling is a very effective tool against aliasing artifacts because it pushes the Nyquist limit towards a higher frequency. But it also introduces a higher computational load which is unwanted considering possible real-time applications of the digital model. Parker et al. introduced a method against aliasing which is perfectly suited for memoryless mapping functions such as the piecewise-defined function used in this work [147].

The idea behind the method is to apply the mapping function to a signal which is an approximation of a continuous-time signal

$$y(t) = m(\hat{x}(t)) \quad , \quad (3.13)$$



**Figure 3.5**

*Output  $m(x)$  for a 1500 Hz sine input signal (frequency-domain). Above: without oversampling. Below: oversampling by a factor of  $L = 8$ .*

with  $\tilde{x}(t)$  as the approximated continuous-time input signal and  $y(t)$  as the approximated continuous-time output. The approximation is created by linearly interpolating the sampled signal to an arbitrary sample point between two time steps. Because  $y(t)$  is only an approximated continuous-time signal it needs to be filtered to remove components above the Nyquist frequency before it is sampled with the original sampling frequency.

This can be done by applying continuous-time convolution with a rectangular low-pass kernel. Taking advantage of the assumption that  $\tilde{x}(t)$  is a piecewise linear signal, this convolution can be summarized to,

$$y(n) = \frac{M(x(n)) - M(x(n-1))}{x(n) - x(n-1)}, \quad (3.14)$$

with  $M(\cdot)$  as the antiderivative of  $m(\cdot)$  [147]. Due to the rectangular convolution kernel, some filtering of high frequencies is applied when using this method. For this reason, oversampling by factor two is used in combination with the antiderivative method. This method can also be extended to higher order antiderivatives as described by Bilbao et al. [148].

The frequency content produced by electric guitars rarely exceeds 2000 Hz. Therefore, the first antiderivative with two times oversampling is used in this work. Since the mapping function is piecewise-defined, the antiderivative of each section has to be calculated separately and can be joined at the same connection points as the mapping function itself.

The antiderivative of  $m(x)$  can be calculated to

$$M(x) = \begin{cases} b_p \cdot x + \frac{a_p}{g_p} \cdot \ln(\cosh(g_p(x - k_p))) + c_p & \text{if } x > k_p \\ \ln(\cosh(x)) & \text{if } -k_n \leq x \leq k_p, \\ b_n \cdot x + \frac{a_n}{g_n} \cdot \ln(\cosh(g_n(x + k_n))) + c_n & \text{if } x < -k_n \end{cases}, \quad (3.15)$$

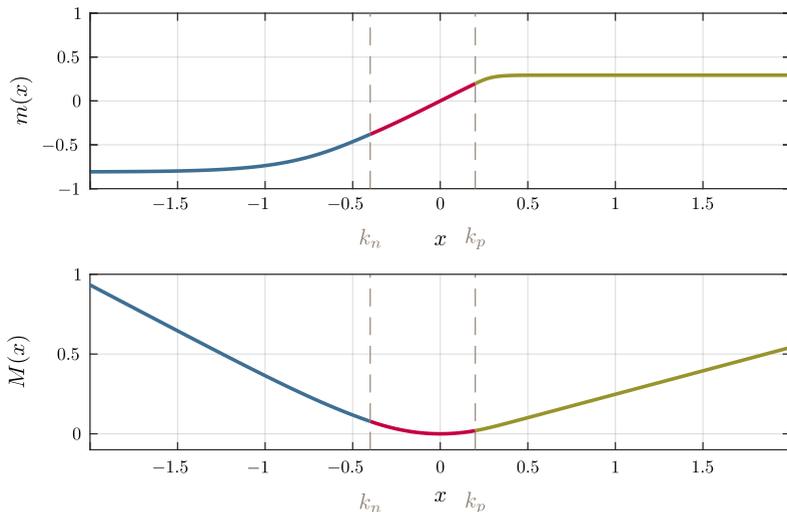
with

$$c_p = \ln(\cosh(k_p)) - k_p \cdot \tanh(k_p), \quad (3.16)$$

$$c_n = \ln(\cosh(k_n)) - k_n \cdot \tanh(k_n), \quad (3.17)$$

as the constants of integration.

The mapping function  $m(x)$  and its corresponding antiderivative  $M(x)$  are shown in Fig. 3.6.



**Figure 3.6**

Antiderivative  $M(x)$  (bottom) to the mapping function  $m(x)$  (top). Shown here for a parameter set of  $k_p = 0.2$ ,  $k_n = 0.4$ ,  $G_p = 20$  dB and  $G_n = 6$  dB. Different colors indicate the piecewise defined functions according to Eq. (3.7) and Eq. (3.15).

Two major numerical issues have to be addressed during calculation of  $y(n)$  with this method. The first is the ill-conditioning of Eq. (3.14) when

$x(n) \approx x(n-1)$ . According to [147] the equation can be substituted by

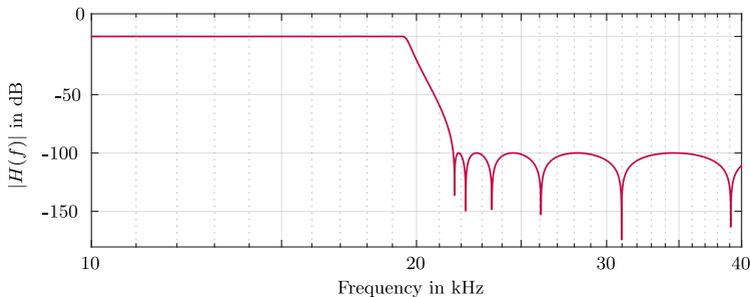
$$\frac{M(x(n)) - M(x(n-1))}{x(n) - x(n-1)} = m \left( \frac{x(n) + x(n-1)}{2} \right) + O((x(n) - x(n-1))^2) , \quad (3.18)$$

with  $O(\dots)$  as the order of the error, when  $x(n) - x(n-1)$  becomes very small.

The second numerical issue that arises is the exponential growth of the  $\cosh(x)$  - term. If the absolute value of the argument  $x$  rises, the hyperbolic cosine results in very large values which can cause numerical instability. This has been circumvented by approximating the antiderivative with a linear function for large values of the argument of the hyperbolic cosine, since  $\ln(\cosh(x)) \approx |x| - \ln(2)$  for  $x > 10$ .

Upsampling by factor  $L$  is usually performed by inserting  $L-1$  zeros between consecutive samples and then filtering the signal with a linear phase FIR filter to suppress image spectra. Downsampling is performed by first filtering the signal with a linear phase FIR filter and afterwards keeping only every  $L^{\text{th}}$  sample from the filtered signal. In a real-time application however, linear phase FIR filters introduce a considerable delay due to their symmetric impulse response, which is undesirable because too much latency creates a disturbing unresponsive feeling for the guitarist.

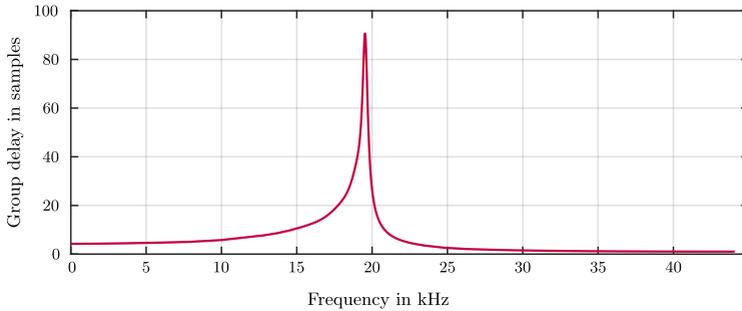
Therefore an elliptic low-pass filter is designed which can be efficiently implemented using six second-order IIR filters for up- and downsampling by factor  $L = 2$ . It operates at  $f_{s,up} = 2 \cdot f_s$  and was designed with a maximum passband ripple of  $\delta = 0.5$  dB, stopband attenuation of  $A_s = 100$  dB, a passband edge frequency of  $f_{\text{pass}} = 19$  kHz and a stopband edge frequency of  $f_{\text{stop}} = 21$  kHz as shown by Fig. 3.7.



**Figure 3.7**

*Elliptic low-pass filter for up- and downsampling.*

The group delay of the filter can be seen in Fig. 3.8. Below  $f = 15$  kHz the group delay is lower than 10 samples which corresponds to an overall

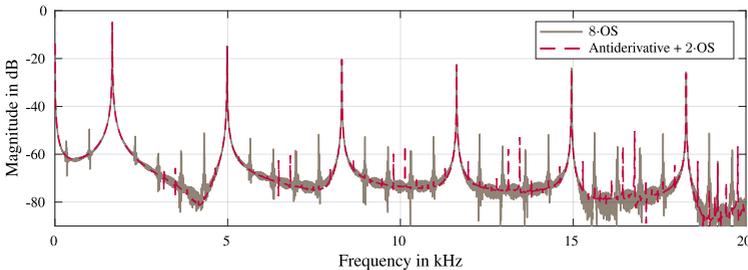


**Figure 3.8**

*Group delay of the elliptic low-pass filter for up- and downsampling.*

latency below 1 ms (since the filter is applied twice). Only for frequencies above 16 kHz the group delay becomes higher which is not relevant for real-time applications when the output of the digital guitar amplifier is filtered with the impulse response of a speaker cabinet which exhibits strong attenuation above 10 to 12 kHz. Furthermore [146] suggests that the upsampling filter can be replaced by linear interpolation equalized with a high frequency shelving filter.

A comparison between the antiderivative method with two times oversampling and 8 times oversampling using the elliptic low-pass filter can be seen in Fig. 3.9. The parameters for  $m(x)$  were set to  $k_p = 0.7$ ,  $k_n = 0.4$ ,



**Figure 3.9**

*Aliasing comparison for a  $f_0 = 1661$  Hz sine wave distorted with  $m(x)$ . Antiderivative method with two times oversampling compared to 8 times oversampling.*

$G_p = 20$  dB,  $G_n = 10$  dB and the input signal was amplified with a pre-gain of  $G_{pre} = 30$  dB. Some aliasing components from the antiderivative method are higher than using 8 times oversampling but the overall energy

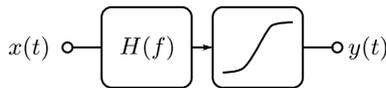
of all aliasing components is shifted to higher frequencies which creates a more pleasant sound. A proper analysis of calculation time was not performed in this work but a comparison for a simple hyperbolic tangent can be found in [148]. The time of computation for a hyperbolic tangent with six times oversampling is compared to the antiderivative method with two times oversampling. The first order antiderivative method is more than twice as fast as six times oversampling.

### 3.3 Simple Block-Oriented Models

Block-oriented models are arbitrarily connected filters and nonlinear blocks. The most simple block-oriented models are called Hammerstein and Wiener models [94]. They are mostly used for system identification in the field of control-theory [149,150]. An overview of block-oriented models is provided in [125] but the applicability of these models to audio signal processing is not mentioned. The use of block-oriented models for audio signal processing is mentioned in [18–20, 35–37, 56, 64, 66, 70, 79, 119, 122–124, 126–128, 151]. In [152] the use of an adaptive block-based model using a feedback structure for speech-coding is described.

Therefore, the most basic block-oriented models and their impact on audio signals shall be mentioned here.

A Hammerstein model consists of a nonlinear block followed by a filter and a Wiener model is a filter in series with a nonlinear block. These models are shown in Fig. 3.10. In the case of a Wiener model, a filter



**(a)** Wiener model: filter in series with nonlinear block



**(b)** Hammerstein model: nonlinear block in series with filter

**Figure 3.10**

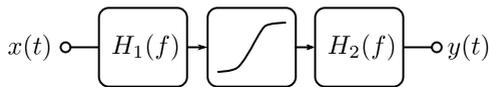
*Simple block-oriented models (a) Wiener model (b) Hammerstein model*

in front of a nonlinear block controls which frequency range of the input signal is distorted. If the filter attenuates a frequency range, the amplitude is reduced and can pass the linear region of the nonlinear mapping

curve [127]. In the opposite case more harmonics are added because the mapping function rather operates in the saturated region. This controls the harmonic components and the intermodulation distortion which are created by the nonlinear block [124].

If a Hammerstein model is considered with a filter following the nonlinear block, all frequencies are distorted equally, but the harmonic overtones are shaped by the output filter. This gives the distortion a specific sound. In [123] a Hammerstein model is used for black-box modeling of distortion systems.

Combining both models to a Wiener-Hammerstein model, which is a nonlinear block sandwiched between two filters, allows flexible reproduction of many analog distortion circuits. The model is able to shape the



**Figure 3.11**

*Wiener-Hammerstein model, consisting of a memoryless nonlinear mapping function sandwiched between two filters (LTI).*

input signal with the input filter controlling the amount of distortion per frequency as well as intermodulation distortion and shape the distorted signal with the output filter. The shape of the nonlinear function also influences the amount and shape of the produced harmonics. In [79] a parametric function is adapted to several nonlinear circuits which are common in distortion effect pedals.

---

## System Identification

---

In this chapter the necessary steps to adjust a digital model according to measured signals are detailed. Different parameter optimization algorithms are explained in this section: the Gauss-Newton method, the gradient-descent method and the Levenberg-Marquardt method. The calculation of gradients with the finite differences scheme and the importance of a defined parameter space as well as the initial parameter vector are discussed. Selected cost-functions are presented and the identification of linear subsystems in a nonlinear system is explained as well. Finally the used method of data acquisition by measuring analog systems is presented.

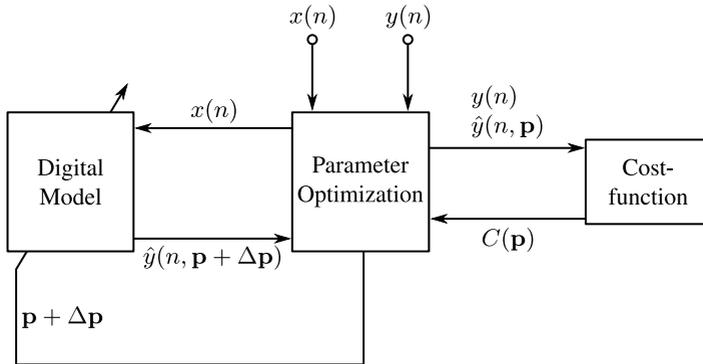
### 4.1 Parameter Optimization: Overview

A typical parameter optimization, independently of a specific optimization method, is shown in Fig. 4.1. The digitized input  $x(n)$  and reference signal  $y(n)$  are the input variables of the parameter optimization algorithm. The algorithm can evaluate the digital model for the input signal  $x(n)$  with a specific parameter set

$$\mathbf{p} = (p_0 \ p_1 \ \dots \ p_M)^T \quad (4.1)$$

and produce the model output  $\hat{y}(n, \mathbf{p})$ . The length of the parameter vector  $M$  is the dimensionality of the minimization problem.

Both output signals from reference device and digital model are given to a cost-function, which calculates the error between these two signals  $C(\mathbf{p})$ .



**Figure 4.1**

*Block diagram of the iterative optimization routine.*

The optimization algorithm calculates the parameter update  $\Delta \mathbf{p}$  by differentiating the cost-function with respect to its parameters. The parameter update equation is unique for each method and will be detailed in the following. If the value of the cost-function for the new parameter set is smaller than for the old parameter set  $C(\mathbf{p} + \Delta \mathbf{p}) < C(\mathbf{p})$  the parameters are updated and the process is repeated until one of several termination criteria is fulfilled.

The termination criteria are listed below:

- Change in parameters smaller than specified tolerance
- Change in cost-function smaller than specified tolerance
- Maximum number of iterations exceeded
- Maximum number of unsuccessful steps exceeded

Please note that the error does not necessarily have to be the time-domain error. The output signals can be processed so that the error between them fits the needs of the optimization. For example, the error between two signal envelopes can be calculated, or the time-frequency error between two short-time Fourier transform (STFT) signals. Different cost-functions will be presented in Sec. 4.3.

## 4.2 Parameter Optimization Methods

Iterative methods which are able to minimize a cost-function will be presented in the following. The aim is to find the minimum of  $C(\mathbf{p})$ , describing the error between the output of a reference system  $y(n)$  and a digital model  $\hat{y}(n, \mathbf{p})$  for a specific input signal  $x(n)$  of length  $N$ .

### 4.2.1 Gauss-Newton Method

For the Gauss-Newton method the minimization is done by approximating  $C(\mathbf{p})$  with the second-order Taylor series, which leads to the update equation

$$C(\mathbf{p} + \Delta\mathbf{p}) = C(\mathbf{p}) + (\Delta\mathbf{p})^T \cdot \mathbf{g} + \frac{1}{2} (\Delta\mathbf{p})^T \cdot \mathbf{H} \cdot \Delta\mathbf{p} \quad , \quad (4.2)$$

with the gradient

$$\mathbf{g} = \left( \frac{\partial C(\mathbf{p})}{\partial p_0} \quad \frac{\partial C(\mathbf{p})}{\partial p_1} \quad \dots \quad \frac{\partial C(\mathbf{p})}{\partial p_M} \right)^T \quad , \quad (4.3)$$

and  $\mathbf{H}$  as the Hessian matrix, containing the second-order derivatives of the cost-function with respect to its parameters. The change of the parameter vector per iteration is expressed as  $\Delta\mathbf{p}$ . In order to minimize the cost-function its first derivative  $C'(\mathbf{p} + \Delta\mathbf{p})$  needs to be zero

$$C'(\mathbf{p} + \Delta\mathbf{p}) = \mathbf{g} + \mathbf{H} \cdot \Delta\mathbf{p} = 0 \quad , \quad (4.4)$$

which leads to the parameter update equation

$$\Delta\mathbf{p} = -\mathbf{H}^{-1} \cdot \mathbf{g} \quad . \quad (4.5)$$

The gradient  $\mathbf{g}$  and the Hessian matrix  $\mathbf{H}$  can also be expressed using the Jacobian matrix  $\mathbf{J}$  which is of size  $N$  by  $M$  and contains the partial derivatives of the model outputs with respect to every single parameter in  $\mathbf{p}$ ,

$$\mathbf{J} = \begin{pmatrix} \frac{\partial \hat{y}(n_0, \mathbf{p})}{\partial p_0} & \frac{\partial \hat{y}(n_0, \mathbf{p})}{\partial p_1} & \dots & \frac{\partial \hat{y}(n_0, \mathbf{p})}{\partial p_M} \\ \frac{\partial \hat{y}(n_1, \mathbf{p})}{\partial p_0} & \frac{\partial \hat{y}(n_1, \mathbf{p})}{\partial p_1} & \dots & \frac{\partial \hat{y}(n_1, \mathbf{p})}{\partial p_M} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \hat{y}(n_N, \mathbf{p})}{\partial p_0} & \frac{\partial \hat{y}(n_N, \mathbf{p})}{\partial p_1} & \dots & \frac{\partial \hat{y}(n_N, \mathbf{p})}{\partial p_M} \end{pmatrix} \quad , \quad (4.6)$$

and the residual

$$\mathbf{r} = \begin{pmatrix} y(n_0) - \hat{y}(n_0, \mathbf{p}) \\ y(n_1) - \hat{y}(n_1, \mathbf{p}) \\ y(n_2) - \hat{y}(n_2, \mathbf{p}) \\ \vdots \\ y(n_N) - \hat{y}(n_N, \mathbf{p}) \end{pmatrix} \quad , \quad (4.7)$$

which leads to new expressions for  $\mathbf{g}$  and  $\mathbf{H}$

$$\mathbf{g} = -\mathbf{J}^T \cdot \mathbf{r} , \quad (4.8)$$

$$\mathbf{H} = \mathbf{Q} + \mathbf{J}^T \cdot \mathbf{J} . \quad (4.9)$$

The matrix  $\mathbf{Q}$  contains the second-order derivatives of Eq. (4.5), it is zero for linear minimization problems and is often ignored. The matrix  $\mathbf{Q}$  is multiplied with the residual  $\mathbf{r}$  which contains random measurement errors, which can be either positive or negative. As soon as  $\mathbf{p}$  is close to its optimum these terms tend to cancel each other out, when summing over  $n$  (assuming that  $\mathbf{r}$  is uncorrelated with the second derivative of  $\hat{y}(n, \mathbf{p})$ ) [153]. Considering this, Eq. (4.5) can be rewritten as

$$\Delta \mathbf{p} = (\mathbf{J}^T \cdot \mathbf{J})^{-1} \cdot \mathbf{J}^T \cdot \mathbf{r} . \quad (4.10)$$

This method has advantageous properties like fast convergence to the nearest optimum, but if the cost-function is not well designed, the Gauss-Newton method can get stuck in saddle-points or even diverge.

## 4.2.2 Gradient-Descent Method

Another method to find the minimum of a cost-function is the gradient-descent method. The Gauss-Newton method has some disadvantageous properties which were discussed in the previous section. In these cases it is desirable to minimize the cost-function by going rigidly downhill, depending on the gradient of the cost-function with respect to the models parameters.

To achieve this, the first-order Taylor expansion of the cost-function is utilized:

$$C(\mathbf{p} + \Delta \mathbf{p}) \simeq C(\mathbf{p}) + \mathbf{g}^T \cdot \Delta \mathbf{p} , \quad (4.11)$$

with  $\mathbf{g}$  as defined by Eq. (4.3). The term  $\mathbf{g}^T \cdot \Delta \mathbf{p}$  has to be smaller than zero to reduce the value of the cost-function

$$\mathbf{g}^T \cdot \Delta \mathbf{p} = -\alpha, \quad \alpha > 0 .$$

From these constraints the adjustment to the parameter vector ( $\Delta \mathbf{p}$ ) can be calculated by multiplying with the gradient vector  $\mathbf{g}$  and rearranging the equation

$$\mathbf{g}^T \cdot \Delta \mathbf{p} = -\alpha , \quad (4.12)$$

$$\mathbf{g} \cdot \mathbf{g}^T \cdot \Delta \mathbf{p} = -\alpha \cdot \mathbf{g} ,$$

$$\|\mathbf{g}\|^2 \cdot \Delta \mathbf{p} = -\alpha \cdot \mathbf{g} ,$$

$$\Delta \mathbf{p} = -\frac{\alpha}{\|\mathbf{g}\|^2} \cdot \mathbf{g} , \quad (4.13)$$

where the operator  $\|\cdot\|$  denotes the Euclidean norm of a vector.

Without the normalization in Eq. (4.13), the optimization would not be able to overcome saddle-points and would be very slow if the gradient becomes small which is the case if  $\mathbf{p}$  is already close to the solution [154].

### 4.2.3 Levenberg-Marquardt Method

To summarize the last two sections, the Gauss-Newton method performs well when the parameter vector  $\mathbf{p}$  is already in the vicinity of the optimum and the gradient-descent method performs robustly even if the current parameter vector is further away from the optimal solution. Levenberg and Marquardt introduced a damping factor  $\mu$  to combine the advantages of the Gauss-Newton and gradient-descent methods [155, 156].

The damping factor is incorporated in Eq. (4.10) and yields

$$\Delta\mathbf{p} = (\mathbf{J}^T \cdot \mathbf{J} + \mu \cdot \mathbf{I})^{-1} \cdot \mathbf{J}^T \cdot \mathbf{r} \quad , \quad (4.14)$$

with  $\mathbf{J}$  and  $\mathbf{r}$  defined in the same way as in Eq. (4.6) and Eq. (4.7). If the damping factor  $\mu$  is larger than zero  $\Delta\mathbf{p}$  points downhill [153]. For large values of  $\mu$ , the new term  $\mu \cdot \mathbf{I}$  becomes dominant and Eq. (4.14) simplifies to

$$\Delta\mathbf{p} \approx \frac{1}{\mu} \cdot \mathbf{J}^T \cdot \mathbf{r} = -\frac{1}{\mu} \cdot \mathbf{g} \quad , \quad (4.15)$$

which coincides with the parameter update of the gradient descent method Eq. (4.12).

If  $\mu$  has a small value, the term  $\mu \cdot \mathbf{I}$  becomes irrelevant and the equation resembles the Gauss-Newton parameter update Eq. (4.10).

The damping factor  $\mu$  is initialized with a small value and is updated after every iteration. If  $C(\mathbf{p} + \Delta\mathbf{p}) < C(\mathbf{p})$  is true, the parameter vector is updated and  $\mu$  is decreased. In the opposite case, the parameter vector is not updated and the value of  $\mu$  is increased [155].

Additionally, Marquardt proposed to replace the identity matrix  $\mathbf{I}$  from Eq. (4.14) with the diagonal elements of  $\mathbf{J}^T \mathbf{J}$  to scale the gradient and create a larger change for parameters with a small gradient, resulting in the Levenberg-Marquardt parameter update equation [156],

$$\Delta\mathbf{p} = (\mathbf{J}^T \cdot \mathbf{J} + \mu \cdot \text{diag}(\mathbf{J}^T \mathbf{J}))^{-1} \cdot \mathbf{J}^T \cdot \mathbf{r} \quad . \quad (4.16)$$

This incorporates the normalization of the gradient scaling from the gradient-descent method (Eq. (4.13)).

### 4.2.4 Finite Differences

Calculating the Jacobian matrix  $\mathbf{J}$  is not a trivial task. An analytical solution would be desirable but is often impractical because it has to be

recomputed for every minor change in the model. For this reason the finite-differences scheme is used for approximating the Jacobian matrix. Each parameter is changed by a small value  $\Delta p_i$  and the influence on the model output is observed. The approximation of the Jacobian can be expressed as

$$\mathbf{J} = \frac{\partial \hat{y}(n, \mathbf{p})}{\partial p_i} \approx \frac{1}{\Delta p_i} (\hat{y}(n, \mathbf{p} + \Delta \mathbf{p}_i) - \hat{y}(n, \mathbf{p})) \quad , \quad (4.17)$$

with the vector  $\Delta \mathbf{p}_i$ , which is all-zero, except for the  $i$  th entry, which is equal to  $\Delta p_i$  and  $i$  is iterated over all parameters  $i = 1, \dots, M$ . The value of  $\Delta p_i$  should be considered carefully. If it is chosen too small, numerical instability can occur. But if it is too big, the result will not approximate the derivative. In this work the parameter adjustment was chosen to be,

$$\Delta p_i = \sqrt{\epsilon} \cdot \|\mathbf{p}\|_2 \quad ,$$

with  $\epsilon$  as the smallest number in floating point arithmetic which produces a result different from 1.0 when added to it.  $\epsilon$  is called machine accuracy and for 64 bit double precision, the value is  $\epsilon = 2^{-52} \approx 2.22 \cdot 10^{-16}$  [157, pp. 28–31].

A more accurate estimate of the Jacobian could be achieved by using the central finite difference scheme

$$\mathbf{J} \approx \frac{1}{2\Delta p_i} (\hat{y}(n, \mathbf{p} + \Delta \mathbf{p}_i) - \hat{y}(n, \mathbf{p} - \Delta \mathbf{p}_i)) \quad , \quad (4.18)$$

instead of the forward finite difference method (Eq. (4.17)), but there will always be the trade-off between computational effort and accuracy. For the central finite difference scheme, the model has to be evaluated twice as often as for forward central differences. If the model has a lot of parameters, approximating the Jacobian will become the most time-consuming part of the optimization procedure.

### 4.2.5 Defined Parameter Space

During optimization it is important to operate in a defined parameter space where every possible combination of the parameters will give a valid model output. Some parameters might cause the model to become unstable if the wrong values are used. As an example, consider optimizing the poles of an IIR filter in the  $z$ -domain. If the pole radius becomes larger than one, the filter becomes unstable and the digital model can not be evaluated any more. The optimization would break off. If the trust-region for the parameters is chosen appropriately this scenario could be circumvented. For the previous example, the pole radius should be clipped to a value of one and the resulting system could not become unstable. The Levenberg–Marquardt algorithm, which is used in this work, is adapted to feature boundaries for the parameters. If a parameter is set to a value outside

the parameter space during optimization, it is clipped to its maximum or minimum value. The lower and upper bounds can be chosen for each parameter.

### 4.2.6 Initial Parameter Vector

All discussed optimization methods are gradient-based, which means that the derivative of the cost-function has to be calculated or approximated. After each iteration the algorithm continues in the direction of decreasing gradient resulting in an inability to escape a local minimum [91].

Therefore, the initial parameter vector plays a very important role during the optimization process. If the initial set is too far from an optimal solution, the algorithm will most likely get stuck in a local minimum, meaning that the model is not able to recreate the reference system to a satisfactory degree.

A typical step to circumvent this problem is the performance of a grid search, where the result of the cost-function for each possible parameter combination is calculated on a coarse grid with size  $g_s$  over the parameter space and the combination with the lowest cost-function value is selected as initial parameter set. But this has to be treated carefully since the computational effort of checking every parameter combination grows exponentially with the number of parameters.

## 4.3 Cost-Functions

The choice of the cost-function is one of the determining factors to ensure a successful optimization. Usually a digital model has a lot of parameters which alter the models output and therefore the result of the cost-function. In most cases it is unlikely to find a cost-function with which all parameters can be optimized at once and still reach its global minimum.

For all nonlinear reference devices, which are modeled in this work, a multi-step optimization with different cost-functions is used. First, the most-important parameters of the model have to be optimized with a corresponding cost-function which depends strongly on the used digital model and the type of reference system. In the following an overview of all used functions is given.

For guitar amplifiers and distortion effects the perceived sound is of importance. Contrary to high-fidelity audio applications it is not important to have a linear and balanced sound. Instead, strong coloration and harmonic distortion are wanted to create the characteristic sound of an analog reference device [158, p. 10-1]. A psycho-acoustically motivated cost-function based on the short-time Fourier transform is proposed as well.

### 4.3.1 Time-domain Cost-Function

The simplest and most straight-forward cost-function is the time-domain error

$$C(\mathbf{p}) = \sum_{n=1}^N (y(n) - \hat{y}(n, \mathbf{p}))^2, \quad (4.19)$$

which is also called sum of squares. The digitized output signals of the reference system and the output of the digital model are subtracted from each other and each sample of the residual is squared and summed up.  $N$  denotes the overall length of the signals.

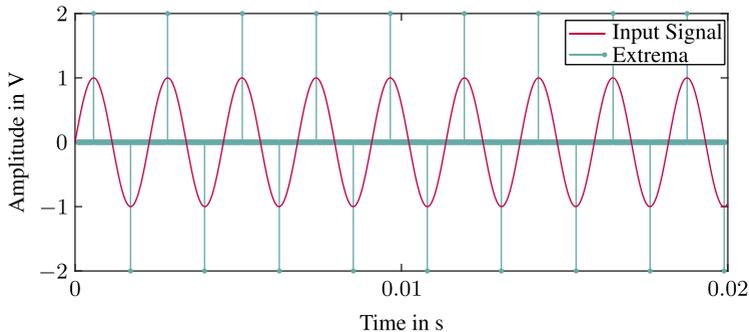
If the (time-domain) sum of squares is low the model will perform very well but only a small misalignment in time will result in relatively large error values. Also small discrepancies in the phase-response of the digital model will increase the error but a human listener would not be able to hear differences between the signals.

Experience has shown that the sum of squares is not suited for optimization of highly nonlinear systems like distortion systems. All digital models which were optimized using the sum of squares cost-function were perceptually well distinguishable from the reference device.

### 4.3.2 Envelope Cost-Function

Some parameters of a digital model have a strong influence on the envelope of the signal. To be able to adjust these parameters in a meaningful way, an envelope-based cost-function can be used. There exist several ways to calculate the envelope of a signal. A straightforward approach would be low-pass filtering of the absolute value of the input signal. The cut-off frequency of the low-pass filter should be set to a low value of  $f_c \leq 10$  Hz. But this method is inaccurate if the signal has transient components. For this reason a second envelope detector was used, which is based on a peak-picking technique and linear interpolation between each detected peak. First, the positions of all extrema are calculated by differentiating the sign of the first derivative of the signal. When the first derivative of a signal changes its sign from positive to negative the original waveform exhibits a maximum. When the sign changes from negative to positive a minimum occurs in the original waveform respectively. When the sign of the first derivative is differentiated again and the result is multiplied with  $-1$ , the extrema are marked with a value of 2 or  $-2$  respectively.

The extrema signal and the incoming waveform are shown in Fig. 4.2. With this technique the envelope of highly transient signals can be detected accurately. The value of each positive extremum is extracted and linear interpolation is used to create the envelope signal. The same can be done for all negative extrema. This is especially useful when some parameters of the digital model only influence positive or negative signal amplitudes.



**Figure 4.2**  
*Extrema of a signal.*

### 4.3.3 Time-Frequency Cost-Function

To be able to make psycho-acoustic considerations it is important to consider not only the time-domain or the frequency-domain separately. Therefore, the short-time Fourier transform (STFT) is used [98, pp. 30–31].

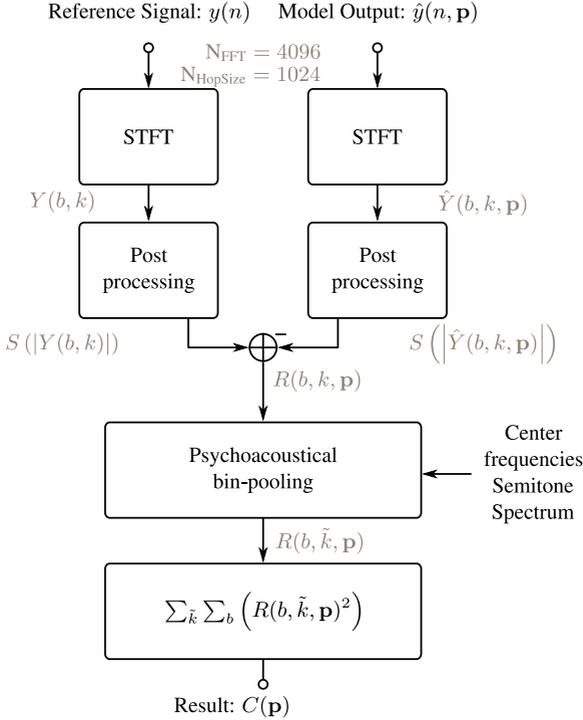
The calculation of the cost-function is illustrated in Fig. 4.3. The reference signal  $y(n)$  and the output of the digital model  $\hat{y}(n, \mathbf{p})$  are transformed into the time-frequency domain

$$Y(b, k) = \text{STFT}(y(n)) \quad , \quad (4.20)$$

$$\hat{Y}(b, k, \mathbf{p}) = \text{STFT}(\hat{y}(n, \mathbf{p})) \quad , \quad (4.21)$$

using a Hann window before applying the Fourier transform with a block size of  $N_{\text{FFT}} = 4096$ . A hop-size of  $N_{\text{HopSize}} = 1024$  samples is used and  $b$  and  $k$  denote the block and frequency indices. The block index  $b$  ranges from  $b_0 = 0$  to  $B = \lfloor \frac{N - N_{\text{FFT}}}{N_{\text{HopSize}}} \rfloor$  with  $N$  as the length of the time-domain signal and  $\lfloor \cdot \rfloor$  denoting rounding to the next smallest integer. The frequency index  $k$  ranges from  $k_0 = 0$  to  $K = \frac{N_{\text{FFT}}}{2} - 1$  resulting in a one-sided spectrum.

After the STFT has been computed the magnitude spectrogram of  $|Y(b, k)|$  and  $|\hat{Y}(b, k, \mathbf{p})|$  is calculated in the post-processing stage. This means the phase of the signal is ignored. Phase differences of tonal signals become audible, when at least three distinct tones are present in the signal and the phase difference between them is at least  $10^\circ$  in an anechoic chamber for a combination of sine waves around 1 kHz with a sound pressure level of 70 dB. In a normal room the value becomes three times as big [159, pp. 187–189]. Thus the phase is not as important as the magnitude when comparing the similarity of two signals. Especially since signals from an electric guitar are monaural and no stereo effects were modeled



**Figure 4.3**

*Spectrogram based cost-function.*

in this work where phase differences between left and right channel are non-negligible because they can create audible artifacts.

In the post-processing step the magnitude of the spectrogram is scaled logarithmically, according to

$$S(|Y(b, k)|) = \max\left(0, \frac{20 \cdot \log_{10}(|Y(b, k)|) + S_{\text{offset}}}{S_{\text{offset}}}\right), \quad (4.22)$$

with  $S_{\text{offset}} = 100$  dB. By using this logarithmic scaling, magnitude differences for low amplitudes have the same effect on the overall error value as magnitude differences for high amplitudes but everything lower than  $-100$  dB is below the noise floor and can be ignored during optimization. In this work the offset was chosen to be  $S_{\text{offset}} = 100$  dB for a 0 dB reference sine wave corresponding to  $\pm 1$  V Amplitude. The time-domain reference

signals were normalized to 0 dB before the optimization was started to ensure that the offset is low enough.

The residual spectrogram

$$R(b, k, \mathbf{p}) = |Y(b, k)| - |\hat{Y}(b, k, \mathbf{p})| \quad (4.23)$$

is calculated by subtracting the model spectrogram from the reference spectrogram.

Before calculating the final score, the frequency bins of the Fourier transform are pooled by calculating the mean value of the bins for a certain frequency region, specified by the center frequencies of a semitone spectrum as suggested by

$$R(b, \tilde{k}, \mathbf{p}) = \text{pool}(R(b, k, \mathbf{p})) \quad (4.24)$$

A semitone spectrum is used throughout this work, starting with the lowest frequency of  $f_0 = 27.5$  Hz, which is A0: the lowest note possible on a standard tuning 5-string bass guitar. The next semitone frequency is calculated by

$$f_{\tilde{k}} = f_0 \cdot 2^{\frac{\tilde{k}}{12}}, \quad (4.25)$$

up to  $f_{\tilde{K}} = 19912.13$  Hz. Frequencies from dc to  $f_0$  and frequencies from  $f_{\tilde{K}}$  to  $f_s/2$  are averaged to one value each.

The final score is calculated by squaring each entry of  $R(b, \tilde{k}, \mathbf{p})$  and summing up all elements in time- and frequency-direction

$$C(\mathbf{p}) = \sum_{\tilde{k}=0}^{\tilde{K}} \sum_{b=0}^B R(b, \tilde{k}, \mathbf{p})^2 \quad (4.26)$$

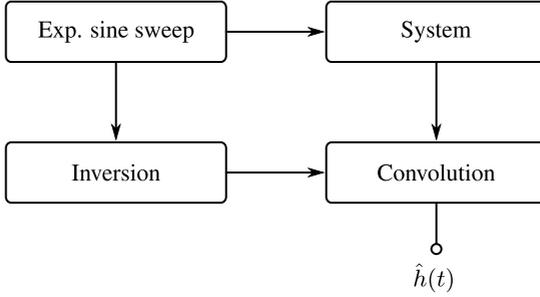
with  $B$  as the number of time-domain blocks and  $\tilde{K}$  as the number of semitone bins.

## 4.4 Identification of Linear Subsystems

The result of the identification of a linear subsystem is an impulse response (IR) which describes the linear system. The most commonly used signals to identify a linear system are pulses, pseudo-random sequences and sine sweeps. The most suitable signal for measuring linear subsystems of non-linear systems is a sine sweep as Holters et al. showed. Pseudo-random sequences produce unreliable results when clipping is involved [160].

In this work, the impulse responses were measured with the method described in [120]. The block-diagram of the measurement procedure is shown in Fig. 4.4. A (nonlinear) system is excited with an exponentially swept sine wave, which can be generated digitally by

$$x_{\text{sweep}}(n) = A \cdot \sin \left( \frac{\omega_1 \cdot (L_s - 1)}{\log(\omega_2/\omega_1)} \cdot \left( e^{\left( \frac{n}{L_s - 1} \log(\omega_2/\omega_1) \right)} - 1 \right) \right) \quad (4.27)$$

**Figure 4.4**

Block-diagram of an impulse response measurement using an exponentially swept sine wave. The sweep is inverted and convolved with the system output to yield the impulse response  $\hat{h}(t)$ .

and the measured output of the system is convolved with an inverted input sweep

$$x_{\text{inv}}(n) = x_{\text{sweep}}(L_s - 1 - n) \cdot (\omega_2/\omega_1)^{\frac{-n}{L_s-1}}, \quad (4.28)$$

where  $L_s$  is the length of the sweep in samples,  $\omega_1 = 2\pi f_{\text{start}}/f_s$  is the start frequency of the sweep,  $\omega_2 = 2\pi f_{\text{stop}}/f_s$  is the highest frequency contained in the sweep,  $A$  is the sweeps amplitude and  $f_s$  is the sampling frequency. If  $x_{\text{sweep}}(n)$  is convolved with the inverse sweep  $x_{\text{inv}}(n)$ ,

$$x_{\text{sweep}}(n) * x_{\text{inv}}(n) \approx c \cdot \delta(n - \Delta n_0). \quad (4.29)$$

the result is a dirac delta function which is shifted in time by the initial shift  $\Delta n_0 = L_s - 1$  and scaled by a scalar factor  $c$ . According to Holters et al. [160] the scaling factor can be found to be

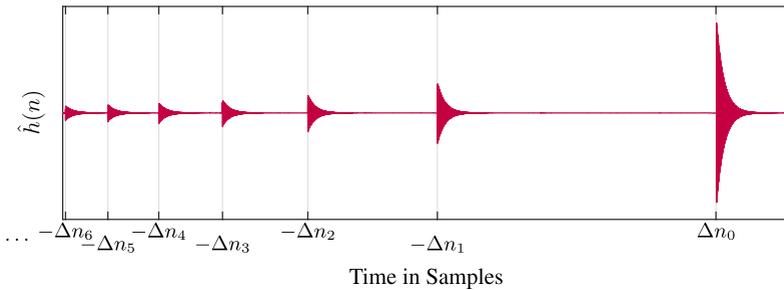
$$c = \frac{\pi L_s \cdot \left(\frac{\omega_1}{\omega_2} - 1\right)}{2(\omega_2 - \omega_1) \log\left(\frac{\omega_1}{\omega_2}\right)}. \quad (4.30)$$

To calculate the delayed impulse response of the system,

$$\hat{h}(n) = \frac{1}{c} \cdot x_{\text{inv}}(n) * y(n). \quad (4.31)$$

the inverted signal has to be convolved with the measured output of the system.

The impulse response  $\hat{h}(n)$  contains the linear response as well as the responses for higher order harmonics. If the measured system would be a linear one, the resulting impulse response would only contain the linear response.



**Figure 4.5**

Resulting impulse response  $\hat{h}(n)$  measured with an exponentially swept sine. The linear impulse response is at  $\Delta n_0$  and the harmonic responses at  $-\Delta n_{1,2,3,\dots}$ .

An exemplary impulse response  $\hat{h}(n)$  is shown in Fig. 4.5. The linear response of the system is placed at the time index  $\Delta n_0 = L_s - 1$ , while the harmonic responses are shifted in time. The first harmonic response is placed at  $-\Delta n_1$ , the second at  $-\Delta n_2$  and so on. The harmonic responses can be used in the digital model to filter the produced harmonics, as described in [119]. In this work only the linear response is used when measuring nonlinear systems because the shape of the generated harmonics can be defined by the shape of the nonlinear mapping curve (see Chap. 3).

#### 4.4.1 Equalization of Measurement Equipment

Kirkeby presented a method to cancel cross-talk between two loudspeakers [161] and Farina has adapted this method in order to negate the influence of the audio interface on the measurements. A ‘loop-back’ connection between an output of the interface and an input has to be established and all measurement signals are recorded with it (see Sec. 4.5, Fig. 4.6). Now the impulse response of the audio interface itself can be measured and its influence can be computationally eliminated [120].

The filter

$$H_c(f) = \frac{H_{\text{lb}}^*(f)}{H_{\text{lb}}^*(f) \cdot H_{\text{lb}}(f) + \gamma(f)} \quad (4.32)$$

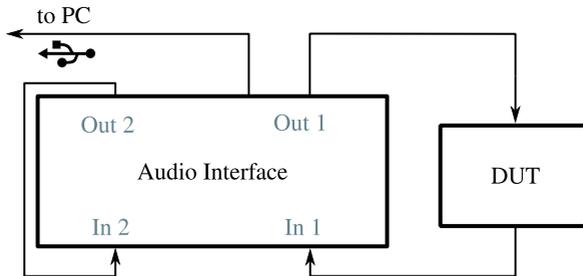
is calculated in frequency domain, where  $H_{\text{lb}}(f)$  is the frequency response of the loop-back measurement and  $H_{\text{lb}}^*(f)$  as its complex conjugate.  $\gamma(f)$  is a regularization parameter, which is usually designed in such a way, that it takes a very small value in the frequency region of the measurement signal and a much larger value outside this frequency range. To reduce the effects of the measurement equipment,  $H_c(f)$  can directly be multiplied with the measured frequency response.

## 4.5 Data Acquisition: Hardware Measurements

To perform any system identification the input and output data from the reference device need to be acquired. In this work, all measurements are taken with a USB audio interface connected to a personal computer. The interface converts the digital signals into analog ones and sends them through the analog reference device. The resulting output is measured and converted back into the digital domain. The digital to analog converter (DAC) and the analog to digital converter (ADC) of the audio interface introduce filtering. But modeling the filters of the interface would not be sensible, since they are not part of the reference device and have no impact on the audible frequency range [162], but they can influence the performance of the system identification.

### 4.5.1 Measurement Setup

All devices are measured with a USB audio interface. In this work the ‘RME – Fireface UC’ is used for all measurements [162]. One of the interfaces outputs is connected to the device under test (DUT) which is connected to the corresponding input of the audio interface. All signals which are sent through the DUT are simultaneously sent through a loop-back connection, which is a direct connection from another output of the audio interface to the corresponding input, as Fig. 4.6 depicts.



**Figure 4.6**

*Used measurement setup. One output is connected to the device under test (DUT) and another output is directly connected to the input of the audio interface.*

At first the interface is calibrated using a digital oscilloscope. A sine wave with a digital amplitude of  $\pm 1$  and a frequency of  $f_0 = 1$  kHz is played back and the outputs of the audio interface are directly connected to the oscilloscope. The gain of the interfaces outputs is then adjusted so that the digital amplitude of  $\pm 1$  corresponds to a voltage of  $V_{pp} = \pm 1$  V.

The loop-back connection has several benefits. All measurement signals are created with a personal computer (PC) and are stored as digital data. The recorded signals from the DUT are shifted in time with respect to the digital input signals because of the input- and output-buffers of the audio interface. If the recorded loop-back signals are considered as input signals no time shift between input and output occurs.

The loop-back connection is also used to equalize the measurement equipment when measuring impulse responses. With this method the influence of the audio interface itself on the measured frequency responses can be eliminated. The equalization is detailed in Sec. 4.4.

## Guitar Amplifiers

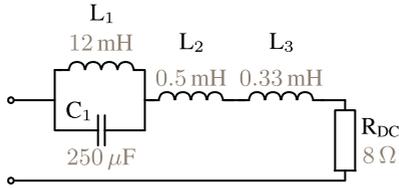
A guitar amplifier could be measured by connecting it to a speaker and recording the output with a microphone. But the influence of the speaker, the room and the microphone would be captured as well. This might be undesirable because each guitar player has different preferences about the speaker, microphone or even microphone position. For this reason the output of the amplifier should be measured directly. After the digital model is optimized to the desired reference device, its output can be filtered with the impulse response of a specific loudspeaker according to the preferences of the user.

When measuring the output of an amplifier some considerations have to be made due to the amplifiers high output power and extreme volume for overdrive of the power amplifier. Another consideration is the behavior of the speaker itself because it is not just an ohmic resistance. The complex impedance of the speaker is frequency dependent. An impedance equivalent filter, also called ‘dummy load’ or ‘power attenuator’, with circuit components dimensioned for the high output power is suited for measuring the guitar amplifier without high sound pressure levels. At the same time neither the influence of the oscillating membrane nor the room resonances influence the measurement and can be simulated after the digital model is optimized.

A ‘dummy load’ according to [163] is used for the measurements of the guitar amplifiers. The circuit is shown in Fig. 4.7. The impedance of the load is measured and compared to a measurement of one loudspeaker. The curves, shown in Fig. 4.8, do not match perfectly but neither would the measurements of two different speakers.

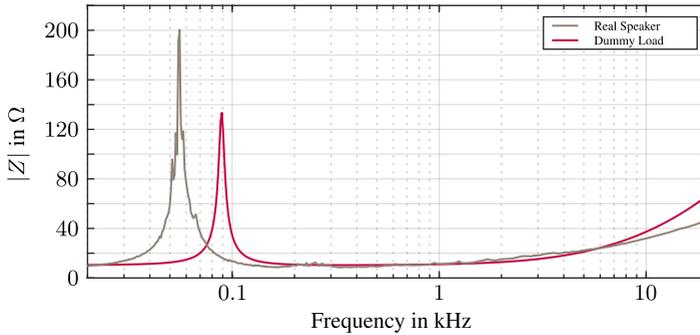
The general shape of both curves remains the same. One prominent peak in impedance can be seen between 50 Hz and 100 Hz and an increase of the impedance towards high frequencies.

The exact circuit of the dummy load is not important for the modeling process however. Different loads or a microphone in front of a real-world speaker cabinet can be used for the measurements. However, the proposed



**Figure 4.7**

*Circuit diagram of the ‘dummy load’. Impedance equivalent filter to a speaker.*



**Figure 4.8**

*Absolute value of the complex impedance in Ohm over frequency. Comparison of the dummy load and a real speaker.*

setup enables the user to simulate different speakers and rooms together with the adapted model.

---

## Evaluation: Error Metrics

---

After a model has been optimized, its performance has to be rated. How similar is the output of the model  $\hat{y}(n, \mathbf{p})$  to the output of the reference system  $y(n)$  for a predefined input signal  $x(n)$ ? There exist several metrics, which describe how similar two signals are or how big the error between the signals is, respectively. There are objective metrics like the root mean square (RMS) error, which are used in optimization methods already. But some of these metrics do not describe the similarity between two signals from a perceptual point of view and might yield relatively high error values even though the signals are perceptually similar. The most common objective error metrics are presented in the following and a perceptually motivated objective score is proposed. Finally techniques for subjective evaluation of audio similarity with listening tests are shown.

### 5.1 Objective Metrics

The root mean square error

$$e_{\text{RMS}} = \sqrt{\frac{\sum_{n=1}^N (y(n) - \hat{y}(n, \mathbf{p}))^2}{N}} \quad (5.1)$$

is often used in optimization problems to describe the differences between the output of a model  $\hat{y}(n, \mathbf{p})$  and the values which are measured or observed  $y(n)$ . It describes the square-root of signal power of the time-domain error between  $\hat{y}(n, \mathbf{p})$  and  $y(n)$ . The RMS error aggregates the

observations made over time into a single measure of error power. Each single error value contributes to the RMS error, which makes this metric sensitive to outliers with large values.

A very similar way of describing the error is the error to signal ratio

$$e_{\text{ESR}} = \frac{\sum_{n=1}^N (y(n) - \hat{y}(n, \mathbf{p}))^2}{\sum_{n=1}^N (y(n))^2} \quad (5.2)$$

which relates the signal power of the error to the signal power of the reference signal. It yields similar values as the root mean square error, but the resulting value will directly express the error power as a percentage of the reference signals power.

The Pearson correlation coefficient

$$r_{y\hat{y}} = \frac{\sum_{n=1}^N (y(n) - \bar{y})(\hat{y}(n, \mathbf{p}) - \bar{\hat{y}})}{\sqrt{\sum_{n=1}^N (y(n) - \bar{y})^2} \sqrt{\sum_{n=1}^N (\hat{y}(n, \mathbf{p}) - \bar{\hat{y}})^2}} \quad (5.3)$$

applied to a statistical sample (measurement) describes the linear correlation between two variables  $y$  and  $\hat{y}$  whose mean values are denoted by  $\bar{y}$  and  $\bar{\hat{y}}$ . It is the covariance of the two variables divided by the product of their standard deviations. The correlation coefficient can take values from  $-1$  to  $1$ , where  $r_{y\hat{y}} = 1$  means that one variable is scaled by a constant positive factor with respect to the other. When the Pearson correlation coefficient takes a value of  $r_{y\hat{y}} = -1$ , one variable is scaled by a constant negative factor with respect to the other and  $r_{y\hat{y}} = 0$  means no correlation at all.

## 5.2 Perceptually Motivated Metrics

Objective metrics provide a mathematical description of an error signal. But the way a human listener perceives the difference between two signals is not always related to these mathematical descriptions. A simple example for this is a sine wave with fundamental frequency  $f_0$  compared to the same sine wave with a phase shift of  $\Delta\phi > 0$ . If the two signals are presented to a human listener, they will sound exactly the same but the root mean square error will exhibit values larger than zero.

Therefore, perceptually motivated error metrics are a good way to rate the quality of the modeling process. Several perceptually motivated metrics already exist, like e.g. PEAQ (perceptual evaluation of audio quality) [164, 165].

ODG	Description
0	Imperceptible
-1	Perceptible but not annoying
-2	Slightly annoying
-3	Annoying
-4	Very annoying

**Table 5.1**

*Objective difference grade (ODG). Results of the PEAQ evaluation with description.*

Two audio signals are fed into the system and they are processed by an ear model, consisting of a transfer function for the outer ear, a time-frequency analysis stage simulating the inner ear and post-processing to enhance the signal for the following stage. The pre-processed signals are fed into a brain model which provides the similarity rating.

The similarity rating is given by the objective difference grade (ODG), which is shown in Tab. 5.1 with the proper descriptions.

PEAQ was initially developed to rate the quality of lossy audio codecs like MP3. Although PEAQ is widely used, the standard is not very well documented and researchers have been trying to re-implement the algorithm but could not explain certain discrepancies between their own implementation and the official implementation from the ITU Radiocommunication Assembly [166].

PEAQ is not suited to rate the quality of virtual analog modeling. The optimization of the algorithm for audio codecs punishes differences in the signals too heavy. While experimenting with PEAQ, very bad scores (ODG around  $-3.5$ ) were achieved for models whose output was not exactly the same as the reference system but certainly not ‘annoying’ or ‘very annoying’.

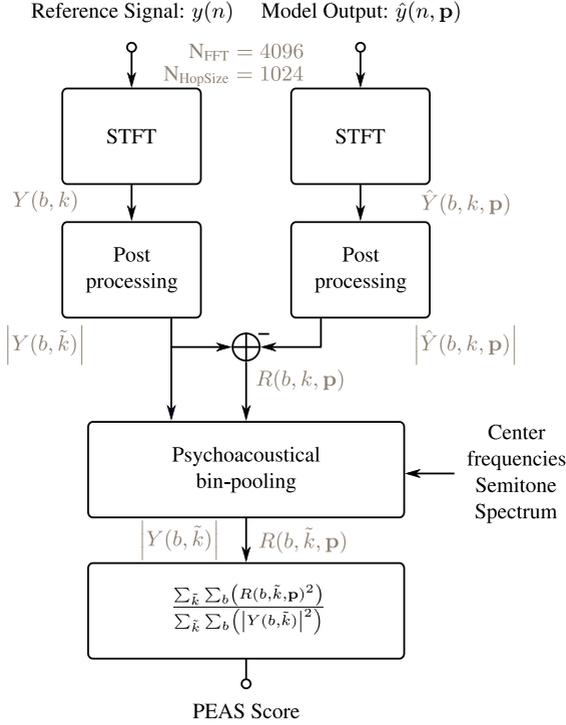
### 5.2.1 Perceptual Evaluation of Audio Similarity (PEAS)

There have been no existing metrics which could be used to evaluate the results of a virtual analog modeling process. Therefore, a simple similarity feature is proposed, giving an estimate to the perceived similarity of two audio signals. It is based on the spectrogram cost-function which was used during the optimization process, as shown by Eq. (4.26).

The input signals for this method are the reference signal  $y(n)$  and the output of the optimized digital model  $\hat{y}(n, \mathbf{p})$  which should sound the same. They are both transformed into the time-frequency domain by a short-time Fourier transform (STFT) with the same block-size and hop-size as in the cost-function, see Sec. 4.3.3. In the post-processing step the magnitude of

the STFT is calculated.

Contrary to the cost-function post processing, the amplitude is kept linear. This ensures that the results of the similarity score are the same if the amplitude of  $y(n)$  and  $\hat{y}(n, \mathbf{p})$  is simultaneously scaled up or down. Even though both signals just exhibit a different loudness they still sound as similar as before so the similarity score should not change.



**Figure 5.1**

*Evaluating audio similarity. Block diagram of the proposed method.*

As Fig. 5.1 depicts, the magnitude spectrograms  $|Y(b, k)|$  and  $|\hat{Y}(b, k, \mathbf{p})|$  are calculated and the residual is calculated by subtracting them from each other. The difference to the cost-function is that the frequency bins of the reference spectrogram  $|Y(b, k)|$  are pooled in the same way as the bins of the residual spectrogram  $R(b, k, \mathbf{p})$  according to the semitone spectrum, which was also used in the cost-function, resulting in the semitone spectrograms  $|Y(b, \tilde{k})|$  and  $R(b, \tilde{k}, \mathbf{p})$ .

The final score

$$\text{PEAS} = \frac{\sum_{\tilde{k}} \sum_b (R(b, \tilde{k}, \mathbf{p}))^2}{\sum_{\tilde{k}} \sum_b (|Y(b, \tilde{k})|^2)} \quad (5.4)$$

is computed by squaring each element of  $R(b, \tilde{k}, \mathbf{p})$  and summing all elements of the resulting matrix and dividing the result by the sum of all squared elements of the reference matrix  $|Y(b, \tilde{k})|$ .

This scales the value computed from the residual spectrogram by the same value computed from the reference spectrogram, resulting in a relative error. If the final score PEAS reaches a value of zero, the two signals  $y(n)$  and  $\hat{y}(n, \mathbf{p})$  are identical. The maximum value of the score could go up to infinity if all values of  $|Y(b, \tilde{k})|$  are zero.

This score is similar to the error to signal ratio but the results are more reliable because phase differences of the signals are ignored. Additionally the frequency spacing is modified by using the semitone spectrum. If this step would be omitted differences for higher frequencies would increase the error value, which does not correspond well to the human perception of frequencies.

### 5.3 Listening Tests

The most reliable way of rating the quality of an audio system from a perceptual point of view, is performing a listening test. This test has to be conducted in a controlled environment, with trained listeners in the best case, to get reliable results. The number of participants needs to be high enough to have statistical significance and the participants also need to know what they are supposed to hear to give meaningful ratings.

Most listening tests have been developed to rate the quality of audio codecs, which is not the goal of this work, although the similarity between two test items should be rated. In contrast to audio codecs, other signals and their characteristics are important when modeling analog audio systems. For an audio codec it might be important that the high-frequency audio content sounds as transparent as possible. When modeling a guitar amplifier, whose cabinet does not exhibit much frequency content above 10–12 kHz, the same considerations do not make sense.

Nevertheless, the fundamental principles of rating audio similarity are the same for audio codecs and virtual analog modeling. For this reason the most common listening tests for evaluating audio similarity are presented in the following sections.

### 5.3.1 ABX – ABC/HR

The ABX listening test is designed to check if a test subject can distinguish a test item from a reference item [167]. The test subject is presented with three items “A”, “B” and “C”. “A” is the reference item with which the test items “B” and “C” have to be compared. The subject has to distinguish which test item is more similar to the reference item. So they have to specify whether “B” or “C” sounds more like the reference, both items cannot be selected. Considering virtual analog modeling, this would be a good test if its purpose was to test if the model is indistinguishable from the reference device. But it would also be beneficial to qualitatively rate the model in comparison to the reference, which is not possible with this test.

Therefore, the ABC/HR listening test was introduced [167]. The test subject is presented with a reference item “A” and has to rate the similarity of items “B” and “C”, where “B” or “C” is randomly selected to be the hidden reference (HR), the remaining item is the output of the digital model. If the test subject rates the hidden reference with a score equal or worse than the model output, the modeling procedure can be considered successful. If there still is an audible difference between items “B” and “C”, the user can rate the difference qualitatively which can be compared to the objective scores discussed in this chapter.

### 5.3.2 MUSHRA

MUSHRA is an acronym for **M**U**l**tiple **S**timuli with **H**idden **R**eference and **A**nchor. It was introduced to rate the perceived quality of lossy audio compression algorithms and is defined by the international telecommunication union recommendation ITU-R BS.1534.3 [168]. A reference item (labeled as such) is presented to the listener and a certain number of test samples have to be rated from 0 to 100 according to the perceived similarity of the items to the reference. 100 means that test item and reference are virtually indistinguishable while 0 stands for an annoying difference between both items. The test items also include a hidden reference and one or more anchors. The ITU recommendation states that one anchor has to be a low-pass filtered version of the reference, where the low-pass has a cut-off frequency of  $f_c = 3.5$  kHz. The anchor is introduced because if all test items are rather similar to the reference, minor artifacts might be rated with very low scores.

MUSHRA is designed to compare several methods (e.g. codecs) with each other, while ABX and ABC/HR are designed to rate one specific method.

---

## Dynamic Range Compression Systems

---

In this chapter, dynamic range compression systems will be analyzed and modeled. First an overview of dynamic range compression in general is given, afterwards the circuit of a specific analog compression system is analyzed in detail. Then the digital model is introduced which is derived from a generic digital dynamic range compressor. A case study, where a popular dynamic range compression system is modeled, is presented before introducing an automated procedure which is able to adapt the digital model to any analog dynamic range compression system. At the end of this chapter, the adapted models are evaluated with objective scores and a listening test for the UREI 1176 limiting amplifier.

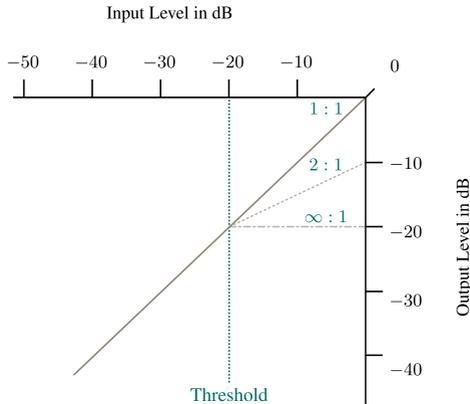
### 6.1 Dynamic Range Compression

A dynamic range compression system controls the gain of the signal depending on the signals characteristics. The system is used to attenuate high signal levels and/or to amplify low signal levels, effectively reducing the dynamic range of the input signal. After loud signal levels have been attenuated a so called make-up gain can be applied to increase the signal power without increasing the maximum level.

A dynamic range compressor can be characterized by its static and dynamic behavior. The static behavior controls how much gain is applied in steady-state while the dynamic behavior determines how fast the compression system acts to sudden changes in the input signal.

### 6.1.1 Static Behavior

The static characteristic relates steady state inputs to steady state outputs. The first characteristic is the threshold. If the input level exceeds the threshold, gain reduction is applied. The compression ratio determines the input output ratio when the gain is reduced. For a ratio of 2 : 1 the output level is increased by 0.5 dB while the input level increases by 1 dB. The ratio controls the amount of compression. For a ratio of  $\infty$  : 1 the output level does not exceed the threshold, the compressor acts as a limiter in this case.

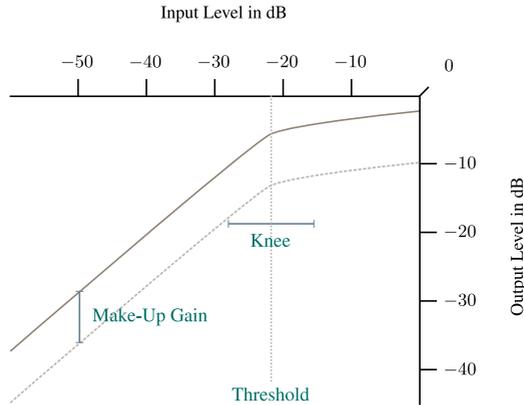


**Figure 6.1**

*Static characteristic of a dynamic range compression system. Signal levels above the threshold can be attenuated by different compression ratios. This example shows a configuration with a hard knee, meaning that the gain reduction is applied instantaneously as soon as the signal level exceeds the threshold.*

Another important characteristic of a compression system is the ‘knee’. It describes the transition from unity ratio to a set ratio. If the level reduction is applied instantly as soon as the input level exceeds the threshold, the knee is referred to as ‘hard knee’ as depicted in Fig. 6.1 which leads to a more noticeable compression. In analog dynamic range compressors, the transition from unity ratio to compression region is usually soft, as depicted in Fig. 6.2. This is called a ‘soft knee’ and the effect of the compression becomes more subtle [169].

Usually the overall volume of the output signal is increased by applying the make-up gain (see Fig. 6.2). With this technique the loudness of the output can be matched to the loudness of the input signal, which is reduced by attenuating higher signal levels. It can also be used to increase the



**Figure 6.2**

*Static characteristic of a dynamic range compression system with soft knee. After the gain reduction is applied the overall signal level is usually increased by a certain factor called the ‘make-up gain’.*

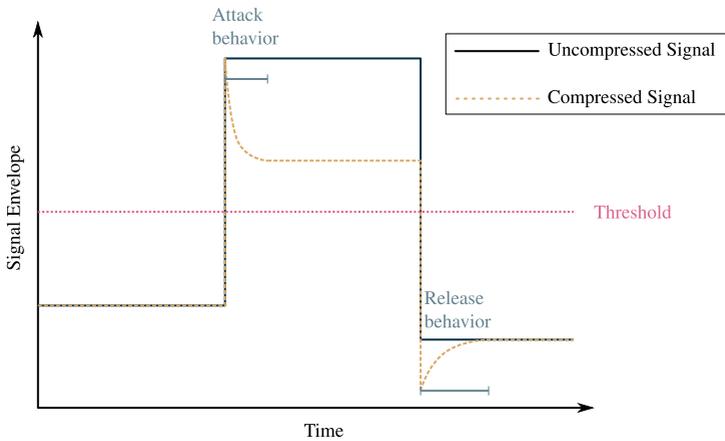
overall loudness of the compressed signal e.g. make an instrument appear more prominent in the mix of an arrangement.

### 6.1.2 Dynamic Behavior

The dynamic parameters determine how the compressor acts over time. If the gain reduction was applied instantly, audible artifacts would be produced in the output signal because the resulting waveform would not be continuously differentiable anymore.

As soon as the input signal level exceeds the threshold, the gain reduction sets in until it has reached the steady state. This is shown in Fig. 6.3 where the solid line depicts the envelope of the input signal (uncompressed) and the dashed line depicts the envelope of the output signal (compressed). The attack time of a dynamic range compressor determines how long it takes the system to reach the steady state. After the signal drops below the threshold again the gain returns to unity. The release time determines how quick the gain returns to its original value [169].

When recreating a dynamic range compression system it is not only important how fast the signal level changes but also the shape of the envelope during attack and release phases. For this reason the attack and release time will be referred to as attack and release behavior, describing also the shape of the envelope.



**Figure 6.3**

*Dynamic behavior of a dynamic range compression system. Envelopes of input signal (uncompressed) and output signal (compressed) over time.*

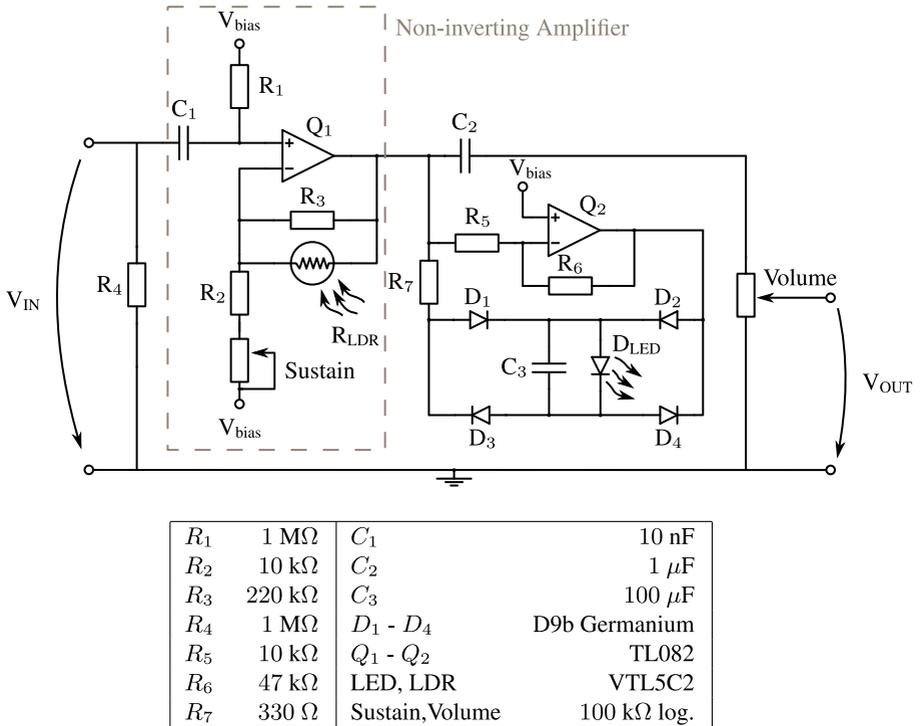
## 6.2 Case Study I: ‘Flatline Compressor’ Circuit Analysis

When an analog system has to be modeled it is always beneficial to analyze said analog system. In [67], the ‘Flatline Compressor’<sup>1</sup>, has been analyzed and modeled in detail. The circuit became quite popular in the ‘do-it-yourself’ (DIY) guitar effect scene, because of its simplistic design with two operational amplifiers and an analog optical isolator (AOI) as the main circuit components.

The most important nonlinear circuit element is the analog optical isolator. It is a two port circuit element which consists of a light emitting diode (LED) and a light dependent resistor (LDR) combined in an opaque enclosure. If the LED is not emitting any light, the resistance of the LDR is high ( $R_{\text{off}} > 1 \text{ M}\Omega$ ). If the LED starts to emit light the resistance of the LDR decreases. For an LED current of  $I_{\text{LED}} = 40 \text{ mA}$  the LDR resistance drops below  $R_{\text{on}} < 300 \Omega$  [171]. Source (LED) and sensor (LDR) are only coupled via the optical channel. Electrical coupling is prevented.

The circuit schematic is shown in Fig. 6.4.  $Q_1$ ,  $Q_2$ ,  $D_{\text{LED}}$  and  $R_{\text{LDR}}$  represent the operational amplifiers and the analog optical isolator. Diodes  $D_1 - D_4$  are **D9b** germanium diodes and the operational amplifiers are **TL082** from Texas Instruments. The bias voltage  $V_{\text{bias}}$  is set to half the supply voltage  $V_+ = 9 \text{ V}$ .  $V_{\text{bias}}$  is used to bias the inputs of the

<sup>1</sup>Designed by John Hollis [170], <http://www.hollis.co.uk/john/circuits.html>



**Figure 6.4**

*Circuit diagram of the 'Flatline Compressor'. Design by John Hollis [170].*

operational amplifiers to avoid a negative voltage supply. The used analog optical isolator is a Perkin Elmer VTL5C2. For readability the enclosure of the device is not depicted but the optical path of the AOI is indicated by the curved arrows in the circuit diagram.

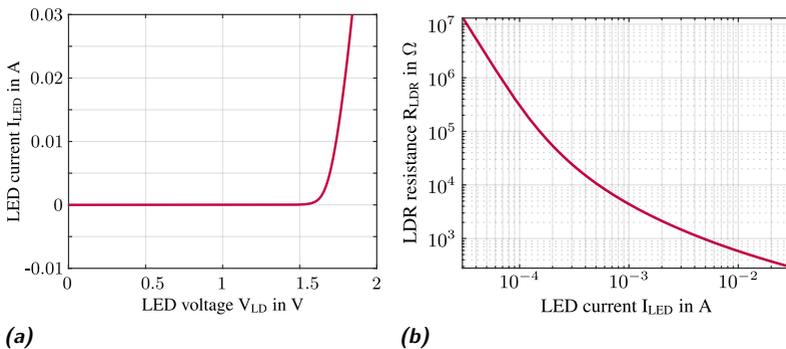
The circuit can be grouped into two main parts. The first main group of circuit elements is formed around the first operational amplifier  $Q_1$ . It forms a non-inverting amplifier consisting of  $Q_1$ ,  $R_2$ ,  $R_3$ , the **sustain** potentiometer and the resistive part of the analog optical isolator  $R_{LDR}$  whose value determines the gain of the non-inverting amplifier.

The second group of circuit elements is used to control the gain value of the first group. The output of the first stage is amplified by an inverting amplifier (around  $Q_2$ ) and rectified by the diodes  $D_1 - D_4$ . The output of the rectifier is smoothed by capacitor  $C_3$ . The smoothed signal controls the

amount of current flowing through the LED  $I_{LED}$ . If the output signal of the non-inverting amplifier overshoots a certain threshold, the LED emits light and the resistance of the LDR decreases. This reduces the gain of the non-inverting amplifier for higher signal levels and thus the dynamic range of the system.

All nonlinear circuit elements were measured before assembling the circuit. The characteristic curves of the diodes have been measured and they matched each other very closely. Nevertheless, the characteristics of the analog optical isolator are more important to determine the overall behavior of the circuit. All nonlinear elements have been measured with the Keithley 2602B system source meter [172].

The current-voltage characteristic of the light emitting diode was measured. The results are shown in Fig. 6.5a. When the voltage across the diode exceeds 1.6 V a current flows and the diode starts emitting light. In



**Figure 6.5**

*Characteristics of the analog optical isolator (a) current-voltage characteristic of the LED (b) resistance-current characteristic of the two port element.*

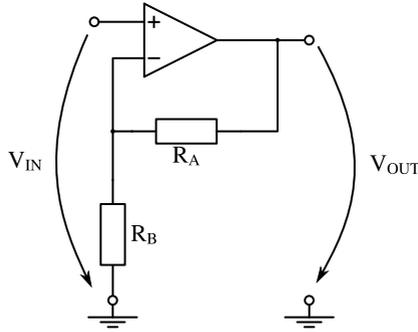
Fig. 6.5b the resistance of the light dependent resistor is plotted against the current flowing through the LED. If the LED is dark only a small current ( $I_{LED} < 0.07$  mA) flows through it and the resistance of the LDR is high,  $R_{LDR} > 1$  M $\Omega$ . The light intensity of the diode increases proportionally to  $I_{LED}$  which leads to a lower resistance of the LDR. For  $I_{LED} > 30$  mA the LDR reaches a value of  $R_{LDR} < 300$   $\Omega$ , which corresponds with the datasheet [171].

### 6.2.1 Static Behavior

The behavior of the circuit in steady state was analyzed by dividing the circuit into two parts. The first part is the non-inverting amplifier around  $Q_1$  as shown in Fig. 6.4. The gain for an ideal non-inverting amplifier, as shown in Fig. 6.6, is

$$g = 1 + \frac{R_A}{R_B} , \quad (6.1)$$

according to [145].



**Figure 6.6**  
*Non-inverting amplifier.*

As the circuit diagram shows, the resistance  $R_A$  consists of the parallel connection of  $R_3$  and  $R_{LDR}$ ,

$$R_A = \frac{R_3 \cdot R_{LDR}}{R_3 + R_{LDR}} . \quad (6.2)$$

The resistance  $R_B$  is the series connection of the sustain potentiometer and  $R_2$

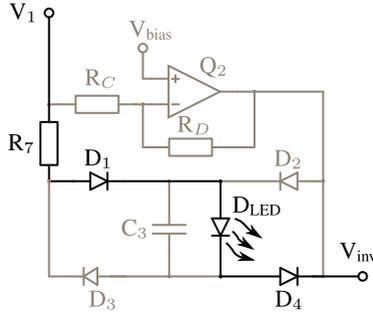
$$R_B = R_2 + R_{sustain} . \quad (6.3)$$

Inserting Eq. (6.2) and Eq. (6.3) into Eq. (6.1) yields

$$g = 1 + \frac{R_3 \cdot R_{LDR}}{(R_2 + R_{sustain}) \cdot (R_3 + R_{LDR})} , \quad (6.4)$$

which is the gain of the non-inverting amplifier in the circuit depending on the value of  $R_{LDR}$ .

Since the value of  $R_{LDR}$  is dependent on the current-flow through the LED, the second stage of the circuit is analyzed. The goal is an expression for the current through the LED. It depends on the output voltage of the



**Figure 6.7**

Second operational amplifier stage of the ‘Flatline Compressor’ circuit, shown in Fig. 6.4.

non-inverting amplifier from the first stage  $V_1$ , relating the value of the LDR  $R_{LDR}$  to a specific LED current  $I_{LED}$  or voltage  $V_{LED}$ .

The second stage is shown in Fig. 6.7. According to [145], the gain of an inverting amplifier can be expressed as

$$g_{inv} = -\frac{R_D}{R_C} . \quad (6.5)$$

This equation can be directly applied to the inverting amplifier around  $Q_2$  to calculate its output voltage  $V_{inv}$

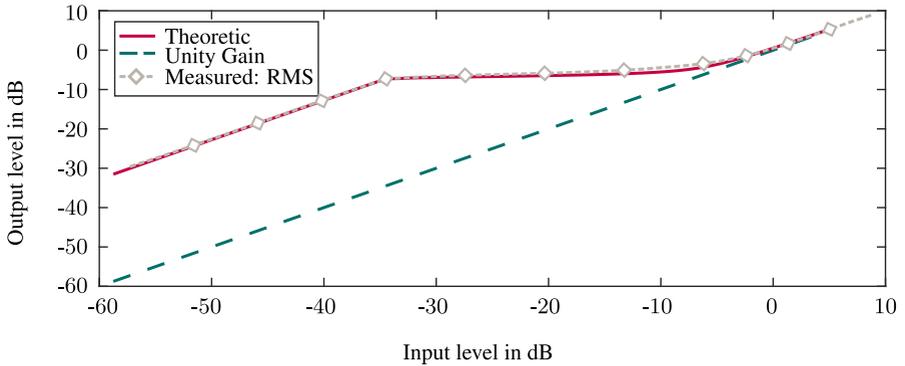
$$V_{inv} = -\frac{R_D}{R_C} \cdot V_1 = -\frac{R_6}{R_5} \cdot V_1 = -4.7 \cdot V_1 . \quad (6.6)$$

Using this knowledge and with the assumption that capacitor  $C_3$  is an open-circuit and can be omitted in steady state, the second stage of the circuit can be simplified. As depicted in Fig. 6.7 the circuit reduces to a series connection of  $R_7$ , two diodes  $D_1$  and  $D_3$  (or  $D_2$  and  $D_4$  respectively) and the LED. An expression for the output voltage of the first circuit stage can be derived, where  $V_D$  are the voltages across the germanium diodes  $D_1$ – $D_4$ .

$$V_1 = \frac{R_7 \cdot I_{LED} + 2 \cdot V_D + V_{LED}}{1 + 4.7} \quad (6.7)$$

With the measured characteristics of the diodes and the LED, Eq. (6.7) can be evaluated for every possible LED current  $I_{LED}$ . Afterwards, utilizing the measured current-resistance characteristic of the AOI, Eq. (6.4) can be solved and the input voltage can be related to the output voltage by calculating,

$$V_{IN} = \frac{V_{OUT}}{g} = \frac{V_1}{g} . \quad (6.8)$$



**Figure 6.8**

*Static behavior of the compressor circuit. Calculated and measured characteristic.*

This yields the static curve for the circuit, relating the level of input voltage to the level of output voltage. This behavior is shown in Fig. 6.8. The solid line shows the result of the presented theoretic calculations which were derived from the circuit and the measurement of the electronic components. The dashed straight line indicates unity gain (output has the same level as input).

To see if the calculations are correct the static mapping curve of the circuit was also measured. A sine wave with a frequency of  $f_0 = 1$  kHz with RMS levels from  $-60$  dB to  $6$  dB was sent through the circuit and the RMS level of the output signal was measured. The resulting curve is shown in Fig. 6.8. It matches the theoretical static curve very closely except for input signal levels between  $-15$  dB and  $-5$  dB, where the measured curve deviates slightly from the calculated one. This deviation might originate from the assumption of ideal operational amplifiers during the calculation of the static curve. When measuring the static curve with a higher supply voltage ( $V_+ \approx 20$  V), the error between calculated and measured curves decreases because the operational amplifiers (especially the inverting amplifier around  $Q_2$ ) are clipping at higher signal levels.

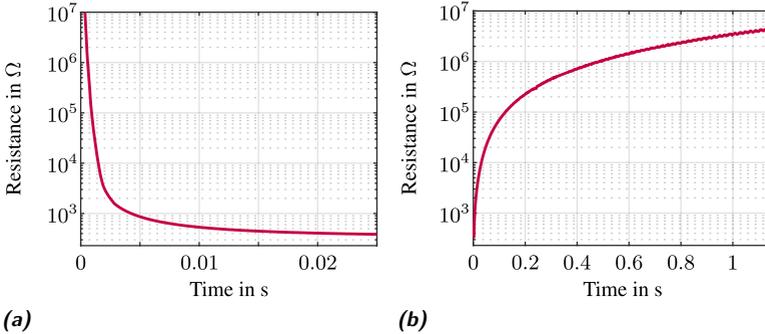
### 6.2.2 Dynamic Behavior

The dynamic behavior of a compression system describes how the system reacts to changes in the input signal over time. For dynamic range compression there are two important cases. If the input signal rises above the threshold the attack behavior determines how fast the level of the signal is reduced. If the input signal drops below the threshold the release behavior

determines how fast the level reduction is reversed.

In case of the circuit, the parallel connection of AOI and  $C_3$  determine the time-dependent behavior. If the LED current changes, the resistance of the LDR changes as well, influencing the amplification of the non-inverting amplifier. If the output signal has a high level, the light intensity inside the optical channel of the AOI gets higher, which reduces the resistance of the light dependent resistor. This, in turn, reduces the gain-factor of the non-inverting amplifier (see Eq. (6.4)). For a silent input signal, the LED is turned off and the LDR reaches its maximum resistance, resulting in the maximum gain factor of the non-inverting amplifier stage.

First, the *turn-on* and *turn-off* behavior of the AOI was measured by tracking the resistance of the second port while turning the LED on or off. The resistance of the LDR when switching the LED current from  $i_{LED} = 0$  mA to  $i_{LED} = 10$  mA at  $t = 0$  s is shown in Fig. 6.9a. The resistance drops relatively fast, after the LED has turned on. After  $t = 5$  ms, the resistance has dropped below  $R_{LDR} < 1$  k $\Omega$  from an initial value of  $R_{LDR} > 10$  M $\Omega$ . It takes about 15 ms for the resistance to reach its final value for this light intensity.

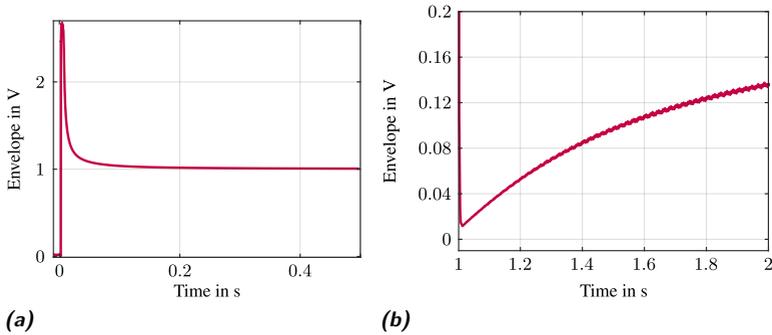


**Figure 6.9**

(a) *turn-on* and (b) *turn-off* behavior of the analog optical isolator. Resistance of the second port was measured over time while altering the LED current at the first port.

In Fig. 6.9b, the *turn-off* case is shown. The current was switched from  $i_{LED} = 10$  mA to  $i_{LED} = 0$  mA at  $t = 0$  s. It is obvious that the AOI is not acting as fast, when the LED is turned off. The resistance rises from its initial value to  $R_{LDR} = 1$  M $\Omega$  after  $t = 500$  ms, which is about 100 times slower as the reversed change in the *turn-on* case.

To relate these measurements to the behavior of the whole circuit, a sine wave with fundamental frequency  $f_0 = 1$  kHz with an abrupt change

**Figure 6.10**

*Dynamic behavior of the 'Flatline Compressor' (a) attack behavior (b) release behavior.*

in amplitude was sent through the reference device. The amplitude of the input signal increases from 0.01 V to 1 V at  $t = 0$  s and drops from 1 V to 0.01 V at  $t = 1$  s. The envelope of the output signal is depicted in Fig. 6.10.

The attack behavior of the whole circuit ( $t = 0$  s) is shown in Fig. 6.10a. When comparing this to the previous measurement (Fig. 6.9a), the reaction time of the whole circuit is considerably slower. The steady-state is reached after approximately  $t = 100$  ms, which is about 10 times slower than the reaction time of the AOI. This behavior corresponds with the effect of capacitor  $C_3$ , parallel to the LED. This capacitor acts as a low pass filter for the voltage across the LED. The attack behavior is proportional to the capacitance of  $C_3$ . Larger values yield longer attack times, because it takes the capacitor longer to charge.

For the release behavior, the AOI is the dominant element. The envelope depicted in Fig. 6.10b, corresponds well with the AOI resistance measurements shown in Fig. 6.9b.

The measured curves for attack and release behavior of the circuit are dependent on the characteristics of  $C_3$  and the AOI. A change in these characteristics, e.g. due to production tolerances, would lead to a different attack and release behavior. To be able to capture the behavior for reference devices with different characteristics, the parameters of a digital model are iteratively optimized.

## 6.3 Case Study II: UREI 1176 LN

In 1966 *Universal Audio* founder Bill Putnam designed a solid state dynamic range compressor, which used a junction field-effect transistor (JFET) as the gain-controlling element. It was the first compressor completely based on transistor circuitry and is still widely used in today's music production.

The circuit has been revised many times over the years. In [173], the popular 'Revision D' circuitry is analyzed and modeled. The main feature of this revision is the introduction of low-noise circuitry in the audio path and its integration onto the main circuit board. The used hardware is a DIY kit of the revision D circuitry [174].

The user controls on the device itself are:

- Input gain from  $-\infty$  to 0 dB to scale the input signal
- Output gain from  $-\infty$  to 0 dB to apply the make-up gain
- Attack and release time from 1 (slow) to 7 (fast)
- Ratio buttons (4 : 1, 8 : 1, 12 : 1, 20 : 1) which change the static characteristic of the device

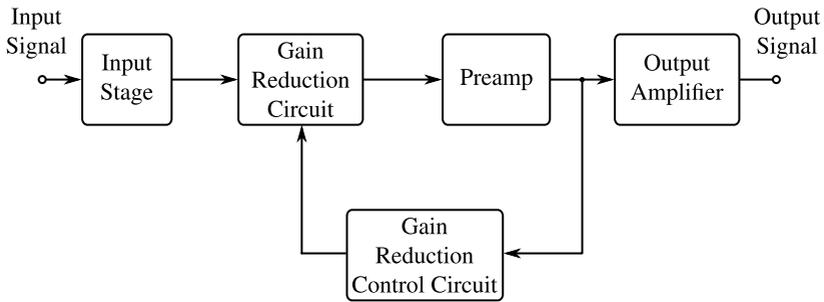
Another option, which was not originally intended, is the so called 'all buttons in' mode, where all four ratio buttons are pressed at the same time. The resulting static curve of the compressor behaves more like a limiter and the signal is slightly distorted, giving it a rich and full sound, which is very popular among audio engineers.

### 6.3.1 Principle of Operation

The 1176LN uses a JFET transistor as a voltage controlled resistor (VCR) which is part of a voltage divider in the audio signals path. The control voltage is generated by a level sensing circuit monitoring the output [174].

The principle of operation of the 1176 is shown in Fig. 6.11. The input signal is fed into an input stage which is used to attenuate the input so that the next stage operates in its linear region. Additional input damping can be applied to control the amount of compression. Afterwards the signal is fed into the gain reduction circuit including the voltage divider with the variable resistor (JFET). The output of this stage is amplified and the amplified signal is the input to the gain reduction control circuit which provides the control voltage for the JFET resistor. The last stage of the signal flow is the output amplifier, which contains an output transformer, which contributes to the specific sound of the compressor [174].

An analysis of the whole circuit would be too extensive and unnecessary, since the aim of this work is to *identify* the system by measuring its input

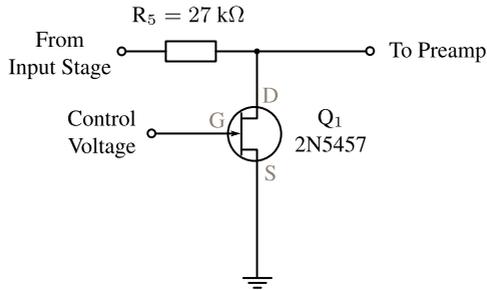
**Figure 6.11**

*Block diagram of the UREI 1176 limiting amplifier.*

output relations. Nevertheless, the most important parts will be discussed in this section.

The most important stage in the circuit is the voltage controlled variable resistor and is shown in Fig. 6.12. It is a voltage divider with  $R_5$  as the series element. The JFET  $Q_1$  acts as a voltage controlled resistor to ground. The resistance of the n-doped JFET channel depends on its gate-source voltage. When no voltage is applied between gate and source, the n-doped channel of the JFET is conducting and the channel resistance is low. If the voltage between gate and source is reduced a depletion layer ‘grows’ into the channel, reducing its conductivity and thus increasing its resistance. When the gate-source voltage reaches a certain threshold  $V_{GS} = V_{th}$  the channel is not conducting and the JFET resistance is very high. In this circuit, the bias voltage of the JFET is set close to the threshold voltage which means that the input signal is not attenuated. When the gate-source voltage increases, the channel of the JFET starts conducting and its resistance is reduced, which means that the signal is attenuated more before passing through the preamp stage of the circuit. This results in an attenuated or compressed signal. The afore mentioned input attenuation results in a very low drain-source voltage (mV range) of the JFET, assuring a linear region of operation.

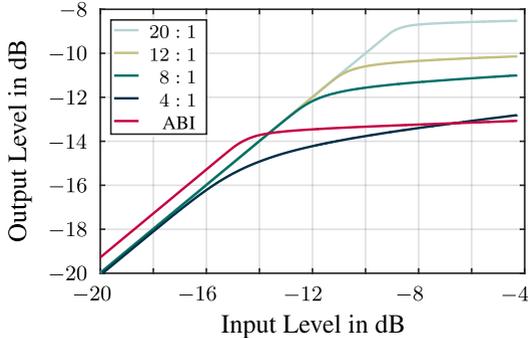
The next important stage is the gain reduction control circuit. its purpose is to supply the control voltage of the JFET transistor which should be proportional to the level of the signal coming out of the preamplifier stage. Without going into too much detail, the signal is sent to two phase-inverted amplifiers and rectified. Signal amplitudes above the threshold modulate the rectified signal while attack and release potentiometers control the DC level of the signal. Louder parts of the input signal will now lead to a conducting JFET channel, decreasing its resistance and thus decreasing the amplitude.



**Figure 6.12**

*Voltage controlled resistor (VCR) stage of the 1176 LN circuit [174]. The channel of the JFET acts as a VCR forming a voltage divider with  $R_5$ . Drain, gate and source terminals of the JFET are marked with  $D$ ,  $G$  and  $S$ .*

The static behavior of the 1176 is measured for all ratio buttons [173], including the ‘all buttons in’ setting, where all four ratio buttons are pressed at the same time. The results of the measurements can be seen in Fig. 6.13.



**Figure 6.13**

*Measured static characteristics for the different ratio buttons of the UREI 1176LN. ABI stands for ‘All Buttons In’.*

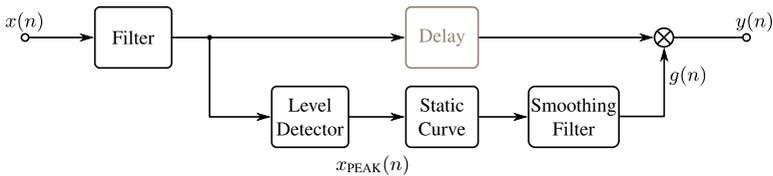
The knee width of the static curves scales with the ratio setting. The 20 : 1 ratio has the lowest knee width which is the transition between unity gain and attenuated region. The 4 : 1 ratio setting has the biggest knee width with a soft transition. The all buttons in (ABI) mode features some

amplification in the linear region and high compression in the saturated region, behaving almost like a limiter. The measured curves for all ratio settings are stored and can be used in the digital model.

The dynamic behavior of the 1176 is very different from the dynamic behavior of the ‘Flatline Compressor’ described in the previous section. The use of the JFET transistor allows much faster attack times than the use of the AOI. The user manual of the 1176 states attack times between 20  $\mu\text{s}$  and 800  $\mu\text{s}$ . A detailed analysis of the dynamic behavior of the 1176 will be conducted in the following.

## 6.4 Digital Model

The digital model used in this work is based on the generic digital dynamic range compressor model from [98]. Its basic structure has been preserved, but several blocks were added or adapted to create a more flexible model which is able to emulate any dynamic range compression system.



**Figure 6.14**

*Generic model of a dynamic range compressor according to [98]. The direct path features an optional delay.*

The model consists of a direct path and a side-chain path. The side-chain features a level detector and the static characteristic of the DRC, mapping the level of the input signal to the gain value. A smoothing filter is used to filter out sudden changes in the time-variable gain signal  $g(n)$ . Please note, that the delay in the direct path of the original model has been removed, because the examined analog DRC systems do not feature look ahead.

The used model is shown in Fig. 6.15. The level detector is a simple peak detector. The absolute value of the current input sample is filtered by a first order digital low-pass filter which returns the weighted average of the absolute value of the current input sample and the previous input sample. The difference equation of the filter is

$$x_{\text{PEAK}}(n) = |x(n)| \cdot c_{\text{LVL}} + x_{\text{PEAK}}(n-1) \cdot (1 - c_{\text{LVL}}) , \quad (6.9)$$

with the filter coefficient  $c_{LVL}$ , which can take different values, depending on the input signals history

$$c_{LVL} = \begin{cases} AT+ & \text{for } x(n) > 0 \text{ and } |x(n)| > |x(n-1)| \\ RT+ & \text{for } x(n) > 0 \text{ and } |x(n)| < |x(n-1)| \\ AT- & \text{for } x(n) < 0 \text{ and } |x(n)| > |x(n-1)| \\ RT- & \text{for } x(n) < 0 \text{ and } |x(n)| < |x(n-1)| \end{cases} . \quad (6.10)$$

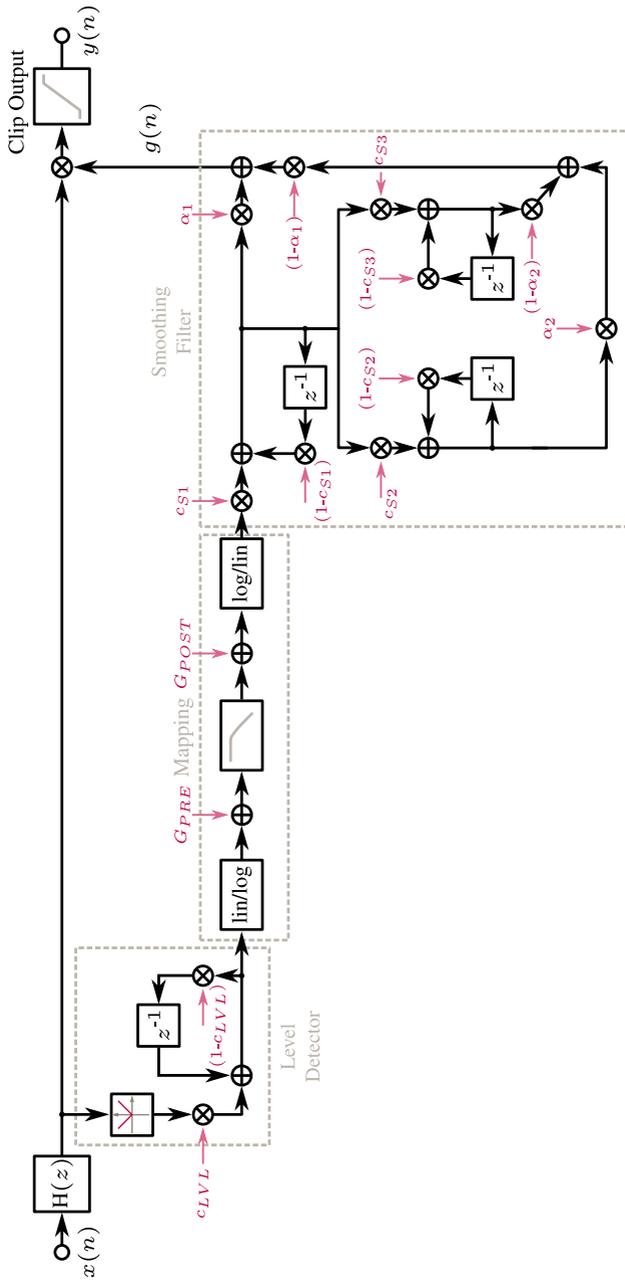
$AT\pm$  and  $RT\pm$  are the actual parameters of the level detector and their range has been limited to  $AT\pm \in [0, 1]$  and  $RT\pm \in [0, 1]$  respectively. Selecting different values of  $AT$  and  $RT$  for positive and negative amplitudes of the input gives the model the possibility of introducing odd harmonics in the output signal.

The static curve, which is utilized in the digital model, maps the level of the input signal (the output of the level detector) to a specific gain value, according to the measured static characteristic of the reference device. To be able to scale the static curve, a pre-gain  $G_{PRE}$  has been added before the actual mapping occurs and a post-gain  $G_{POST}$  afterwards. The input signal levels for the measurement of the static curve range from  $-80$  dB to  $0$  dB. The mapping curve which is used in the digital model is calculated by dividing the output voltages by the corresponding input voltages. The distribution of amplitudes of the measurement input is logarithmic, to have a good amplitude resolution for low signal levels. The mapping curve is saved and used as a lookup table with linear interpolation in the digital model.

After calculating the gain factor it is necessary to prevent it from changing too fast. Sudden changes in the gain factor will result in an unpleasant listening experience for the user. The smoothing filter is the component which has been changed the most in respect to the model proposed in [98]. The original smoothing filter was a simple first order digital low-pass filter, with different coefficients for attack and release. Regarding  $x(n)$  and  $y(n)$  as the input and output signal of the first order low-pass, its difference equation can be expressed as,

$$y(n) = c_{si} \cdot x(n) + (1 - c_{si}) \cdot y(n-1) . \quad (6.11)$$

The new smoothing filter was developed in [175] and is a combination of three first order low-pass filters, which are connected as shown in Fig. 6.15. The output of the mapping curve is smoothed by the first low-pass filter and is then fed into two parallel low-pass filters. The output of the second and third filter are weighted with complementary weighting factors  $\alpha_2$  and  $(1 - \alpha_2)$  before they are added together. The resulting signal and the output of the first low-pass are weighted with the complementary weighting factors  $\alpha_1$  and  $(1 - \alpha_1)$  to form the variable gain signal  $g(n)$ . By using



**Figure 6.15**

*Detailed model of the proposed dynamic range compression model. Variable parameters are shown in red.*

the complementary weighting factors  $\alpha_{1,2}$  in each branch no energy is lost or added by the smoothing filter on average. Combining the filters in this structure allows for flexible shaping of the dynamic behavior, being able to adapt to many different dynamic range compression systems.

The parameters for the smoothing filter are the coefficients of each first order low-pass filter and the two weighting parameters  $\alpha$ . Each low-pass has different coefficients for attack and release case

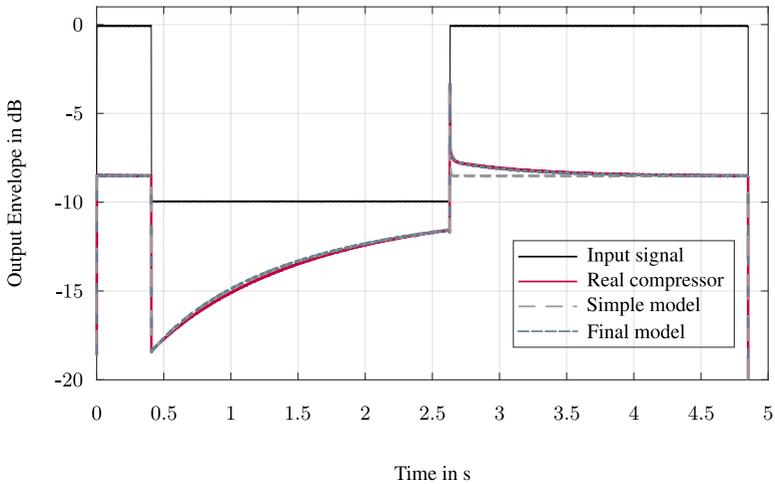
$$c_{si} = \begin{cases} AT_i & \text{for } |x(n)| > |x(n-1)| \\ RT_i & \text{for } |x(n)| < |x(n-1)| \end{cases} . \quad (6.12)$$

$c_{si} \in [0, 1]$  and  $\alpha_j \in [0, 1]$  is valid for all parameters of the smoothing filter.  $i = 1, 2, 3$  is the index of each first-order low-pass filter and  $j = 1, 2$  is the index of the weighting factors.

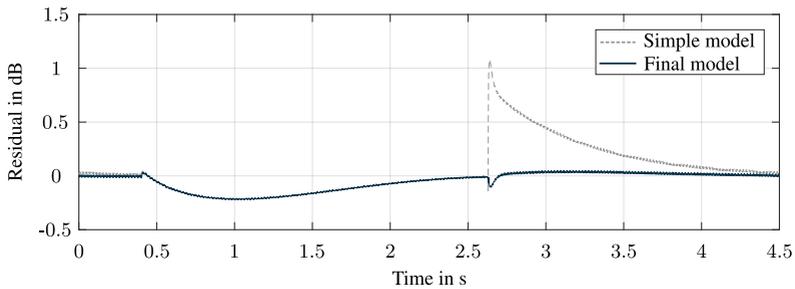
The examination of the dynamic parameters of the UREI 1176LN shall serve as an example for the parallelized smoothing filter [173]. The shape of the envelope suggested that it can be attained by combining several first order smoothing filters with different decay times. To achieve the desired response, three different smoothing filters are combined together as shown in Fig. 6.15. The three low-pass filters are now able to recreate the more complex attack behavior of the 1176LN, as shown in Fig. 6.16, for  $t > 2.5$  s. The three attack times of the smoothing filters are combined to reduce the error between digital model and the measured curves from the reference device. In Fig. 6.16b the residual between the envelope of the measured reference signal and the model with a simple smoothing filter (only one first order low-pass filter) and the residual for the proposed smoothing filter. The magnitude of the residual decreases from a maximum value of 1 dB at  $t \approx 2.6$  s to a magnitude below 0.1 dB with the proposed smoothing filter.

Because analog compression systems only have limited amplification capabilities the output is hard-clipped to the amplitude range of the analog reference system after the adaptive gain is applied to the input signal. This is only relevant in attack cases with high level differences, before the compression system has reduced the attack overshoot.

To be able to incorporate the linear behavior of the reference system, its impulse response has been measured, as described in Sec. 4.4. It is directly used in the digital model, before the actual compression is performed, as shown by Fig. 6.15.



(a) Signal envelopes



(b) Residuals between analog reference and the two digital models

**Figure 6.16**

Signal envelopes and residual of the 1176 with simple and modified smoothing filter. (a) shows the envelopes of a sine wave with  $f_0 = 1000$  Hz and varying amplitudes and the responses of the reference system as well as the digital model. (b) depicts the residuals between original signal and digital model outputs.

All parameters shown in Fig. 6.15 are collected in a parameter vector

$$\mathbf{p}_{\text{comp}} = \begin{pmatrix} G_{\text{PRE}} \\ G_{\text{POST}} \\ c_{\text{LVL,AT+}} \\ c_{\text{LVL,RT+}} \\ c_{\text{LVL,AT-}} \\ c_{\text{LVL,RT-}} \\ c_{\text{s1,AT}} \\ c_{\text{s1,RT}} \\ c_{\text{s2,AT}} \\ c_{\text{s2,RT}} \\ c_{\text{s3,AT}} \\ c_{\text{s3,RT}} \end{pmatrix} \quad (6.13)$$

The parameters  $G_{\text{PRE}}$  and  $G_{\text{POST}}$  can scale the static mapping curve of the model to correct any measurement inaccuracies. The  $c_i$  parameters are the filter coefficients for level detector and smoothing filter.  $\alpha_1$  and  $\alpha_2$  are used to weight the outputs of the parallel smoothing filter structure.

The model features a feedforward structure, while both analyzed DRC systems have a feedback architecture. A feedforward architecture uses the input signal as side-chain input while the feedback architecture uses the output signal, after the time-variable gain has been applied as side-chain input. Early compression systems were built with a feedback side-chain, because possible errors from the gain stage could be rectified like this. According to Giannoulis et al. it will not affect the performance of a digital system if it is implemented with a feedforward or feedback design [169].

## 6.5 Modeling

In this section the procedure to model an analog dynamic range control system will be explained. The modeling procedure is able to model the reference device in a *specific* setting. That means that a potential user will not be able to use the same knobs and controls as on the reference device. The second part of this section will show a way to integrate the user controls into the emulated version of the DRC system.

### 6.5.1 Input Signals

To model the reference device in a specific setting, appropriate input signals have to be designed to excite the relevant characteristics of the DRC system and to calculate the error between digital model and analog reference device. In the case of dynamic range compression systems the input signal has to cover all relevant signal levels. Otherwise the optimization would be meaningless because the behavior of the reference device to specific input conditions would be unknown. To ensure that the model is able to adapt to various input signal types the parameter set is optimized for different input signals. At first a multi-frequency sine wave which changes its amplitude abruptly is send through the reference device to record the desired output. The sine wave is synthesized using Schroeder phases [176] which minimize the crest-factor for a multi-frequency sine wave signal. The sine wave is comprised of frequencies from  $f_0 = 50$  Hz to  $f_{39} = 2000$  Hz (the 40<sup>th</sup> harmonic of  $f_0$ ), to cover the typical frequency range of an electric guitar.

First  $K = 40$  sine waves with Schroeder phases are generated and added up

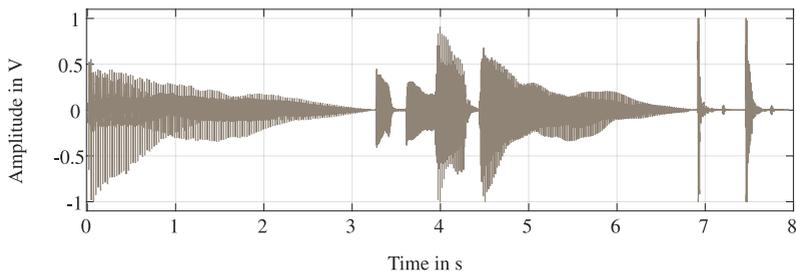
$$x_{\text{sin}}(n) = \sum_{k=1}^K \frac{1}{K} \sin \left( 2\pi k \frac{f_0}{f_s} n + \pi \frac{k^2}{K} \right) . \quad (6.14)$$

Afterwards  $x_{\sin}(n)$  is normalized to exhibit maximum amplitudes from  $x_{\min} = -1$  to  $x_{\max} = 1$ . An amplitude weighting sequence  $a(n)$  is generated which contains 9 randomly selected signal levels from  $a_{\min} = -50$  dB to  $a_{\max} = 0$  dB. The input signal

$$x_{\text{in}}(n) = a(n) \cdot x_{\sin}(n) \quad (6.15)$$

is generated by multiplying the amplitude weighting sequence with the sine wave. Each signal level changes abruptly after  $\Delta t \approx 0.55$  s resulting in a signal of an overall length of 5 s.

The second input signal  $x_{\text{in},2}(n)$  is used to refine the parameters adapted to the synthesized multi-frequency sine wave to more realistic signal conditions, since the reference device is intended to be used on electric guitars. The time domain plot of the input signal is shown in Fig. 6.17 and it consists of dry recordings from an electric bass (0 s – 3.1 s), electric guitar (3.1 s – 6.8 s) and a short drum recording (6.8 s – 8 s). The transient signals from bass drum, snare drum and hi-hat are helpful when adjusting the dynamic parameters of the model.



**Figure 6.17**

*Time domain plot of the second optimization input  $x_{\text{in},2}(n)$ .*

### 6.5.2 Initial Parameter Set

Before optimizing the parameters of the digital model a grid search is performed. To contain the computational load of the grid search, the result of the cost-function for the parameters with the highest impact on the model output  $G_{\text{POST}}$  and  $G_{\text{PRE}}$  is calculated on a grid  $G_{\text{MIN}} = -60$  dB to  $G_{\text{MAX}} = 60$  dB and step size  $\Delta G = 4$  dB for the measured gain mapping curve. The combination yielding the lowest error is chosen for the initial parameter set. The input signal for the grid search was  $x_{\text{in}}(n)$ .

The filter coefficients of the first low-pass in the smoothing filter,  $c_{S1,AT}$  and  $c_{S1,RT}$  are chosen for a second grid search because  $\alpha_1$  is initially set to 1 so that the other low-pass filters have no effect on the output. The

parameter range in this case is  $c_{S3} \in [0, 1]$  with a step size of  $\Delta c \approx 0.033$  and the results corresponding to the smallest error value are chosen as initial parameter values.

The cost-function is set to give back the difference of the output signals envelopes. First, the envelopes of reference output and model output are calculated with a peak-picking technique and afterwards their quadratic error is the result of the cost-function which has to be minimized.

All other time constants of the filters were initialized according to Zölzer et al. [177]. The step response of a first order recursive low-pass is used to relate a time constant in seconds to the filter coefficients

$$c = 1 - e^{\frac{-2.2}{f_s \cdot t_c}} \quad , \quad (6.16)$$

where  $f_s$  is the sampling frequency and  $t_c$  is the time constant in seconds. The initial values of  $t_c$  were empirically analyzed and set to  $t_{c,\text{attack}} = 1$  ms and  $t_{c,\text{release}} = 0.5$  s.

### 6.5.3 Modeling Procedure

Before optimizing the parameters itself, all measurable characteristics of the reference device are extracted and directly used in the model. The impulse response of the reference device is extracted, as described in Sec. 4.4. The amplitude of the identification signal is set to a value below the threshold to ensure that no compression occurs while measuring the impulse response. It is used as the first block,  $H(z)$ , in the DRC model shown in Fig. 6.15.

Then the static characteristic of the reference device is measured. A sine wave with frequency  $f_0 = 1$  kHz and signal levels from  $-60$  dB to  $6$  dB is sent through the reference device. Each measurement is repeated five times. The RMS value of each measurement is calculated and the mean value of the repeated measurement is used as the final result. The mapping curve which relates input level to gain value is calculated from the static characteristic and used as a look up table with linear interpolation in the digital model.

Now the parameter optimization is performed, according to the Levenberg – Marquardt method (see Sec. 4.2.3). The optimization is divided into two parts for the different input signals. At first, the error between the envelopes is minimized for the first input (multi-frequency sine wave) with the peak-picking envelope cost-function described in Sec. 4.3.2. After the parameters have been optimized the parameter set is saved and used as the initial parameter set for the next optimization.

The second optimization is used to improve the parameters for real-world signals, like electric guitars. The second input signal with length of

$N$  is used and the cost-function returns the time-domain error

$$C(\mathbf{p}) = \sum_{n=1}^N (y(n) - \hat{y}(n, \mathbf{p}))^2 \quad . \quad (6.17)$$

The optimized parameter set is the final result of the modeling procedure. Together with the filter  $H(z)$  and the measured static characteristic the model emulates the behavior of the reference device in one specific setting.

#### 6.5.4 Integrating User Controls: UREI 1176LN

The model is now able to emulate a reference device in one specific setting, but if the user controls are altered the model is not valid anymore and it is not clear how the parameters of the model should be altered to recreate the new set of user controls. Therefore, the user controls of a specific reference device have been analyzed to see how much they influence each other [173, 175]. The chosen reference device was the UREI 1176LN. As mentioned in Sec. 6.3, the user controls are input- and output-gain, attack- and release-time and the ratio buttons.

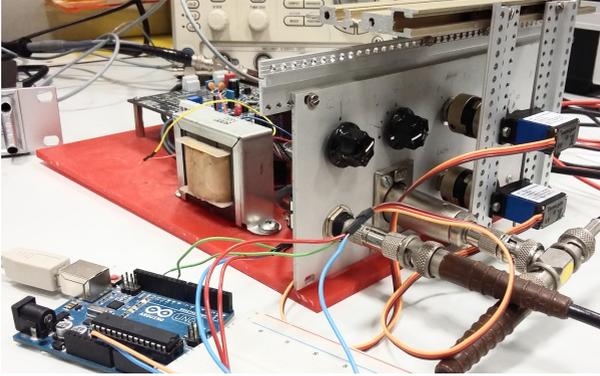
After measuring and using the reference device the hypothesis was constructed that the dynamic parameters influence each other, namely the release control knob position alters the attack time of the reference device. To correctly model the behavior of the reference, this influence has to be identified and incorporated into the digital model.

To quantify the influence the user controls have on the parameters of the digital model a test machine was built. It consists of servo motors mounted to the front plate of the reference device, which are driven by a micro-controller and automatically adjust the angle of the knobs.

The test machine is shown in Fig. 6.18. The evaluation script of the dynamic evaluation sets the position of the knobs and the reference device is measured and the optimization routine is carried out. The model parameters and user control positions are stored and the process is repeated for the next user control combination.

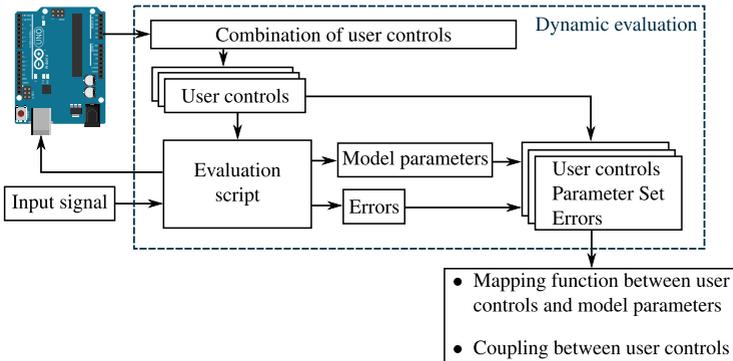
A set of model parameters is adapted for every possible combination of the attack and release user controls to find the relationship between the angles of the control knobs and the set of model parameters. This relationship can then be used to adjust the parameters of the model. This evaluation of the dynamic parameters is depicted in Fig. 6.19. The result of this analysis is a mapping function, between user controls and model parameters and the coupling between the user controls.

The results confirmed the initial hypothesis that the release control influences the attack time, as Fig. 6.20 shows. The effective measured attack time is shown over the attack and release control knobs (whose angular position is labeled from 1 to 9).



**Figure 6.18**

Test machine for altering the user controls [173]. An Arduino micro-controller board can be seen on the left. The servo motors are attached to the front plate of the reference device, a DIY kit of the UREI 1176LN Rev. D.

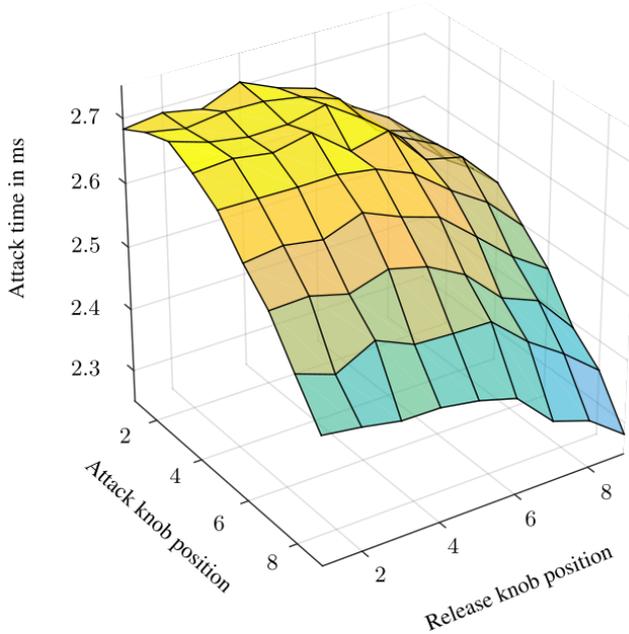


**Figure 6.19**

Dynamic evaluation of the user controls [173].

The value of the attack time is decreasing along the attack control axis as it is expected to be, since a high knob position implies a shorter attack time. The attack also decreases for a higher value of the release control. This curve suggests a second order relationship between the attack time and the angles of the attack and release control knobs  $\theta_{at}$  and  $\theta_{rt}$ .

The model is flexible enough to reach the behavior of a reference device, but in case of the UREI 1176LN only five user controls are available to



**Figure 6.20**

*Relation between user control positions and attack time [173].*

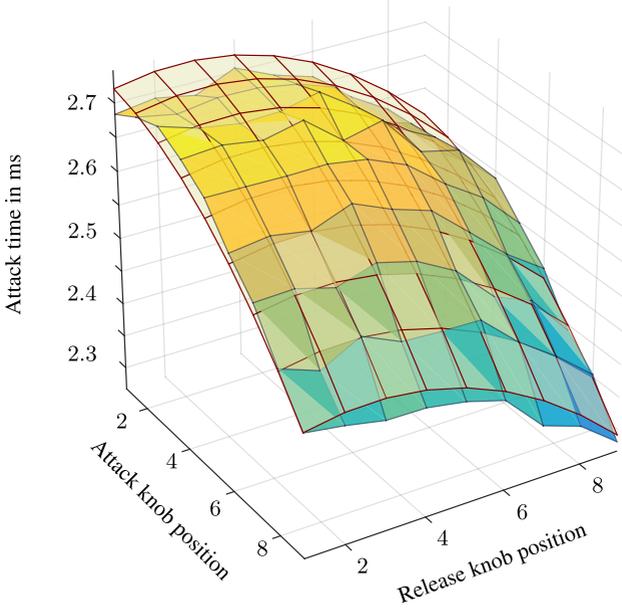
control the 14 model parameters, mentioned in the previous sections. With the obtained model parameters for all user control positions a mapping function can be designed, which relates five user controls to the value of each model parameter.

This relationship can be expressed as a multidimensional polynomial up to second degree

$$p_i(\theta, a_i) = \sum_{j=1}^m a_{i,2j+1} \cdot \theta_j^2 + a_{i,2j} \cdot \theta_j + a_{i,1} \quad , \quad (6.18)$$

where  $p_i(\theta, a_i)$  describes the value of the  $i^{th}$  model parameter,  $\theta_j$  is the angle of the  $j^{th}$  user control knob and  $a_i$  the corresponding polynomial coefficient. The Levenberg–Marquardt method was then used to fit the polynomial coefficients to the measured data, as shown by the mesh in Fig. 6.21.

If all user controls would be analyzed simultaneously in this manner, the amount of combinations would rise uncontrollably. Without simplifi-



**Figure 6.21**

*Relation between user control positions and attack time with optimized second degree mapping polynomial [173].*

cations, the number of possible combinations becomes very high

$$N_e = \prod_{j=1}^{m_c} S_j . \quad (6.19)$$

In this case  $S_j$  represents the number of subdivisions of the  $j^{th}$  user control range and  $m_c$  the total number of user controls. If  $m_c = 5$  user controls are considered and their range is divided into  $S_j = 10$  steps  $N_e = 10^5$ . This leads to  $N_e$  optimization routines. Considering that one optimization takes about one minute to complete, this would result in an unrealistic total model optimization of 2.3 months. But with some simplifications, this amount of time can be drastically reduced.

The ratio buttons can easily be integrated into the digital model, by loading different measured curves for different ratio knobs. The gain controls on the reference device do not influence any other parameter and can be directly mapped to the gains  $G_{PRE}$  and  $G_{POST}$ . The number of inter-dependent user controls has been reduced to  $m_c = 2$  and with  $S_j = 10$

subdivisions of their range, the number of optimizations can be reduced to  $N_e = 100$ , which will amount in roughly 1.67 hours of model optimization.

In this case the optimized digital model is able to emulate the reference device with user controls but this might not be true for a different reference device with more interdependent user controls. This means that there is no unique and robust way to integrate user controls for block-oriented gray box models. The structure of the model has to be considered as well as the interdependency of user controls on the reference device.

## 6.6 Results

In this section the results of the modeling process are presented. Several dynamic range compression systems with different principles of operation have been modeled and the quality of the emulation is rated in this section. Ultimately four devices have been selected:

- UREI – 1176LN (JFET variable resistor)
- Flatline compressor (analog optical isolator)
- MXR – Dynacomp (operational transconductance amplifier)
- Aguilar – TLC compressor (unknown schematic)

After the modeling process, the error between digital model and reference device is expressed with objective metrics. But these metrics do not always correspond to the error perceived by a human listener. This is why a listening test was performed for the digital version of the UREI 1176LN since the model featured the integrated user controls.

### 6.6.1 Objective Metrics

The objective metrics which are used to evaluate the quality of the adapted digital model have been introduced in Chap. 5. Independent of the used error metric the values are very low, as shown by Tab. 6.1.

The similarity score is in almost all cases below 0.1 which indicates that the model output is indistinguishable from the reference signal. Only the MXR – Dyna Comp exhibits a similarity score slightly above 0.1. The reason for this score might be the high noise floor of the Dyna Comp, which is not reproduced by the digital model. The digital model was optimized with a signal consisting of electric guitar and bass inputs which can explain the good scores. The error to signal ratio ( $\epsilon_{ESR}$ ) confirms the good quality of the model. It stays below 0.1 for all inputs from electric guitar and bass.

	Device	Input	Similarity	$e_{\text{RMS}}$	$r_{\hat{y}\hat{y}}$	$e_{\text{ESR}}$
01	Aguilar – TLC	Guitar	0.0417	0.0001	0.9990	0.0021
02	Aguilar – TLC	Bass	0.0818	0.0002	0.9907	0.0087
03	Flatline	Guitar	0.0239	$6 \cdot 10^{-6}$	0.9994	0.0012
04	Flatline	Bass	0.0804	0002	0.9954	0.0092
05	MXR – DynaComp	Guitar	0.1156	0.0002	0.9891	0.0218
06	MXR – DynaComp	Bass	0.0814	0.0001	0.9953	0.0091
07	UREI – 1176LN	Guitar	0.0509	0.0005	0.9975	0.0064
08	UREI – 1176LN	Bass	0.0981	0.0011	0.9914	0.0176
09	UREI – 1176LN	Vocals	0.1464	0.0031	0.9601	0.0790
10	UREI – 1176LN	Drums	0.3463	0.0058	0.8569	0.2660
11	UREI – 1176LN	Piano	0.2457	0.0005	0.9916	0.0429

**Table 6.1**

*Objective scores for evaluation of the optimized digital compressor model.*

When processing signals which were not used during the optimization process, like vocals or piano, the digital model does not perform as good. The sound files which were used in the listening test (see Sec. 6.6.2) are also rated with the objective scores and the digital model does not perform as good as for guitar and bass. An explanation for the relatively bad performance of the model for the drum signal with an  $e_{\text{ESR}} = 0.266$  and a similarity score of 0.3463 might be the setting of the compressor. When processing drum sounds the compressor has to act faster and more intense than for signals which change slower, exhibiting more nonlinear behavior.

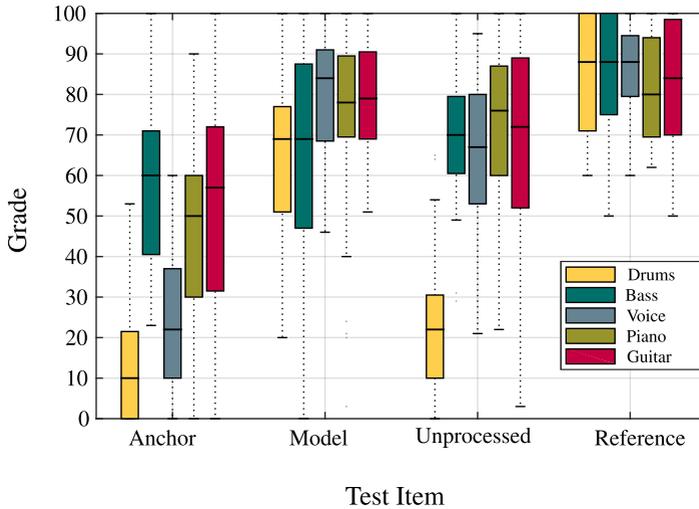
### 6.6.2 UREI 1176LN: Listening Test

To rate the similarity between digital model and analog reference device from a subjective, perceptual point of view a listening test was conducted [175]. The selection of the test-item sound files is not a trivial task. First, the sound files must correspond to real use-cases of a dynamic range compressor. Secondly, they have to be recordings of a real instrument on which an audio engineer would apply dynamic range compression.

The selected sounds should not all have the same pitch, amount of transients and amount of harmonics. A drum sample contains considerably more transients than a guitar signal and a piano exhibits more harmonics than a bass guitar, for example. DRC systems are also used on vocals, especially the UREI 1176LN is often used on singing voices. Human ears are also most sensitive in the frequency region of voices.

Therefore five unprocessed audio samples from drums, electric bass guitar, voice, piano and electric guitar were selected as items in the listening test.

The listening test methodology which was used is called **MUSHRA** [168] as described in 5.3.2.



**Figure 6.22**

*Results of the MUSHRA listening test comparing the optimized digital model with the reference device [175].*

BeaqlJS (browser based evaluation of audio quality and comparative listening environment) is the framework which was used to implement the listening test [178]. It features an ABX and MUSHRA configuration and the results can be transmitted on-line. Overall 63 persons took part in the test but the results of 31 participants have been discarded, because they rated the hidden reference item with a score lower than 50.

The results of this test are shown in Fig. 6.22. The test items have been grouped into the categories ‘Anchor’, ‘Model’, ‘Unprocessed’ and ‘Reference’. The different colors indicate the five sound examples, which were used. The unprocessed items (no compression applied) were rated relatively high which suggests that dynamic range compression is a quite ‘transparent’ effect for an unexperienced listener. Since the test was not conducted in a controlled listening environment some test subjects could have been using low-end audio gear, which is reflected in the scores for the hidden reference which were rated between 80 and 90 for all test items. But the output of the digital model was, on average, always rated better than the ‘Unprocessed’ and ‘Anchor’ items. The results of this listening test with small differences between the test items depends a lot on the trained ears of the test subjects. Although nearly half of the test subjects have been removed from the evaluation, the results of this test can still contain uncertainty. Nevertheless, in general the digital model is rated relatively

close to the hidden reference, which confirms the good performance of the digital model.

Especially vocals, piano and guitar are rated nearly as good as the reference signal which is another indication for the performance of the adjusted model. One particularity is the discrepancy between the similarity scores and the result of the listening test for the vocals and the bass guitar signal. The similarity score for the bass item is 0.0981 and the vocal item has a score of 0.1464. In the listening test the rating for the vocal item is better than for the bass guitar. This could be explained with the experience level of the test subjects. It might be easier for an untrained listener to hear differences for the bass guitar item than for the vocal item.

Another reason for the audible difference between the unprocessed bass and drum items and the compressed ones in comparison to the vocal item are the settings of the reference device. It is set up to compress these items more than guitar or vocals exhibiting a more nonlinear behavior which makes it more difficult to adapt the digital model with a gray-box approach.

## 6.7 Conclusion

An automated procedure for dynamic range compressor modeling was presented in this chapter. A generic model has been used to recreate the behavior of an analog reference device. The quality of the model has been rated for several devices in different settings with objective scores. A listening test has been conducted for one specific reference device.

The objective scores and the results of the listening test show that the adapted digital model performs quite good. For the intended use-case of electric guitars the objective scores showed good results. In the listening test, the digital model was not rated worse than 'minor differences' on average, while the anchor was rated with 'annoying differences', 'major differences' and 'differences'. A real-world input signal is very important for the optimization of the models parameters. The objective scores showed that the model performs quite well for signals which are similar to the optimization signal (electric guitar and bass) but does not perform as well for different signals (vocals and piano) which suggests that these results could be improved by designing a more versatile input signal for the optimization.

It is possible to integrate the user controls of a specific reference device by creating a digital model for every possible combination of user parameters and then interpolating between the models parameters according to the user settings. But this process scales exponentially with the number of user parameters. Making simplifications is advisable e.g. assuming that a volume knob just scales the output signal.

---

## Basic Distortion Circuits

---

In this chapter, the fundamental building blocks of distortion circuits will be analyzed and modeled. Three different circuits, employing three different concepts of creating harmonic distortion, are analyzed in detail. A block-oriented Wiener model has been chosen to recreate each circuit's characteristics. The analyzed circuits are a simple diode clipper, a distortion stage from the Electro Harmonix – Big Muff and the basic circuit of the Ibanez – Tube Screamer.

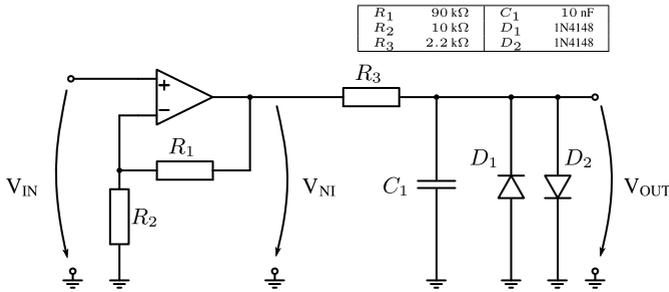
### 7.1 Circuit Analysis

All circuits are simple building blocks of larger systems. The simplest circuit is the diode clipper, where only the anti-parallel diodes introduce distortion. The other two circuits introduce more nonlinear elements as well as a slightly more complex topology. Their schematics will be presented and analyzed in detail.

All circuits are simulated with 'LTSpice', which is a circuit simulation software, modeling the physical parameters of each circuit element. A transistor is for instance described by e.g. the width and depth of the doped areas or its channel width. Real measurements are not performed in this study.

#### 7.1.1 Diode Clipper

The first analyzed circuit is the diode clipper whose schematic is shown in Fig. 7.1.

**Figure 7.1**

*Schematic of the diode clipper with pre-amplification. The distortion is introduced by the two anti-parallel diodes.*

The input signal is amplified by an ideal non inverting amplifier with a gain of

$$V_{NI} = \left(1 + \frac{R_1}{R_2}\right) \cdot V_{IN} = 10 \cdot V_{IN} \text{ ,}$$

afterwards it is filtered by a first-order low-pass filter with cut-off frequency

$$f_c = \frac{1}{2\pi R_3 C_1} \approx 7.23 \text{ kHz} \text{ ,}$$

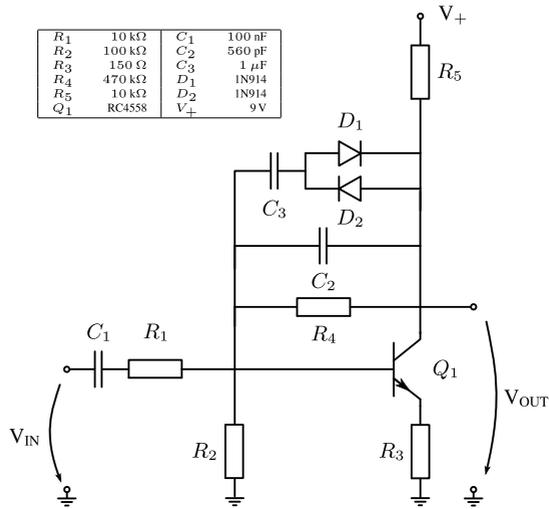
and the last stage consists of the two anti-parallel diodes which create the distortion by clipping the amplified and filtered waveform.

### 7.1.2 Big Muff – Distortion Stage

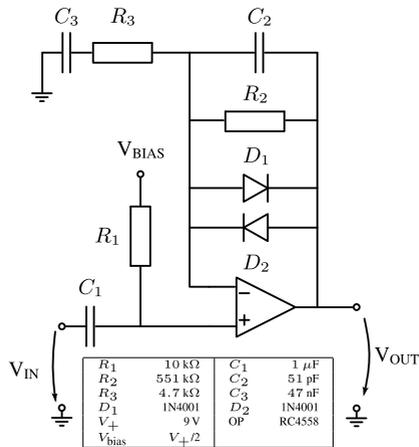
The distortion stage of the Electro Harmonix – Big Muff is an extended BJT transimpedance stage, which was analyzed in [102]. The schematic of the Big Muff stage is shown in Fig. 7.2. The distortion is introduced by the anti-parallel diodes in the feedback path and for high signal levels ( $V_{in} > 1.5 \text{ V}$ ) by the NPN transistor itself.  $C_2$  and  $R_4$  are used to shape the frequency response of the circuit.

### 7.1.3 Tube Screamer

The distortion stage of the Tube Screamer is based on an operational amplifier gain stage, which was also analyzed in [53, 102]. The input signal is amplified and distorted by the two anti-parallel diodes in the feedback path from output to negative input of the operational amplifier. The operational amplifier introduces distortion as well. Additionally  $R_2$ ,  $R_3$ ,  $C_2$  and  $C_3$  form a band-pass filter, which is parallel to the nonlinear elements. In Fig. 7.3 the circuit is depicted. Resistor  $R_2$  is the ‘drive’ potentiometer

**Figure 7.2**

Schematic of the Big Muff distortion stage [179].

**Figure 7.3**

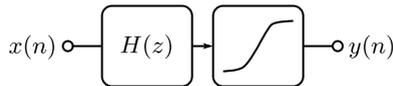
Schematic of the Tube Screamer distortion stage [180].

in the original circuit. It takes values from  $R_{2,min} \approx 51$  k $\Omega$  to  $R_{2,max} \approx 551$  k $\Omega$ . In this work the potentiometer is always set to  $R_{2,max}$  to maximize the amplification of the operational amplifier and thus the distortion of the

output signal. The bias voltage is set to  $V_{\text{BIAS}} = 4.5 \text{ V}$ , which is half the supply voltage of the operational amplifier. The used op-amp is a general-purpose amplifier RC4558 by Texas Instruments.

## 7.2 Digital Model

The digital model chosen to represent these circuits is an extended Wiener model. It consists of a series connection of a linear time invariant block and a mapping function as the nonlinear block. Its block diagram is shown in Fig. 7.4.



**Figure 7.4**

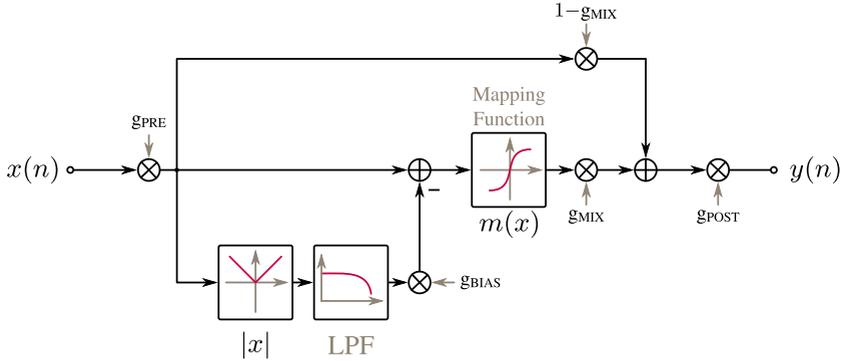
*Block diagram of the extended Wiener model.*

The Wiener model is used because none of the circuits has a filter at the output and in case of the diode clipper it exactly recreates the reference circuits topology. Furthermore, it should be ascertained how well such a simple model can recreate the circuits of the Big Muff and the Tube Screamer, which do not exhibit the exact same topology as the model.

This model is not a typical Wiener model, because the nonlinearity can not be considered completely memoryless. The mapping function, which is the main part of the nonlinear block, has been extended. The nonlinear block contains a low-pass filter, which introduces memory. The usual Wiener model only features memoryless nonlinear blocks.

The signal flow graph of the nonlinear block is depicted in Fig. 7.5. After the input of the nonlinear block  $x(n)$  is multiplied with the parameter  $g_{\text{PRE}}$ , the signal path splits up in dry path, wet path, and side-chain. In the side-chain path the absolute value of the input signal is calculated and low-pass filtered. The low-pass filter has a cut-off frequency of  $f_c = 5 \text{ Hz}$ , which makes the side-chain a simple envelope detector. The output of the envelope detector is weighted with the parameter  $g_{\text{BIAS}}$  and then subtracted from the wet path just before the mapping function, which introduces the distortion, is applied. This side-chain envelope detector simulates a signal dependent bias-point shift which can be found in vacuum tube amplifiers for electric guitars. A comparable signal processing technique has been proposed in [181] to achieve a similar effect like a varying cathode voltage which alters the plate current and thus moves the bias point of a vacuum tube.

After the distortion is introduced by the mapping function  $m(x)$ , dry



**Figure 7.5**

Signal flow graph of the nonlinear block.  $m(x)$  is the mapping function. The ‘wet’ signal path is weighted with  $g_{\text{MIX}}$  and the ‘dry’ signal path with  $(1-g_{\text{MIX}})$ .

and wet signal path are mixed together. The parameter  $g_{\text{MIX}} \in [0, 1]$  is used to determine the mixture ratio. Finally the signal is weighted with the post-gain  $g_{\text{POST}}$  which yields the output  $y(n)$ .

The mapping function which is used in this work is the piecewise-defined hyperbolic tangent, which is described in detail in Sec. 3.2.1 and uses four parameters. Together with the parameters from the extensions, the nonlinear block has a total amount of eight parameters, which are combined in the parameter vector

$$\mathbf{p}_{\text{nl}} = (k_p \quad k_n \quad g_p \quad g_n \quad g_{\text{PRE}} \quad g_{\text{POST}} \quad g_{\text{MIX}} \quad g_{\text{BIAS}})^{\text{T}}.$$

## 7.3 Modeling

To ensure a robust modeling procedure, the process is divided into several steps. First, the linear part of each circuit is measured with an exponentially swept sine. The amplitude of the sweep is kept low enough that no harmonic distortions are created by the nonlinear circuit elements. Afterwards, the impulse response is extracted as described in Sec. 4.4 and directly used as an FIR filter to form the first block of the Wiener model.

When designing the input signal it is important to consider the influence of the parameters on the output. The nonlinear block in the extended Wiener model is frequency independent, which means that, for the purpose of parameter optimization, it is not necessary for the input signal to excite more than one frequency. But it is most important to excite all possible amplitudes of the input signal, so their modification by the reference

system can be observed. A single frequency sine wave with logarithmically decreasing amplitude was used as the input signal

$$x_{nl}(n) = a(n) \cdot \sin\left(2\pi \frac{f_0}{f_s} \cdot n\right). \quad (7.1)$$

The fundamental frequency was set to  $f_0 = 1$  kHz. The amplitude scaling function  $a(n)$  is logarithmically decreasing from a start value of  $a(0) = 1$  to the lowest value  $a(N - 1) = 1 \cdot 10^{-5}$ , with  $N$  as the total amount of samples. The sampling frequency is  $f_s = 44.1$  kHz.

Because the Levenberg–Marquardt method is gradient-based, it is highly unlikely to find initial parameters that will ensure convergence into a global optimum of the cost function, as mentioned in Sec. 4.2.6. Therefore, the parameters of the nonlinear block are optimized in three steps and with different input signals and cost-functions.

The choice of the cost function is crucial for the robustness of the optimization process. The straight-forward approach would be to calculate the difference of digital model output and analog reference output in time-domain. But if the phase characteristic of reference and model is not matched perfectly, the time-domain error is quite high, which does not necessarily represent the human perception of the difference between the two signals. To neglect any phase shift between reference system and digital model, the cost function is designed to match the envelopes of both systems with the straightforward approach mentioned in Sec. 4.3.2. The envelopes are calculated for positive and negative signal amplitudes separately, because the nonlinear mapping function is able to shape positive and negative amplitudes independently.

To calculate the envelope, the signals are half-wave rectified and low-pass filtered by a second order IIR low-pass with a cut-off frequency of  $f_c = 5$  Hz. For negative amplitudes the signals are multiplied with  $-1$  before half-wave rectification. In the final optimization step, the parameters are refined for

Step #	Input Signal	Cost-Function	Parameters
1	$x_{nl}(n)$	Envelope: pos. amplitudes	$k_p, g_p, g_{PRE}, g_{POST}$
2	$x_{nl}(n)$	Envelope: neg. amplitudes	$k_n, g_n, g_{PRE}, g_{POST}$
3	$x_{real}(n)$	Magnitude spectrogram	all

**Table 7.1**

*Overview of the different parameter optimization steps to ensure a robust identification.*

a real-world signal with the cost-function described in Sec. 4.3.3. The input signal  $x_{real}(n)$  is a dry recording of different electric guitars and an electric bass. This step should not be omitted, because the synthetic sine

wave from the previous optimization steps does not offer realistic signal conditions in terms of frequencies and transients.

Finally a summary of all optimization steps with their corresponding cost-functions is given in Tab. 7.1.

## 7.4 Results

The results of the modeling process are shown in Tab. 7.2. The diode clipper obtained the best results with an error to signal ratio of  $e_{\text{ESR}} = 5.78\%$ , a correlation coefficient of  $r_{y\hat{y}} = 0.9983$  and a similarity score of  $\text{PEAS} = 0.0727$ .

The results for the other circuits are not as good. The Big Muff BJT gain stage has an  $e_{\text{ESR}} = 9.01\%$ , a correlation coefficient of  $r_{y\hat{y}} = 0.9578$  and a similarity score of  $\text{PEAS} = 0.3425$  which already leads to a perceptible difference between the signals. This can be explained by the feedback path from the collector of the NPN transistor to its base (see Fig. 7.2). This feedback path is not modeled in the extended Wiener model, so the result of the modeling process is only an approximation of the real circuit.

Circuit	$e_{\text{ESR}}$	$r_{y\hat{y}}$	PEAS
Diode Clipper	0.0578	0.9983	0.0727
Big Muff	0.0901	0.9578	0.3425
Tube Screamer	0.1832	0.9062	0.3231

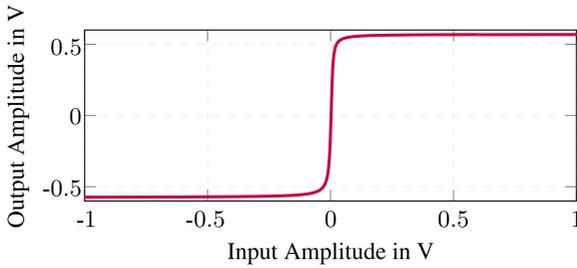
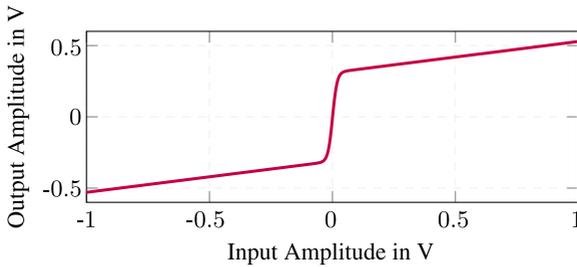
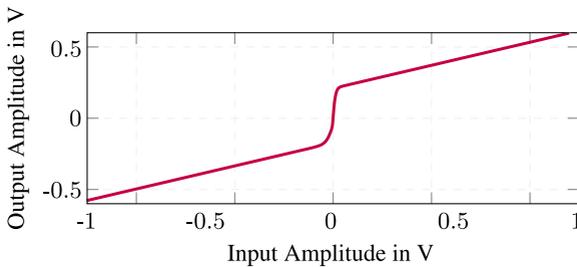
**Table 7.2**

*Results of the modeling process.*

The Tube Screamer has an  $e_{\text{ESR}} = 18.32\%$ , a correlation coefficient of  $r_{y\hat{y}} = 0.9062$  and a similarity score of  $\text{PEAS} = 0.3231$ , which also leads to a noticeable difference between digital model and circuit output. This difference can also be explained by the simplicity of the extended Wiener model. In the original circuit,  $R_2$ ,  $R_3$ ,  $C_2$  and  $C_3$  form a band-pass filter, which is parallel to the nonlinear elements. To achieve better results this filter should also be placed parallel to the nonlinear block. But its frequency response is unknown, because only the global frequency response can be measured, without detailed measurements of the analog circuit, which is not the aim of this study.

To see how the digital model adapts to the different circuits, the output of the nonlinear block

$$y = g_{\text{MIX}} \cdot m(x) + (1 - g_{\text{MIX}}) \cdot x$$

(a) *Diode Clipper*(b) *Big Muff*(c) *Tube Screamer***Figure 7.6**

Output of the nonlinear block  $y = g_{MIX} \cdot m(x) + (1 - g_{MIX}) \cdot x$  after optimization of the parameters.

is shown in Fig. 7.6 for all adapted digital models. Please note that the mapping function is shown **after** the optimization is finished and the model is adapted.

The mapping function of the diode clipper shows, that no dry signal is mixed with the distorted signal, because the slope of  $m(x)$  is 0 at the

amplitude maxima for  $x(n) = 1$  and  $x(n) = -1$ . This corresponds well with the circuit, because it has no filters parallel to the distortion.

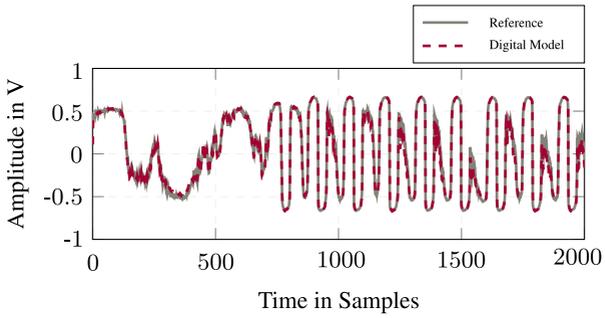
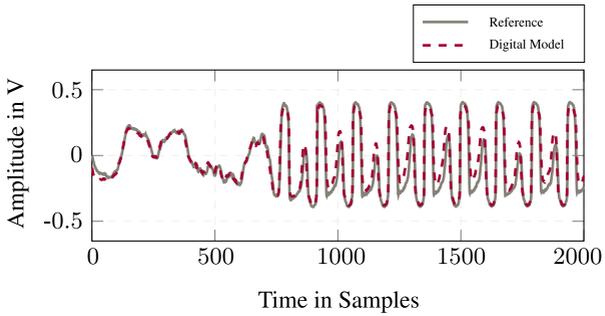
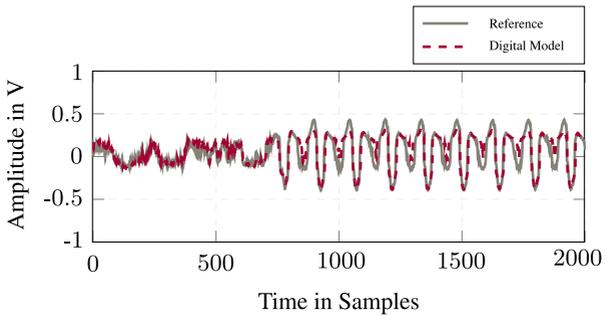
The Big Muff circuit has a little dry signal mixed together with the wet signal, which can be explained as a ‘compensation’ of the missing feedback path in the model. The mapping function for the Tube Screamer exhibits even more dry signal mixed together with the distorted wet signal, which confirms the assumption that the circuit can be modeled with a parallel dry signal, which can be mixed with the distorted signal, even though there is no filter in the dry path of the digital model.

All mapping functions look symmetrical, which is due to using the same model for the two diodes in each Spice simulation, leading to the same shape for positive and negative amplitudes. Only in the Tube Screamer circuit there is a slight difference between the shape of positive amplitudes and negative amplitudes, which is also visible in the mapping function, Fig. 7.6c, because the transition from steep middle part of the mapping function to higher amplitudes is a little softer for negative amplitudes.

The time-domain signals for each reference circuit and their comparison to the corresponding model are shown in Fig. 7.7. The input signal was a self-recorded riff played on a stratocaster-type electric guitar using the humbucker bridge-pickup. The guitar was directly connected to an RME – Fireface 800 audio interface. For the diode clipper, the digital model waveform is nearly identical to the waveform of the reference signal, which leads to no perceivable difference between the two signals. With a rising  $\epsilon_{\text{ESR}}$  and PEAS value for Big Muff and Tube Screamer the waveform of the digital model differs more and more from the reference output. Generally, it can be observed that the difference in the waveforms is proportional to the input amplitude, because for the first 500 samples of the test signal, all models are close to the reference signal, while for higher input amplitudes (sample 600 to 2000) the model is not accurate enough to recreate the more complex reference circuits perfectly.

In addition to these scores, an informal listening test was conducted. The participants of the test were five experienced researchers in virtual analog modeling, who should test if they are able to hear a difference between digital model and reference signal. In case of the diode clipper none of the participants was able to hear a difference between simulation and reference. For Big Muff and Tube Screamer, the results were not as convincing, since every test subject was able to hear a difference but no quantitative analysis was performed. Nevertheless all of the participants confirmed that the overall characteristic of the reference device could be captured by the corresponding digital version.

One particularity in the objective scores in relation to the results of the informal listening test can be pointed out. All participants confirmed that the differences for the Big Muff circuit were easier to detect, while the Tube Screamer performed better. This is not reflected by the error to

(a) *Diode Clipper*(b) *Big Muff*(c) *Tube Screamer***Figure 7.7**

*Time-domain response to a recorded guitar signal for all circuits. The same part of the waveform is shown in all plots.*

signal ratio ( $e_{\text{ESR}}$ ) or the visual comparison of the time-domain waveform shown in Fig. 7.7 but the similarity score (PEAS) shows a lower value for the Tube Screamer than for the Big Muff, which confirms that the similarity measure is better suited to rate the quality of a virtual analog model than the straight forward time-domain error.

## 7.5 Conclusion

Three distortion circuits have been modeled with an extended Wiener model consisting of a linear time-invariant block and a nonlinear block. The results for the simple diode clipper circuit are very good. The waveform of the reference device is nearly perfectly recreated because the digital model has the same topology as the circuit. The good performance of the digital model is also reflected in the similarity score.

The nonlinear mapping function proves to be very versatile and is able to adapt to various nonlinear circuit elements using only a few parameters. The extensions that have been made to the mapping function, like pre- and post-gains and the dry/wet mixing, are used to make the nonlinear block even more flexible.

The results for more complex circuits, which do not have the same topology as the digital model, are not perfect but the digital model is able to recreate their characteristic sound to a satisfying degree given the simple digital model. The model should be expanded to be able to capture the characteristics of more complex distortion circuits.



---

## Distortion Guitar Effect Pedals

---

In this chapter the focus will shift from parts of distortion circuits to whole systems which are used to distort an electric guitar signal. By using guitar effect pedals, also called ‘stompboxes’, a guitarist is able to create a distorted sound by stepping on the switch of the device. In this chapter these devices are modeled with a fully parametric Wiener–Hammerstein model utilizing a multi-step optimization routine.

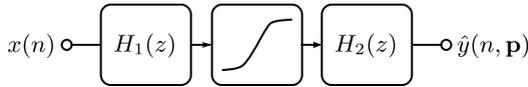
A multitude of distortion guitar effect pedals are available as commercial products. Some of them enjoy broad popular appeal. A few well known distortion pedals will be analyzed and modeled in this study.

The aim of black- or gray-box modeling is to recreate the reference system without complete knowledge about the schematic and every circuit element. This is the reason the circuits of the devices are not analyzed in detail in this chapter.

### 8.1 Digital Model

The basic idea behind the model used in this study is to have a completely parametric model, which is flexible enough to recreate many distortion effects but still simple enough to be computationally efficient during runtime and optimization.

The structure of a distortion effect can be described by a Wiener–Hammerstein model. This model consists of linear-time-invariant (LTI) blocks and a nonlinear block. The blocks are ordered in series where the nonlinear block is lined by two LTI blocks, see Fig. 8.1. The LTI blocks are



**Figure 8.1**

*Block diagram of a digital Wiener Hammerstein model.*

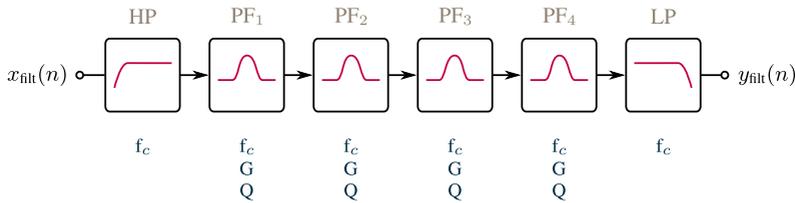
filters and the nonlinear block is the extended mapping function, which was already discussed in the previous chapter (see Sec. 7.2). The signals  $x(n)$  and  $y(n)$  denote the input and output. The incoming signal  $x(n)$  is filtered by  $H_1(z)$  which emphasizes or attenuates certain frequency regions. Usually a high-pass filter is used in this stage, to block out the direct component and because distortion of low frequencies results in a crowded spectrum because the harmonics are relatively close together. The output of the first filter is then processed by the nonlinear block which adds the distortion. The output of the nonlinear block is then filtered again to shape the overtones produced by the mapping function and leads to the characteristic sound of the distortion.

Block-oriented Wiener-Hammerstein models are successfully used in commercial products due to their flexibility and expandability. Fractal Audio Systems calls a Wiener-Hammerstein system consisting of two filters and a distorting nonlinearity the ‘fundamental paradigm of electric guitar tone’ [129]. They extend this model with more linear and nonlinear blocks to include frequency responses of cabinets and model the different nonlinear stages (pre-amp, power-amp) of an analog guitar amplifier.

### 8.1.1 Linear Blocks

The filters of the Wiener-Hammerstein model are designed to be flexible but still stay stable in the identification process. Hence, the parameters which will be varied during optimization are not the coefficients of a Direct Form filter structure, because the identification procedure will not converge if the filter coefficients yield an unstable filter during optimization. Instead, they are expressed as the parameters of a simple band-limited equalizer.

Fig. 8.2 shows the structure of one LTI block of the Wiener-Hammerstein model. The input signal  $x_{\text{filt}}(n)$  is processed by a high-pass filter, a series of four peak filters and finally by a low-pass filter which yields the output signal  $y_{\text{filt}}(n)$ . The adjustable parameters of the filters are expressed in terms of cutoff frequency  $f_c$  for the low-pass and high-pass filters. The peak filters can be modified in terms of gain  $G$ , Q factor and center frequency  $f_c$ . All filters are second order IIR filters and their coefficients are computed according to [98].

**Figure 8.2**

Block diagram of the parametric filter block used in the Wiener-Hammerstein model.

All parameters are aligned in a filter parameter vector

$$\mathbf{p}_{\text{LTI}} = \begin{pmatrix} f_{c,\text{hp}} \\ f_{c,\text{pf1}} \\ G_{\text{pf1}} \\ Q_{\text{pf1}} \\ f_{c,\text{pf2}} \\ G_{\text{pf2}} \\ Q_{\text{pf2}} \\ \vdots \\ f_{c,\text{lp}} \end{pmatrix}. \quad (8.1)$$

## 8.2 Modeling

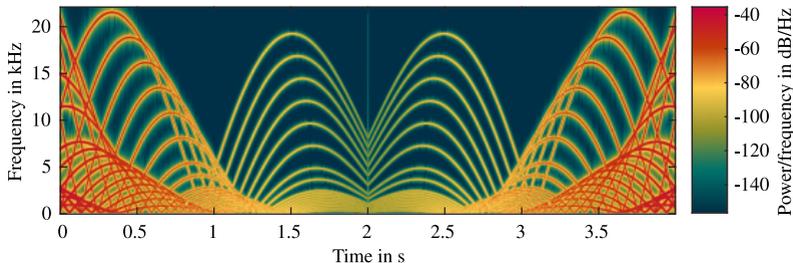
Before optimizing the models parameters the initial parameters have to be found, as it is crucial for a successful optimization with convergence to the global minimum of the cost-function. A reliable way to find good initial parameters are grid searches but the effort of performing a grid search is exponentially increasing with the grid resolution and the number of parameters. This is why only the parameters of the mapping function  $m(x)$  and the pre-gain  $g_{\text{PRE}}$  of the nonlinear block are tested in the grid-search. These parameters have the strongest influence on the output signal of the digital model. The result of the cost-function  $C(\mathbf{p})$  is calculated for every possible combination of the five parameters on a coarse grid and the one with the lowest value of  $C(\mathbf{p})$  are chosen as the initial parameters. The parameters of the filters are not changed during the grid search and they are configured to yield a neutral frequency response in the audible frequency range.

### 8.2.1 Optimization of Nonlinear Parameters

The optimization itself is carried out according to Fig. 4.1 (see Sec. 4.1). It is performed in several steps to ensure a reliable result. The first step of the optimization procedure uses a multi-frequency sine wave with amplitude weighting as the input signal. The signal is generated using Eq. (8.2), where  $a(n)$  is an amplitude weighting sequence, with logarithmically decreasing amplitudes from  $a(0) = 0$  dB to  $a(N - 1) = -60$  dB

$$x_{ms}(n) = a(n) \cdot \sum_{m=0}^{M-1} \sin \left( 2\pi \cdot \frac{f_m + f_m \cdot \sin \left( 2\pi \frac{f_{\text{mod},m}}{f_s} \cdot n \right)}{f_s} \cdot n \right) . \quad (8.2)$$

After the signal has been created it is concatenated with a modified version of itself, which is  $-x_{ms}(N - n - 1)$  to see how the reference system reacts to rising signal amplitudes. The fundamental frequencies for each sine wave are denoted by  $f_m$ . The lowest fundamental frequency is  $f_0 = 100$  Hz up to the highest frequency, which is  $f_{M-1} = 8$  kHz. These frequencies are modulated by adding a weighted version of the fundamental frequency to itself. The weight itself is another sine wave with frequency  $f_{\text{mod},m}$ , which contains linearly spaced frequencies between  $f_{\text{mod},0} = 0.8$  Hz and  $f_{\text{mod},M-1} = 0.35$  Hz. The magnitude spectrogram of the signal is shown in Fig. 8.3.



**Figure 8.3**

*Time-frequency representation of the multi-frequency sine input  $x_{ms}(n)$ .*

Please note that the signal  $x_{ms}(n)$  is normalized to exhibit a maximum amplitude of one. The aim of this design is to cover a lot of relevant frequencies and amplitudes for an electric guitar in one signal.

A grid search is performed for some parameters of the mapping function. The envelope of model and reference output is calculated with the straightforward approach described in Sec. 4.3.2. All combinations of the parameters of the mapping function  $k_p$ ,  $k_n$ ,  $g_p$ ,  $g_n$  and the pre-gain are compared on a coarse grid and the combination yielding the lowest error is chosen as initial parameters for the first optimization step. These pa-

parameters are chosen for the grid-search because they directly influence the envelope of the output signal. The input signal  $x_{\text{ms}}(n)$  is chosen because all relevant amplitudes are covered.

The iterative optimization is started and only the parameters of the non-linear block are adapted. The parameters of input and output filter are not changed and they are configured to exhibit a flat frequency response in the audible spectrum. The cost-function in this first optimization step is the spectrogram-based cost-function described in Sec. 4.3 and all calculations are performed with  $x_{\text{ms}}(n)$  as the input signal.

## 8.2.2 Optimization of Filter Parameters

The input signal for the parameter optimization of the LTI blocks is white noise with signal levels below -50 dB, because we assume, that for low signal levels the nonlinear part of the reference system operates in its linear region and  $\tanh(x) \approx x$  for  $|x| \ll 1$  is also true for the nonlinear block of the digital model.

In this case optimization of time domain error signals is challenging because the signal can look different in time domain, due to deviating phase characteristics of simulation and reference system. The signal can still be perceived as similar for a human listener, however. For this reason the output signals  $y(n)$  and  $\hat{y}(\mathbf{p}, n)$  need special treatment before the actual minimization procedure can be started.

The time domain output sequence for the white noise identification input is recorded. The power spectral density (PSD) of the output is computed by calculating a 16384-point discrete Fourier transform (DFT) with a hop size of 4096 samples. All calculated spectra are averaged and multiplied by its complex conjugate to yield the PSD. Afterwards the bins of the Fourier transform are averaged to yield a semi-tone spectrum as described in Sec. 4.3.3.

The initial values for the identification procedure are chosen in such a way that the filter is flat and the cutoff frequencies of high- and low-pass filters are set to  $f_{\text{c,HP}} = 10$  Hz and  $f_{\text{c,LP}} = 18$  kHz. After the filter parameters are adapted, they are stored for further use.

## 8.2.3 Optimization of All Parameters

In this final step the stored parameter vectors for both LTI blocks and the nonlinear block are loaded and used as the initial parameters in the final parameter vector

$$\mathbf{p} = \begin{pmatrix} \mathbf{p}_{\text{LTI1}} \\ \mathbf{p}_{\text{LTI2}} \\ \mathbf{p}_{\text{NL}} \end{pmatrix}. \quad (8.3)$$

The Levenberg-Marquardt algorithm is started and now all parameters of the model can be modified to refine the results from the previous optimization runs. The input signal for this step is a dry guitar signal, composed of different guitar types and playing styles. The cost-function in this optimization step is the time-frequency spectrogram cost-function described in Sec. 4.3.3.

### 8.3 Results

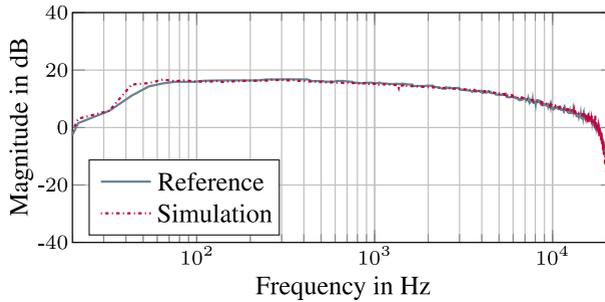
The optimization process has been tested thoroughly on different distortion effect pedals in various settings. The aim is to find effect pedals which create different kinds of distortions with miscellaneous nonlinear circuit elements, to verify how reliable the optimization process is.

The ‘Electro Harmonix – Big Muff’ [179] consists mainly of two BJT gain stages, which were already discussed in the previous chapter. A DIY clone of the ‘Jim Dunlop – Fuzz Face’ is a simple two-stage amplifier with a feedback network between the two stages, which controls gain and frequency response [182]. The amplification is achieved with BJT PNP transistors. The ‘Hughes & Kettner – Tube Factor’ [183] is a tube-based pre-amplifier as a guitar effect pedal with a 12AX7 vacuum tube. It has two modes of operation called ‘Factor 1’ and ‘Factor 2’, where ‘Factor 2’ exhibits higher amplification before the nonlinear circuit element (12AX7 tube) which leads to more distortion. And finally a clone of the Ibanez – Tube Screamer [184], which was also analyzed in the previous chapter.

The frequency response of the ‘Fuzz Face’ is shown in Fig. 8.4. The parametric filter adapts quite well to the noise PSD, because in the important frequency range the error between both frequency responses is always lower than 1 dB. Larger error values can only be found for frequencies below 60 Hz and above 18 kHz. In the case of the ‘Fuzz Face’ these error values did not result in any unpleasant auditory impression.

A time domain comparison is shown for the ‘Tube Factor’ in Fig. 8.5a. The response to a 440 Hz sine wave with exponentially decreasing amplitude from 1 V to 0 V is shown. The digital model follows the analog reference system very well except for the transient part at the start of the signal ( $t = 0.5$  s). For the highest amplitudes the peaks of the reference waveform seem to exhibit more high-frequency content, since they do not look as rounded as the peaks for the digital simulation.

In Fig. 8.5b the response of the ‘Tube Factor’ to a real-world guitar signal is shown. The digital emulation recreates the reference system well, but not perfectly. Please note, that these results are only small excerpts from the range of time-domain signals and frequency responses which could be shown and only give an impression of the performance of the digital model. Therefore, the objective scores, presented in Chap. 5, are used to



**Figure 8.4**

*Frequency response comparison of the ‘Jim Dunlop – Fuzz Face’ and its digital version.*

rate the quality of the parametric Wiener-Hammerstein model.

The scores are presented in Tab. 8.1. They are calculated for the same input signal, which is composed of several recorded guitars in different playing styles and a decaying tone played on an electric bass guitar. The output of the reference device is measured and compared to the output of the digital model according to the objective metrics: the similarity score (PEAS), the error to signal ratio ( $e_{\text{ESR}}$ ), the root mean square error ( $e_{\text{RMS}}$ ), and the Pearson correlation coefficient  $r_{y\hat{y}}$ .

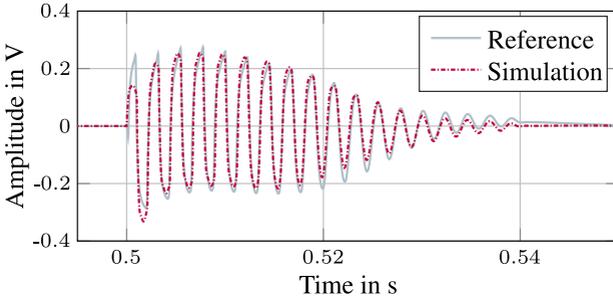
DUT	PEAS	$e_{\text{ESR}}$	$e_{\text{RMS}}$	$r_{y\hat{y}}$
Big Muff	0.1204	0.0125	0.0004	0.9937
Fuzz Face	0.1056	0.0292	0.0004	0.9854
Tube Factor (Factor 1)	0.1714	0.0368	0.0035	0.9816
Tube Factor (Factor 2)	0.2863	0.1166	0.0079	0.9417
Tube Screamer	0.2311	0.1194	0.0053	0.9403

**Table 8.1**

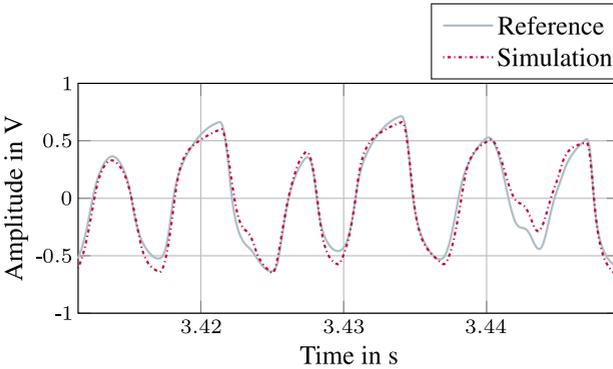
*Results of the optimization for different devices under test (DUTs). The ‘Big Muff’, ‘Fuzz Face’, ‘Tube Factor’ and ‘Tube Screamer’.*

The RMS error is very low for all reference devices, for instance the values for the Big Muff and the Fuzz Face are the same ( $e_{\text{RMS}} = 0.0004$ ), while all other metrics show a difference for both emulations. This confirms the assumption that this metric is not very well suited to rate the quality of the emulation.

The error to signal ratio is based on the root mean square error. It relates



(a) Tube Factor – 40 ms, 440 Hz sine with exponentially decaying amplitude.



(b) Tube Factor – Excerpt from recorded guitar signal.

**Figure 8.5**

*Time-domain comparison of digital model and analog reference device for the Hughes & Kettner – Tube Factor.*

the energy of the error signal to the energy of the reference signal. It shows a difference between Big Muff and Fuzz Face but drastically increases for Tube Screamer ( $e_{ESR} = 11.66\%$ ) and ‘Factor 2’ of the Tube Factor ( $e_{ESR} = 11.94\%$ ) but the quality of the emulation is not as bad as these scores would suggest.

This is verified by the similarity score (PEAS). The values for Tube Screamer and ‘Factor 2’ are still higher than for the other emulations but comparable to the results from the last chapter. A trained listener would be able to hear differences between the signals but would not rate them as annoying.

The general trend for all emulations is that for higher distortion the

quality of the emulation is impaired because the system becomes more nonlinear than before which makes the reference device harder to measure and identify. The comparison between ‘Factor 1’ and ‘Factor 2’ of the Tube Factor also shows this trend. The reference device had the same settings for both modes, but in the ‘Factor 2’ mode the signal is amplified by 2 (in comparison to ‘Factor 1’) before passing through the nonlinear circuit elements which leads to more distortion and hence more nonlinear behavior of the reference device.

## 8.4 Conclusion

Different analog distortion pedals for electric guitars have been modeled with a completely parametric Wiener–Hammerstein model which consists of a series connection of a linear time-invariant block, a nonlinear block and another linear time-invariant block. The model is not able to recreate every device perfectly but the results were satisfactory. The digital model is suited to recreate simple analog distortion systems.

Even though the parametric filters can be adapted with the proposed approach it would be preferable to measure the filters directly. This would drastically reduce the dimensionality of the optimization problem, which would make the procedure more robust and reliable. The nonlinear block proved to be a valuable part of the model because of its flexibility using only a few parameters.

The evaluation shows that the results of the modeling process get measurably worse when the reference system exhibits more nonlinear behavior.



---

## Guitar Amplifiers

---

Guitar amplifiers have been the subject of virtual analog modeling for years and there already exist methods to recreate the analog circuit with digital means. The focus of this chapter lies on automatically adjusting a digital model to recreate the reference amplifier.

Two block-oriented models are analyzed in this chapter. The first one is a Wiener–Hammerstein model which was also used to recreate distortion guitar effect pedals (see Chap. 8). It contains only one nonlinear block and two filters. The second one is a more complex cascaded model which resembles the internal structure of a guitar amplifier. It contains two nonlinear blocks and three filter blocks.

The optimization procedure for both models is described in this chapter. The optimized models are evaluated with objective scores as in the previous chapters and additional listening tests are performed to rate the subjective quality of each model. Objective and subjective results are compared and conclusions are drawn.

### 9.1 Digital Models

The following section describes the topology of the digital models. Each model is a series connection of nonlinear blocks and filters. First, the simpler Wiener–Hammerstein model is described in the context of guitar amplifier modeling and afterwards the more complex cascaded model is described.

### 9.1.1 Wiener – Hammerstein Model

The digital model used in this work is an extended Wiener–Hammerstein model and describes a distortion system (see Fig. 9.1) like e.g. a guitar amplifier. The fundamental principle of guitar amplifiers are pre- and post filters sandwiching a distorting element like a vacuum-tube, for example. The extended Wiener–Hammerstein model recreates this structure.



**Figure 9.1**

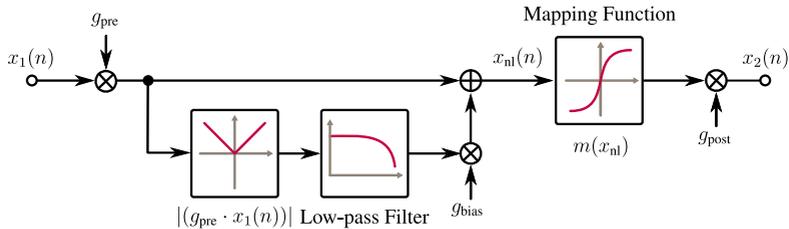
*Digital model describing the structure of a generic digital distortion system. The input signal is filtered with the input filter  $H_1(z)$ , distorted by the nonlinear block and the distorted signal is again filtered by the output filter  $H_2(z)$ .*

The input signal  $x(n)$  is filtered by the input filter  $H_1(z)$  resulting in the signal  $x_1(n)$ . This determines the amount of distortion per frequency. If the input filter has a lot of attenuation for a certain frequency region, these frequencies will not produce as many harmonics since the nonlinear mapping function will not be driven into saturation as much. The input filter also defines the amount of intermodulation distortion which is created when the input signal features more than one tonal component.

The nonlinear block, shown in Fig. 9.2, features a pre- and a post-gain to scale input and output signal respectively. The concept of the side-chain envelope detector has been proposed in [181]. The envelope of  $g_{\text{pre}} \cdot x_1(n)$  is calculated, weighted with the parameter  $g_{\text{bias}}$  and added to  $g_{\text{pre}} \cdot x_1(n)$  directly before the nonlinear mapping function. This simulates the signal-dependent bias-point shift that is happening in tube amplifiers due to a varying cathode voltage which alters the plate current. It enables the digital model to feature effects that are produced by vacuum-tube distortion systems such as blocking distortion where the bias-point of the tube shifts after a transient signal like e.g. palm-mutes. The intensity of this effect can be controlled by the parameter  $g_{\text{bias}}$ .

The distortion is introduced by the mapping function of the nonlinear block  $m(x_{\text{nl}})$ . A piecewise defined mapping function based on the hyperbolic tangent is used Eq. (3.7), which allows to shape positive and negative half-waves separately. It is described in detail in Sec. 3.2.1.

In a tube-based guitar amplifier the push-pull amplification stage uses one tube for each half-wave of the incoming signal where each stage might shape the signal differently depending on the surrounding circuitry, the type of the tubes or even production differences of the circuit elements. To

**Figure 9.2**

Signal flow graph of the nonlinear block. A similar nonlinear block has been used in Chap. 7 and Chap. 8. Here the dry/wet mixing has been omitted.

be able to reproduce this asymmetrical behavior the hyperbolic tangent mapping function is used which has proven to be beneficial when using an iterative optimization routine to adapt to various reference amplifier characteristics.

The parameters of the nonlinear block are aggregated in the parameter vector for the nonlinear block

$$\mathbf{p}_{\text{nl}} = (g_{\text{pre}} \quad g_{\text{bias}} \quad k_p \quad k_n \quad g_p \quad g_n \quad g_{\text{post}})^{\text{T}}. \quad (9.1)$$

The model is called *extended* Wiener–Hammerstein model because these models usually feature a memoryless nonlinear block while the configuration of the nonlinear block in this work employs a low-pass filter with memory.

Finally the distorted signal  $x_2(n)$  is filtered again by the output filter  $H_2(z)$  which determines the spectral shape of the overtones creating the specific sound of the distorted signal. The parameter vector  $\mathbf{p}$  contains the adjustable parameters of the digital model

$$\mathbf{p} = \begin{pmatrix} \mathbf{p}_{\text{H},1} \\ \mathbf{p}_{\text{nl}} \\ \mathbf{p}_{\text{H},2} \end{pmatrix}, \quad (9.2)$$

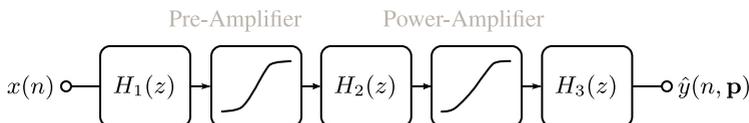
which leads to the output signal  $\hat{y}(n, \mathbf{p})$  which should be the same as the digitized output of the analog reference device  $y(n)$  (ground truth) if  $\mathbf{p}$  is adjusted properly.

### 9.1.2 Cascaded Model Topology

The basic structure of the cascaded digital model originated from [129] and can also be found in [128] or [185] and is a straightforward approach, which

gives convincing results if the digital model is properly adjusted. The model is a series connection of linear and nonlinear blocks which recreates the structure of a generic guitar amplifier. The process of adjusting these models manually is very tedious and would be nearly impossible without analysis and measurements of the circuitry. This study aims to investigate the performance of this model when subjected to an automated gray-box modeling procedure, using only input/output measurements as the source of information.

As Fig. 9.3 depicts, the input signal is filtered by the first filter  $H_1(z)$ , afterwards it is distorted by the first nonlinear block, which corresponds to the nonlinear behavior of the pre-amplifier. The output of the pre-amplifier is then filtered by  $H_2(z)$ , which corresponds to the tone-section of the guitar amplifier. Finally, the signal passes through the second nonlinear block, corresponding to the power stage of the guitar amplifier and is then filtered by the output filter  $H_3(z)$  resulting in the model output  $\hat{y}(n, \mathbf{p})$ .



**Figure 9.3**

*Digital model describing the structure of a generic guitar amplifier. The first mapping function corresponds to the pre-amplifier and the second one to the power amplifier. The linear blocks consist of input filter  $H_1(z)$ , tone filter  $H_2(z)$  and output filter  $H_3(z)$ .*

The first filter in an analog amplifier is mostly a first order RC-highpass or RC-bandpass filters described by  $H_1(z)$  and the output filter  $H_3(z)$  is determined by the frequency behavior of the output transformer.

For analog guitar amplifiers, the distinction between the blocks is not always this clear. For example many amplifiers are designed in such a way, that turning the drive- or gain-knob down, the frequency response of the input filter is also changed.

The last filter  $H_3(z)$  does not correspond to the impulse response of a loudspeaker. As described in Sec. 4.5, the amplifiers were measured without the influence of a real-world cabinet.

### Pre-Amplifier Nonlinearity

The basic structure of the nonlinear blocks remains unchanged from the studies presented in the previous chapters. The most important part of each nonlinear block is the mapping function because it defines the spectral

shape of the harmonics. As described in [128], a different mapping function has been used for the pre-amplifier block of the digital model.

The mapping function used in the first nonlinear block is a polynomial. The advantage of polynomial wave-shaping functions is the mathematical relationship between the coefficients of the polynomial and the shape of the harmonic overtones in the spectrum which is described in the following. Consider a polynomial function

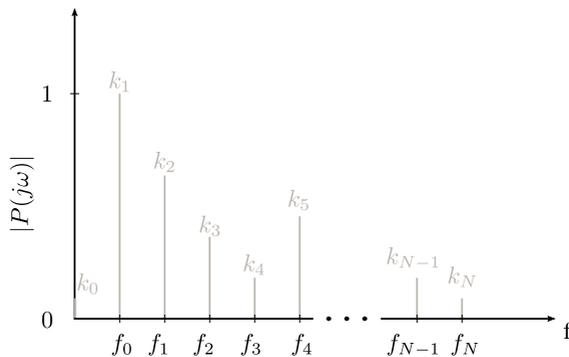
$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_Nx^N , \quad (9.3)$$

where  $x$  is the input variable (corresponding to the amplitude of the input signal) and  $a_n$  with  $n \in [0, N]$  are the coefficients of the polynomial.

Substituting  $x$  with  $\tilde{x}(t) = u \cdot \cos(\omega t)$ , it is possible to separate the different harmonic oscillations of the fundamental frequency  $\omega = 2\pi f$

$$p(\omega t) = k_0 + k_1\cos(\omega t) + k_2\cos(2\omega t) + \dots + k_N\cos(N\omega t) , \quad (9.4)$$

where the variables  $a_n$  and  $u$  have been combined into the harmonic variables  $k_n$ . Each  $k_n$  describes the amplitude of the  $n$ -th harmonic to the fundamental frequency  $f_0$  or  $\omega_0$  for a fixed input amplitude  $u$ . Fig. 9.4 depicts the harmonic variables  $k_n$  in frequency domain.



**Figure 9.4**

*Overtones of a sinusoidal signal after the polynomial mapping function.*

The relationship between harmonic variables  $k_n$  and polynomial coefficients  $a_n$  can be written in matrix form

$$\begin{pmatrix} A_{11} & A_{12} & \dots & A_{1N} \\ A_{21} & A_{22} & \dots & A_{2N} \\ A_{31} & A_{32} & \dots & A_{3N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \dots & A_{NN} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix} = \begin{pmatrix} k_0 \\ k_1 \\ k_2 \\ \vdots \\ k_N \end{pmatrix},$$

and solved for every  $a_n$ . With this technique it is possible to calculate the polynomial mapping function which creates the desired shape of overtones.

As an example the matrix equation is shown for 4 harmonics

$$\begin{pmatrix} 1 & 0 & u^2/2 & 0 & 3u^4/8 \\ 0 & u & 0 & 3u^3/4 & 0 \\ 0 & 0 & 0 & u^3/4 & 0 \\ 0 & 0 & 0 & 0 & u^4/8 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \begin{pmatrix} k_0 \\ k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix}.$$

The benefit of expressing the mapping functions with spectral coefficients instead of polynomial coefficients is the stability of the optimization process. If the polynomial coefficients  $a_n$  would be used as model parameters during optimization, unsuitable coefficient sets might be produced. By optimizing the spectral coefficients  $k_n$  as the models parameters no set of polynomial coefficients resulting in an unsuitable model output occurs which leads to better convergence during optimization.

### Power-Amplifier Nonlinearity

The second nonlinear block of the digital model shown in Fig. 9.3, corresponding to the power stage of the guitar amplifier is slightly different from the first nonlinear block. Instead of employing a polynomial mapping function the piecewise defined function based on the hyperbolic tangent is used. It was already used in [126, 127] to model distortion audio circuits and is described in Sec. 3.2.1. As for the Wiener–Hammerstein model, the blend stage was omitted in this nonlinear block. This relatively simple nonlinearity was chosen because the power stage of most guitar amplifiers is rarely driven into saturation due to the overall volume.

## 9.2 Parameter Optimization of the Wiener–Hammerstein Model

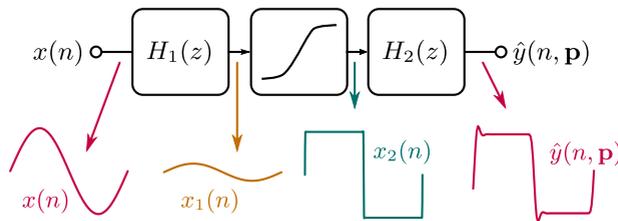
The optimization of the extended Wiener–Hammerstein model is carried out in two steps. First the linear blocks are adapted and afterwards the parameters of the nonlinear block are optimized with the Levenberg–Marquardt method [155, 156].

The filters  $H_1(z)$  and  $H_2(z)$  are measured with two exponentially swept sines according to [120]. First, the small signal impulse response is measured with an exponential sine sweep from 10 Hz to 21 kHz and an amplitude lower than 0.01 V. It is important that the amplitude of the sweep is kept low enough that no harmonic distortions occur during the measurement. The amplitude has to be chosen carefully so that absolutely no distortion, like crossover distortion from the power amplifier occurs.

The linear part of the sine sweep measurement is used as the impulse response  $h_{\text{low}}(n)$ . Afterwards the same measurement is repeated, but with an amplitude of 1 V. The linear part of the resulting impulse response is used as  $h_{\text{high}}(n)$ .

The low amplitude sweep is not exposed to the nonlinear behavior of the reference system and contains the influence of all its filters. The high amplitude sine sweep gets distorted. The saturation of the nonlinear elements in the circuit removes the influence of the filters in front of them by amplifying the sinusoid into a square wave which is then filtered by all the filters behind the nonlinear element.

This behavior can be simulated with a Wiener–Hammerstein model as depicted in Fig. 9.5. The input signal  $x(n)$  is filtered by  $H_1(z)$  which alters only amplitude and phase creating signal  $x_1(n)$ . The nonlinear block amplifies and saturates  $x_1(n)$  to the maximum possible amplitude, creating signal  $x_2(n)$ , thus negating the influence of the preceding filter  $H_1(z)$  on the amplitude of the signal. The output of the nonlinear block  $x_2(n)$  is then filtered by the last filter in the cascade,  $H_2(z)$ , creating the output signal  $\hat{y}(n, \mathbf{p})$ . The only measurable influence on the amplitude of  $\hat{y}(n, \mathbf{p})$  comes from the last filter after the input signal is distorted (in this case  $H_2(z)$ ).



**Figure 9.5**

*Influence of filters and nonlinear blocks while measuring with a sinusoidal signal.*

The obtained impulse responses are transformed into frequency-domain with a discrete Fourier transform using  $l_{\text{IR}} = 16384$  samples. To calculate the input filter, the small signal frequency response is divided by the large

signal frequency response,

$$H_1(z) = \frac{H_{\text{low}}(z)}{H_{\text{high}}(z)}. \quad (9.5)$$

The resulting filter is transformed back into time-domain and can directly be used in the digital model. The output of the nonlinear block is filtered with the second filter

$$H_2(z) = H_{\text{high}}(z) \quad (9.6)$$

which is derived from the linear part of the impulse response obtained with the high amplitude exponential sine sweep measurement.

According to Primavera et al. [124] if the frequency response measured with a high amplitude and a low amplitude sine sweep are the same, a nonlinear reference system can be well described by a Hammerstein model. This is implicitly done by calculating the first filter of the Wiener–Hammerstein model with Eq. 9.5. If the low amplitude and the high amplitude sweep produce the same frequency response, the first filter  $H_1(z)$  will be totally flat, reducing the Wiener–Hammerstein model to a Hammerstein model. Even if the assumption that the influence of  $H_1(z)$  is completely negated after the nonlinear block is not entirely true for all frequencies, the resulting model should still reproduce the behavior of the reference system adequately.

Usually a guitar amplifier has several cascaded amplification stages which introduce nonlinear distortion. By using input/output measurements it is not possible to quantify which stage produces how much distortion. It is assumed that the digital model is able to concentrate the effects from several stages into the combination of one nonlinearity with two filters.

In view of a real-time implementation of the digital model, the measured filters are approximated by second order IIR filters (biquads).

### 9.2.1 Adapting Second-Order IIR Filters

To achieve a proper frequency resolution, the measured impulse responses need to be quite long. In view of a real-time implementation of the digital model, the measured frequency responses can be approximated using second order IIR filters since the amount of additions and multiplications can be drastically reduced when using IIR filters instead of FIR filters. A measured impulse response with a proper frequency resolution can easily exceed a length of more than  $l_{\text{FIR}} > 4000$  samples at a sampling frequency of  $f_s = 44100$  Hz, which would lead to  $l_{\text{FIR}}$  multiplications and  $l_{\text{FIR}} - 1$  additions to calculate one output sample. One output sample of a second-order IIR filter can be computed using 5 multiplications and 4 additions. This means that  $n_{\text{Sec}} = 100$  second-order IIR filters in series are still more efficient to calculate than one FIR filter of length  $l_{\text{FIR}} = 4096$ . The total

FIR				IIR			
$l_{\text{FIR}}$	+	$\cdot$	$\sum$	nSec	+	$\cdot$	$\sum$
1024	1023	1024	2047	20	80	100	180
2048	2047	2048	4095	50	200	250	450
4096	4095	4096	8191	100	400	500	900

**Table 9.1**

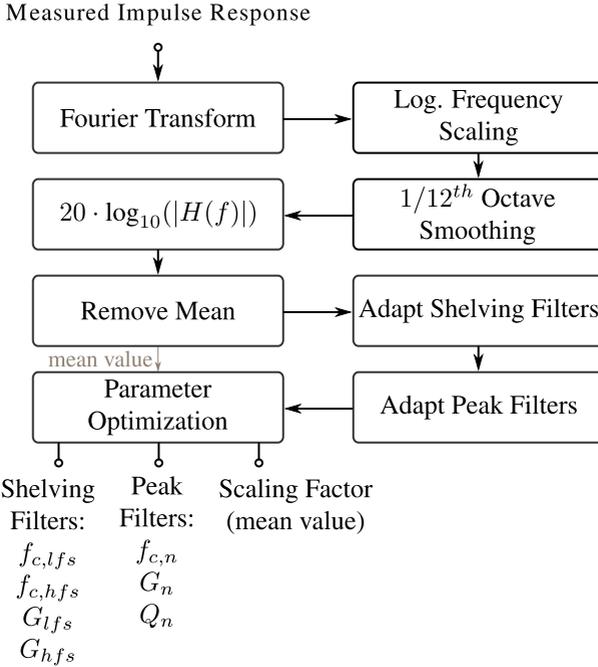
*Number of operations when calculating one output sample for different digital filter implementations.*

number of additions for the FIR filter is 4095 and the number of multiplications is 4096, while  $nSec = 100$  second-order IIR sections would yield 500 multiplications and 400 additions. If additions and multiplications are considered as operations, an FIR filter of length 4096 would have 8191 operations in contrast to the IIR version with only 900 operations. So with second-order sections the number of operations are reduced drastically. The number of necessary second-order sections  $nSec$  strongly depends on the measured impulse response and its application. For example, a distortion circuit usually contains first or second order analog filters which will have a much smoother frequency response than the impulse response of a speaker cabinet in a room, which contains considerably more peaks and notches in frequency domain. In this case more second-order sections are needed to accurately describe the behavior of the reference system.

For this reason a routine was developed which iteratively adjusts parametric second-order IIR sections to recreate the magnitude frequency response. It is similar to the procedure used in [186] to equalize loudspeakers but in this work it was used to recreate the desired frequency response rather than equalizing it. The flow diagram of this procedure is shown in Fig. 9.6.

All filters are connected in series and the cascade contains a high frequency shelving filter (HFS), a low frequency shelving filter (LFS) and  $N_{\text{filt}}$  peak filters, designed according to Zölzer et al. [98, pp. 63 – 66]. The filter design allows to adjust gain and cut-off frequency of the shelving filters as well as Q factor, gain and center-frequency of the peak filters. All filters are initialized to have a gain of 0 dB which will result in a flat frequency response. Please note that the filter design yields minimum phase filters. Most analog filters also feature minimum phase and therefore it is not expected that the filter cascade will introduce a much higher latency than the filters inside the reference device itself.

At first, 1/12th octave smoothing is performed on the logarithmically spaced frequency response using a triangular window. Then the ordinate is transformed into logarithmic domain. The mean value of the frequency response is removed to center it around 0 dB. The mean value is stored and



**Figure 9.6**

*Flow diagram of the biquad adaptation to reproduce the magnitude frequency response.  $n \in [1 \ N_{\text{filt}}]$ .*

can be used to properly scale the frequency response back to its original value.

Afterwards the filters are adapted starting with the shelving filters. Their gain is set to the values of the frequency response at  $f_{dc} = 0$  for the LFS and  $f_{max} = f_s/2$  for the HFS. Then the cut-off frequency of both filters is altered from  $f_{c,min}$  to  $f_{c,max}$  and the frequency yielding the lowest error is chosen. The cut-off frequency ranges from  $f_{c,min} = 15$  kHz to  $f_{c,max} = 21$  kHz for the HFS and  $f_{c,min} = 50$  Hz to  $f_{c,max} = 1$  kHz for the LFS with a grid size of 500. The error is expressed as the sum of squares between the frequency response of the current shelving filter and the preprocessed frequency response.

In the second filter adaption step a cascade of  $N_{\text{filt}} = 10$  peak filters is adjusted iteratively. The center frequency and gain is directly determined from the magnitude frequency response by picking the largest peak or notch. Afterwards the sum of squares of the error between prepro-

cessed frequency response and the frequency response of the peak filter is calculated for 500 Q factors between  $Q_{\min} = 10^{-1}$  and  $Q_{\max} = 10^2$ . The Q factor producing the lowest sum of squares between the frequency response of the peak filter and the current frequency response is chosen. To compute the frequency response of the next iteration the influence of the peak filter is removed from the frequency response and the next filter is adapted.

When this procedure is finished, the filter parameters are optimized with the Levenberg–Marquardt method. The cost-function describes the quadratic error between the measured magnitude frequency response and that of the filter cascade. During optimization the error is only calculated between  $f_{\min} = 20$  Hz and  $f_{\max} = 20$  kHz because errors outside this range produce no audible differences. The control parameters of each second order IIR filter and the removed mean value are the result of this routine. The coefficients of the filters can be calculated according to the aforementioned filter design [98]. These parameters are collected in the filter parameter vectors  $\mathbf{p}_{H,1}$  and  $\mathbf{p}_{H,2}$  for input and output filter block.

### 9.2.2 Wiener–Hammerstein Model: Optimization of Nonlinear Parameters

After the linear blocks of the digital model have been adjusted, the parameters of the nonlinear block are optimized by minimizing the cost-function  $C(\mathbf{p})$  (see Sec. 4.3.3). Please note that the filter parameters  $\mathbf{p}_{H,1}$  and  $\mathbf{p}_{H,2}$  are fixed during optimization to reduce its complexity.

First, a grid search is carried out for the pre- and post-gain of the nonlinear block, because these parameters control the amount of overtones and the output amplitude, which have the highest influence on the model output. The result of the spectrogram-based cost-function  $C(\mathbf{p})$  from Sec. 4.3.3 is calculated for each combination of the two parameters and the one resulting in the lowest error is chosen as initial values for the Levenberg–Marquardt method.

The mapping function is initialized with  $k_p = k_n = 0.3$ ,  $g_p = g_n = 6$  dB and  $g_{\text{bias}} = 0$ . It proved to be beneficial to allow a limited number of control parameters for the filters to be changed during optimization, which are the cut-off frequencies  $f_{c,\text{LFS}}$ ,  $f_{c,\text{HFS}}$  and gains  $g_{\text{LFS}}$ ,  $g_{\text{HFS}}$  of the low- and high-frequency shelving filters, which were used when approximating the measured frequency responses of  $H_1(z)$  and  $H_2(z)$ . At the band edges, where the filters of the analog reference device exhibit strong attenuation, there *might* be some uncertainty in the measurement, meaning that characteristics which should belong to the first filter of the digital model might be mapped to the second filter and vice versa. Using the control parameters of the filters during optimization helps reducing possible measurement errors.

The optimization is then started with the parameters of the nonlinear block  $\mathbf{p}_{\text{nl}}$  (Eq. (9.1)) and the 8 control parameters of LFS- and HFS-filters

$$\mathbf{p} = \begin{pmatrix} \mathbf{p}_{\text{nl}} \\ f_{c,\text{LFS1}} \\ g_{\text{LFS1}} \\ f_{c,\text{LFS2}} \\ g_{\text{LFS2}} \\ f_{c,\text{HFS1}} \\ g_{\text{HFS1}} \\ f_{c,\text{HFS2}} \\ g_{\text{HFS2}} \end{pmatrix}.$$

The input signals used for identification consisted of short recordings from different guitars, a bass guitar and a multi-frequency sine wave (see Eq. (8.2) in Sec. 8.2.1).

### 9.3 Parameter Optimization of the Cascaded Model

This section describes the steps needed to adapt the cascaded model to a reference system. The process is subdivided into several steps to assure that the iterative optimization does not get stuck in a local minimum. At first the linear part of the reference system is measured and afterwards the parameters of the nonlinear blocks are optimized.

#### 9.3.1 Cascaded Model: Filter Optimization

An exponential sine sweep is sent through the reference device and the recorded output is convolved with an inverse filter (see [119]). The resulting impulse response contains the linear impulse response as well as different impulse responses for higher order harmonics. In this work only the impulse response corresponding to the linear part of the circuit is used because all nonlinear behavior should be modeled with the mapping function.

At first the same filter identification as for the Wiener–Hammerstein model is carried out to measure the first  $H_1(z)$  and last  $H_3(z)$  filter of the model. A 256 tap finite impulse response (FIR) filter  $H_2(z)$  is additionally adapted to further refine the small signal frequency response of the reference device. The FIR filter is located between the nonlinear blocks of the model.

The parameters for the linear part of the model are the FIR filters coefficients. The input signal for the adaptation is the above mentioned low amplitude sine sweep. The cost function calculates the difference of the magnitude spectrum of the reference system and the digital model

output. The filter coefficients are adjusted to minimize the error between both spectra.

The length of the filter is a trade-off between computational complexity during optimization and frequency resolution and was chosen empirically. Finite differences are used to approximate the derivative of the model output with respect to each parameter (in this case 256 coefficients). This leads to time-consuming calculations, because the digital model has to be evaluated once for each derivative.

### 9.3.2 Cascaded Model: Optimization of Nonlinear Parameters

After the small signal frequency response has been adapted, the filter coefficients of the FIR filter are not changed anymore, only the parameters of the nonlinear blocks can be altered.

Each nonlinear block features a multiplication with a variable gain (pre- and post-gain) of input signal and output signal. The intensity of the bias-point shift is adaptable for each nonlinear block Fig. 9.2. For the pre-amplifier nonlinearity the mixing stage has an additional parameter which is used to mix dry and wet signal.

The first nonlinear block of the digital model also uses the parameters mentioned in Sec. 9.1.2. To limit the number of parameters, only the first 39 overtones  $k_0 - k_{40}$  can be adapted (see Fig. 9.4). The optimization routine only alters the  $k_n$  parameters from which the polynomial coefficients  $a_n$  are computed. If the polynomial coefficients are used as parameters, too many unsuitable solutions would be possible and the optimization routine would not converge as well. The typical fundamental frequency region of an electric guitar in standard tuning ranges from 80 Hz to 1100 Hz, depending on the number of frets. 40 harmonics do not cover the whole frequency region for the tones with the lowest pitch, but usually the contribution to the overall spectrum of the 40<sup>th</sup> harmonic is negligible.

The nonlinear parameters are adapted for different input signals (of different complexity) and with different cost functions to assure convergence of the parameters into their global minimum. The used algorithm is the Levenberg–Marquardt optimization routine, as described in Chap. 4 with different cost functions.

- At first, a grid search for the pre- and post-gain of the power-amplifier nonlinearity is performed because these parameters have the strongest influence on the shape of the output envelope of the digital model. The cost function calculates the difference of the envelopes of the output signals and the gain combination with the lowest error is chosen. The envelope is calculated by low-pass filtering the absolute value of the signal. The cut-off frequency of the used low-pass filter is  $f_c = 10$  Hz as described in Sec. 4.3.2.

- Afterwards all nonlinear parameters are adapted at the same time. The cost function, however, was designed differently in this optimization step. It calculates the sum of squares between digital model and reference system

$$C(\mathbf{p}) = \sum (y(n) - \hat{y}(n, \mathbf{p}))^2, \quad (9.7)$$

with  $y(n)$  as the (digitized) output of the reference system and  $\hat{y}(n, \mathbf{p})$  as the output of the digital model.  $\mathbf{p}$  is the parameter vector of the digital model. In this optimization step, all filters except the input filter  $H_1$  are bypassed during optimization. The chosen input signal is a 1000 Hz sine wave with amplitudes from 1 V down to 0.001 V. This first step helps to find a set of parameters which can be used as initial parameters in the next optimization step where all filters in the digital model are turned back on.

- The next optimization step is done with a multi-frequency sine wave described by Eq. (6.14). The phase shift between each frequency is chosen in such a way that the peak-factor of the sum of the different frequencies is minimal and the signal has a flat power-spectrum [176]. The spectrogram-based cost-function described in Sec. 4.3.3 is used in this step.
- For guitar amplifiers with nonlinear behavior it is beneficial to add a last step, where the spectrogram-based cost-function is used (see Sec. 4.3.3) but a recorded guitar track is used as input, which consists of a combination of guitar tones and chords to further refine the parameters of the digital model.

## 9.4 Results

After performing the system identification the adapted digital models have to be evaluated. At first objective error metrics are calculated. If these metrics are close to zero (or one for the correlation coefficient), the quality of the model is good. Sometimes the objective scores give a ‘bad’ result which can not be confirmed by informal listening tests. This is why a subjective listening tests was performed additionally to rate the perceived quality of the models.

### 9.4.1 Wiener–Hammerstein Model

The amplifiers chosen for the Wiener–Hammerstein model evaluation were modeled in different settings and evaluated with different guitars using different pick-ups.

Different input signals are used for each amplifier. Input signals from three guitars with different pick-ups are tested. Additionally some input signals were taken from the Fraunhofer IDMT guitar database [187]. The guitar used in the IDMT dataset is the Ibanez – RG2820 with DiMarzio humbuckers which will be referenced as (HM3). An overview of the used guitars and the acronyms which will be used for each guitar in this section is given in Tab. 9.2.

#	Guitar	Acronym
1	Ernie Ball stratocaster-style guitar singlecoil pick-ups	SC
2	Reverend – Buckshot humbucker pick-ups, medium output voltage	HM1
3	Hagstrom – Super Swede Gibson humbuckers, higher output voltage	HM2
4	IDMT database: Ibanez – RG2820 DiMarzio humbuckers	HM3

**Table 9.2**

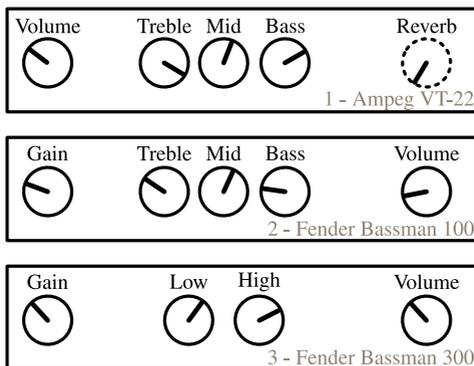
*Overview of the guitars which are used for evaluating the optimized models. The acronyms will be used throughout the section to refer to the type of input signal.*

For the evaluation of the Wiener–Hammerstein model the ‘Gain’ and ‘Volume’ knobs were altered from a low value (9 o’clock position) to medium value (12 o’clock position) to high value (3 o’clock position). The settings are marked with acronyms, e.g. ‘HGMV’ meaning ‘high gain medium volume’. The tone section was set to a 12 o’Clock position which is the middle value of the knob. In this study the settings focus more on distorted amplifiers which are more challenging to model.

### 9.4.2 Cascaded Model

The results are evaluated for different amplifier models and for several guitar signals. Some amplifiers are tested in multiple settings, creating distortion with the pre-amplifier, the power-amplifier, or both. Other amplifiers are tested in an artist preferred setting, where the user controls of the amplifiers are not altered from the settings the artists used in the rehearsal room. Please note that the amplifiers in the artist preferred setting did not introduce a lot of distortion which would make the modeling process more challenging.

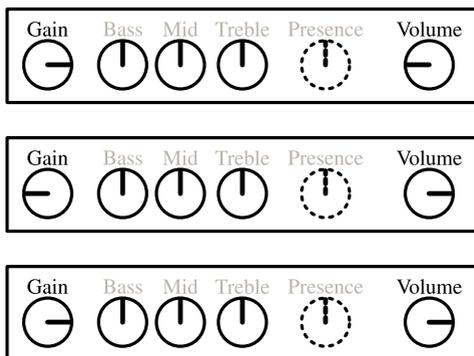
The settings of the amplifiers in the ‘artist preferred’ configuration are shown in Fig. 9.7. All amplifiers produced very little distortion in the output signal. The first amplifier (top), the Ampeg VT-22, did not feature



**Figure 9.7**

*Settings for amplifiers in clean setting. 1.) Ampeg VT-22 (top) 2.) Fender Bassman 100 (middle) 3.) Fender Bassman 300 (bottom)*

separate controls for pre-amplifier and power-amplifier and the reverb was turned off. The Fender Bassman 100 (middle) and Fender Bassman 300 (bottom) were set up to introduce almost no distortion, as can be seen by the gain and volume controls.



**Figure 9.8**

*Settings for pre-amp distortion (top), power-amp distortion (middle) and heavy distortion (bottom).*

Only the amplifiers which were modeled in the ‘artist preferred’ setting, were set up to have a clean sound, introducing very little distortion in the output signal. The amplifiers which introduced more distortion in the output signal were modeled in multiple settings:

1. High gain and low volume (pre-amp distortion)
2. Low gain and high volume (power-amp distortion)
3. high gain and high volume (heavy distortion)

These settings are illustrated in Fig. 9.8. The tone-section of the amplifiers was set to a medium value in 12 o’clock position and the presence knob is depicted with a dashed line, because only the Marshall – JCM 900 featured a presence control. The presence knob controls an additional filter in the amplifier resulting in more high-frequency content in the output signal.

### 9.4.3 Objective Scores: Cascaded Model

The results in Tabs. 9.3 and 9.4 show that the cascaded model works very well with clean or almost clean amplifiers. For the Bassman 100  $e_{\text{ESR}}$  remains below 0.1 and the correlation coefficient never drops below 0.96. The RMS and PEAS values are also very low which indicates that the identification was successful.

<b>Fender Bassman 100</b>	<b>PEAS</b>	$e_{\text{RMS}}$	$r_{y\hat{y}}$	$e_{\text{ESR}}$
Single coil (SC)	0.0541	0.0003	0.9668	0.0668
Humbucker medium (HM1)	0.0146	0.0001	0.9656	0.0689
Humbucker high (HM2)	0.0002	0.0001	0.9752	0.0518

**Table 9.3**

*Objective scores for the Bassman 100 with no distortion.*

<b>Ampeg VT-22</b>	<b>PEAS</b>	$e_{\text{RMS}}$	$r_{y\hat{y}}$	$e_{\text{ESR}}$
Single coil (SC)	0.0954	0.0001	0.9631	0.0745
Humbucker medium (HM1)	0.0200	0.0006	0.9582	0.0840
Humbucker high (HM2)	0.0240	0.0012	0.9419	0.1159

**Table 9.4**

*Objective scores for the VT-22 with very little distortion.*

The VT-22 also gives very good results, but when the input signal level is high, the error becomes higher too. This can be seen from the results in

Tab. 9.4. For the single coil guitar input the  $e_{\text{ESR}} = 0.0745$  but if the input voltage is higher, which leads to more distortion,  $e_{\text{ESR}}$  gets above 0.1. The correlation coefficient and  $e_{\text{RMS}}$  have the same tendency as  $e_{\text{ESR}}$ . One particularity is that the PEAS value for the single coil input is higher than for both humbuckers but this does not lead to a perceptible difference.

<b>Madamp A15Mk2</b>	<b>PEAS</b>	$e_{\text{RMS}}$	$r_{y\hat{y}}$	$e_{\text{ESR}}$
Single coil (SC)	0.5047	0.0619	0.1150	1.7863
Humbucker medium (HM1)	0.4365	0.3121	-0.2350	2.4995
Humbucker high (HM2)	0.4423	0.3172	-0.2849	2.6122

**Table 9.5**

*Objective scores for the A15Mk2 with power-amp distortion (low gain, high volume).*

Any reference device can add distortion either by increasing the gain, which leads to a clipping pre-amplification stage or by increasing the volume, which leads to a clipping power-amplification stage. The power-amplifier in the reference device is rarely turned up to high values, because it reaches very high sound pressure levels when the amplifier is connected to a speaker [163], but while measuring, only a dummy-load was connected to the reference device.

<b>Marshall JCM 900</b>	<b>PEAS</b>	$e_{\text{RMS}}$	$r_{y\hat{y}}$	$e_{\text{ESR}}$
Single coil (SC)	0.4625	0.03	0.001	2.0127
Humbucker medium (HM1)	0.3727	0.1075	-0.1237	2.2390
Humbucker high (HM2)	0.4106	0.0998	-0.0051	1.9873

**Table 9.6**

*Objective scores for the JCM 900 (Channel B) with pre-amplifier distortion (high gain, low volume).*

<b>Marshall JCM900</b>	<b>PEAS</b>	$e_{\text{RMS}}$	$r_{y\hat{y}}$	$e_{\text{ESR}}$
Single coil (SC)	0.5969	0.1725	-0.1120	2.5827
Humbucker medium (HM1)	0.4144	0.2423	-0.3836	2.7626
Humbucker high (HM2)	0.4310	0.2396	-0.3253	2.6443

**Table 9.7**

*Objective scores for the JCM900 with maximum distortion (high gain, high volume).*

The results of the modeling process are shown in Tab. 9.5. Due to the nonlinear behavior of the reference device, the error does not increase proportionally with a rising input level and is already quite high.

When using the pre-amplifier to create a distorted guitar sound (see Tab. 9.6), the results are roughly comparable to the clipping power amplifier. For the humbucker guitars the scores are nearly the same for the same input signals. Merely the score for the single coil input deviates slightly between the clipping pre-amplifier and power-amplifier but in general no conclusion can be drawn from a single test.

For all amplifiers which introduced distortion,  $e_{\text{ESR}}$  is always greater than one which would mean that the time-domain error signal has more energy than the corresponding reference signal. This confirms the assumption that the error to signal ratio is not a suitable metric to evaluate the similarity between two audio signals.

Finally, the objective scores for the reference amplifier which introduced the most distortion in the output signal are shown in Tab. 9.7. In this case, the error energy is always higher than the actual signal energy, since  $e_{\text{ESR}}$  is always greater than 2 for all test items. This is also the model which has the greatest deviation from the reference device from a perceptual point of view.

A negative correlation coefficient  $\mathbf{r}_{\mathbf{y}\hat{\mathbf{y}}}$  can occur when the model output has a phase shift of more than  $\pi$  with respect to the reference system. This can occur since the phase is ignored during the calculation of the spectrogram cost-function.

#### 9.4.4 Objective Scores: Wiener–Hammerstein Model

In the following the objective scores for the Wiener–Hammerstein model are discussed. They are presented in Tab. 9.8. All test-items from the table are also presented to the test subjects in the listening test.

The error to signal ratio is not very well suited for evaluation because the results do not represent the human perception of differences between model and reference. Item 8 in Tab. 9.8 shows an  $e_{\text{ESR}}$  value of 1.7936 which means that the time-domain error has more energy than the reference signal. As the listening test will confirm, this item is not that bad from a subjective point of view. This is reflected by the similarity score (PEAS) as well. If  $e_{\text{ESR}}$  has a value close to zero, however, the results are good.

The evaluation of the cascaded model showed that the objective scores for heavily distorted reference amplifiers were not as good as for amplifiers in a clean or a slightly distorted setting. This trend can also be seen for the Wiener–Hammerstein model. The items are sorted with increasing distortion per amplifier. The PEAS-score shows an increasing tendency if the amount of distortion increases as well. But the scores are lower when compared to the results with the cascaded model. This indicates

	<b>Amp (Setting)</b>	<b>Pick-Up</b>	<b>PEAS</b>	$e_{\text{RMS}}$	$r_{\hat{y}y}$	$e_{\text{ESR}}$
01	Bassman 100 (LGLV)	HM3	0.0580	0.0001	0.9952	0.0100
02	Bassman 100 (LGLV)	SC	0.0058	0.0001	0.9972	0.0077
03	Roost SR22 (HGMV)	HM2	0.1353	0.0032	0.9851	0.0296
04	Roost SR22 (HGMV)	HM3	0.1838	0.0039	0.9719	0.0555
05	JCM 900 (LGMV)	SC	0.0423	0.0005	0.9742	0.0546
06	JCM 900 (MGLV)	HM1	0.2131	0.0034	0.8583	0.2846
05	JCM 900 (MGMV)	HM3	0.2902	0.0261	0.7034	0.5919
08	JCM 900 (HGLV)	HM3	0.3172	0.0151	0.6571	1.7936
09	JCM 900 (HGMV)	HM3	0.3942	0.0567	0.5619	0.9386
10	JCM 900 (HGHV)	HM2	0.4175	0.0342	0.6282	0.7413
11	A15Mk2 (LGLV)	HM3	0.1151	0.0007	0.9898	0.0206
12	A15Mk2 (LGLV)	Bass	0.1397	0.0113	0.9211	0.1518
13	A15Mk2 (MGMV)	HM3	0.2210	0.0155	0.9296	0.1370
14	A15Mk2 (MGHV)	HM3	0.2807	0.0269	0.9112	0.1741
15	A15Mk2 (HGHV)	HM3	0.2947	0.0095	0.8536	0.6577

**Table 9.8**

*Objective scores for evaluation of the optimized Wiener–Hammerstein model.*

that the Wiener–Hammerstein model performs better than the cascaded model which will also be confirmed by the results of the listening test.

### 9.4.5 Listening Test: Cascaded Model

A listening test was conducted to see how well the adapted models perform for a human test subject. An ABC/HR style listening test was chosen. It aimed at rating the adapted model in relation to the analog reference device. The test subjects were presented with a reference item and two test items. The items should be rated according to how **similar** they sound to the reference, where 100 represents no detectable difference between the item and the reference and 0 represents a very annoying difference. One of the test items was a hidden reference, which was the same audio-file as the reference item. The similarity rating is divided into 5 categories which are shown in Tab. 9.9 with their identifier and corresponding score.

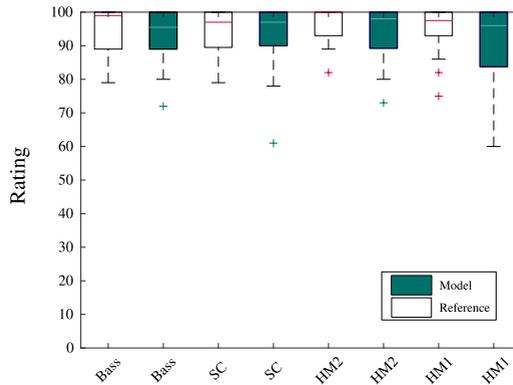
The listening test featured 20 listening examples with hidden reference and digital model output. 15 participants have taken the test from which

Identifier	Score
Imperceptible	80 – 100
Minor Differences	60 – 80
Differences	40 – 60
Major Differences	20 – 40
Annoying Differences	0 – 20

**Table 9.9**

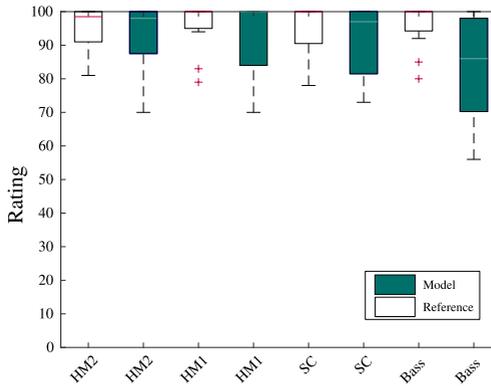
*Listening test rating score with corresponding identifier.*

8 were experienced listeners, 4 were musicians and 3 were unexperienced listeners. At the end of the test, each participant had the option to comment on the test. The framework for the listening test was the ‘BeagleJS’ framework, described in [178]. It features example configurations for ABX and Mushra style listening tests. The Mushra configuration was adapted to fit the needs for model – reference comparison. The test results have

**Figure 9.9**

*Cascaded model: results of the listening test for the Fender – Bassman 100 (clean setting).*

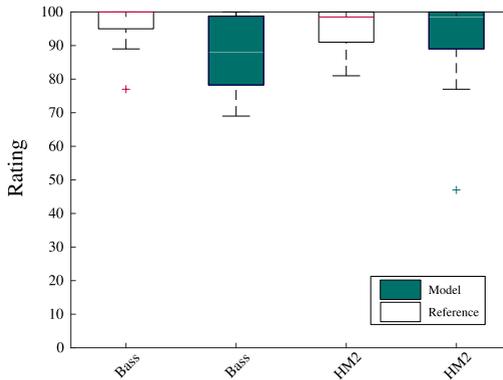
been cleaned by deleting the ratings where the hidden reference was rated with a score lower than 80, but only one test subject was removed from the evaluation completely, because for 13 of 20 items, the hidden reference was rated with scores much lower than 80. Figures 9.9 – 9.12 show the results of the listening test. The bar displays the 50% quantile (median) for each item. The lower and upper bounds of the box represent the 25% quantile or the 75% quantile respectively. Outliers are depicted as crosses. Figures 9.9 and 9.10 show the results for the reference amplifiers in clean



**Figure 9.10**

*Cascaded model: results of the listening test for the Ampeg – VT-22 (clean setting).*

setting and the adapted models. The results show that the digital model is always rated in the same range as the analog reference device for the test items ‘Bass’, ‘Single Coil (SC)’, ‘Humbucker 1 (HM1)’ and ‘Humbucker 2 (HM2)’. These results confirm that the model is very well adapted as the objective scores suggest. The results for the Ampeg VT-22 are similar to



**Figure 9.11**

*Cascaded model: results of the listening test for the Fender – Bassman 300 in clean setting.*

the results of the Fender Bassman 100. In some cases there was an unwanted ‘crackling’ noise in the recording of the reference amplifier, which was not reproduced by the digital model. This made it possible to identify the difference between the hidden reference and the digital model output.

The last amplifier which is in an almost clean setting was the Fender Bassman 300 (Fig. 9.11). Nevertheless, the HM2 (humbucker with high output voltage) test item had a nearly identical rating as the hidden reference. Only for the input signal from an electric-bass, there were minor audible differences in the output signal. These results are in agreement with the comments from the participants. Several stated, that they could not perceive any difference when the amps were in a ‘clean’ or ‘almost clean’ setting.

The results of the optimization routine for distorted reference amplifiers are not as good as the results for the clean ones. The more nonlinear the amplifier becomes (more distortion), the higher is the perceptible difference between digital model and analog reference device.

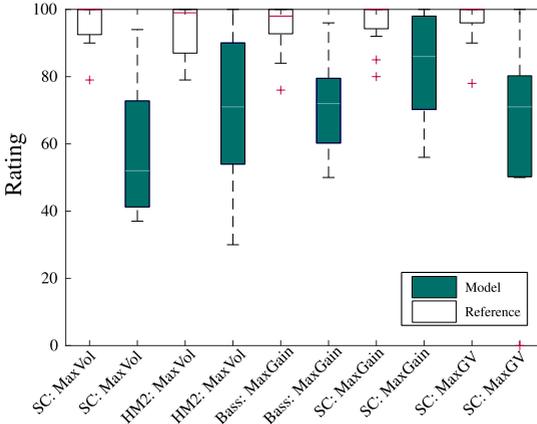
The clipping power-amplifier of the Madamp A15Mk2 is rated worse than the clipping pre-amplifier, as shown by the single-coil (SC) items in Fig. 9.12a. In the listening test, the clipping pre-amplifier was rated with  $\approx 90$  (median) and the clipping power-amplifier with  $\approx 50$  (median), in comparison with the hidden reference, which had a median of 100 in both cases.

The listening test results for the last amplifier confirm the assumption that a reference device with highly nonlinear behavior is not identified as well as a system with little nonlinear behavior. The Marshall JCM 900 was rated worse if both pre- and power-amplifier were at high values, in comparison to the first 2 test items where only the pre-amplifier was set to a high value. A common comment from the participants was that a difference in the noise floor between digital model and reference device made it possible to distinguish the reference from the model.

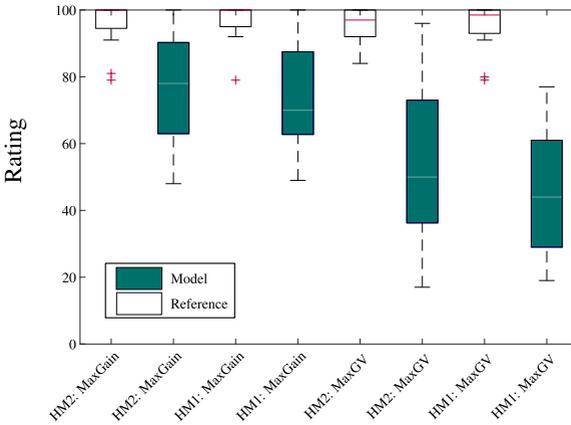
It is possible to tune the digital model by hand, although it is not recommended. This is an indication that the model is able to recreate also highly nonlinear systems. All signals were recorded while the amplifier was not connected to a cabinet. The influence of a cabinet could lead to reduced high frequency content in the output signal, which could lessen the perceived difference between reference device and digital model.

#### 9.4.6 Listening Test: Wiener–Hammerstein Model

The same style of listening test that was performed for the cascaded model was conducted for the Wiener–Hammerstein model to be able to compare the results. The test subjects were presented with a reference item and two test items. The items should be rated according to how **similar** they sound to the reference, where 100 represents no detectable difference between



(a) Madamp A15Mk2 with different distortion settings.



(b) Marshall - JCM 900 with different distortion settings.

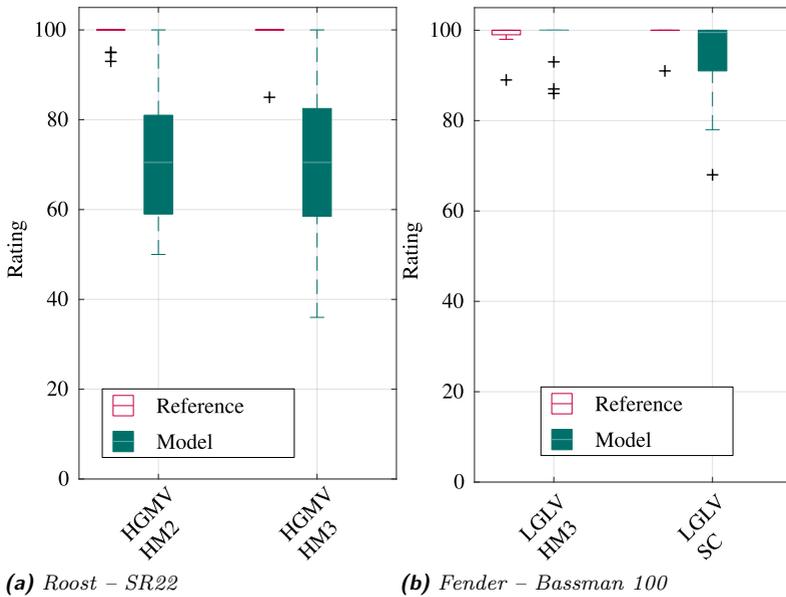
**Figure 9.12**

Cascaded model: results of the listening test for distorted guitar sounds

(a) Madamp - A15Mk2 and (b) Marshall - JCM 900

the item and the reference and 0 represents a very annoying difference according to Tab. 9.9.

One of the test items was a hidden reference, which was the same audio-file as the reference item, the other item was the output of the digital model. The test had 35 participants from which only 19 were used for the final evaluation. The other participants were not able to detect the hidden

**Figure 9.13**

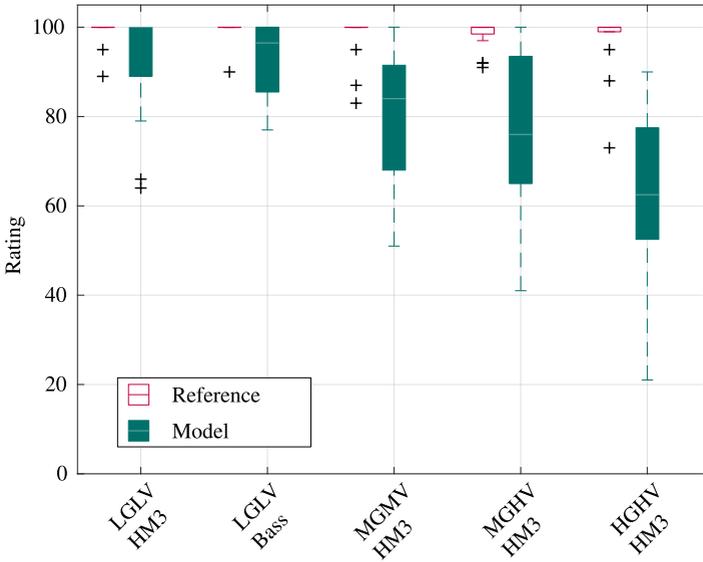
*Wiener-Hammerstein model: results of the listening test (a) Roost - SR22 (b) Fender - Bassman 100 for two different input signals.*

reference or always rated the hidden reference with scores below 80 and were therefore excluded from the evaluation.

The results of the listening test are visualized in Fig. 9.13 – Fig. 9.15. The bar displays the 50% quantile (median) for each item. The lower and upper bounds of the box represent the 25% quantile or the 75% quantile respectively. Outliers are depicted as +.

The general trend for each amplifier is that the results get worse if the nonlinearity of the reference system increases (listening test and objective scores) but no amplifier was rated worse than ‘minor differences’ on average. The similarity score corresponds quite well to the results of the listening test, for example when comparing the objective scores for the Madamp A15Mk2 (Fig. 9.14) to the results of the listening test. If the similarity score is below 0.15, differences between model and reference signal are nearly imperceptible, as is shown by the results of the listening test.

The listening test was performed without an anchor, which is an intentionally bad test item usually used to scale the subjective range for each



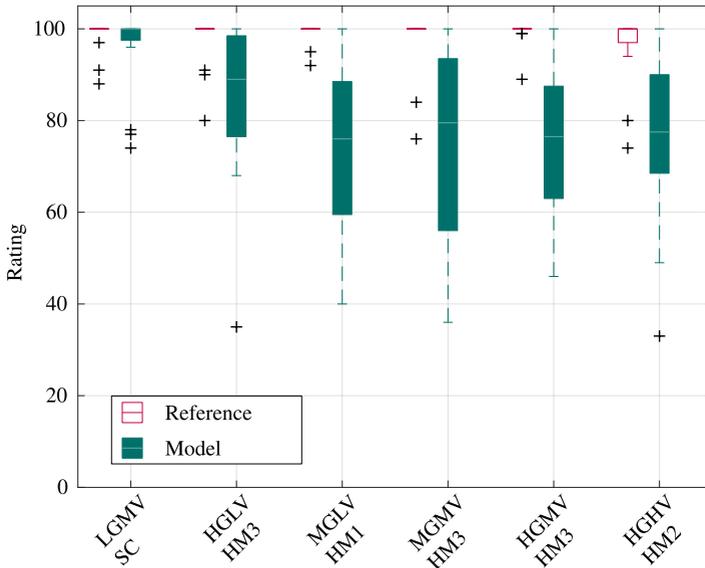
**Figure 9.14**  
*Wiener–Hammerstein model: results of the listening test: Madamp A15Mk2.*

test subject. This can explain the relatively large spread of the results. In future work an anchor should be used for the listening test.

### 9.5 Conclusion

Different guitar amplifiers have been modeled using two different digital models. In a first attempt a digital model is used which recreates the structure of a generic guitar amplifier using three filters and two nonlinear blocks. Guitar amplifiers in a clean setting or introducing little distortion can be modeled properly and the listening test proved that they are indistinguishable from the reference device. But when modeling guitar amplifiers introducing a lot of distortion with an automated optimization routine the results are not convincing. The objective scores and the listening test show a noticeable difference between digital model and analog reference device for amplifiers with distortion. They are not rated as ‘annoying’ or as ‘major differences’ but are noticeable even for an untrained listener.

The polynomial nonlinearity which is used in the pre-amplifier block of the digital model adapts very well to weakly nonlinear systems but fails



**Figure 9.15**

*Wiener–Hammerstein model: results of the listening test: Marshall – JCM900.*

to yield good results for highly distorted amplifiers. The main part of the distortion comes from the power amplifier block of the digital model even if the reference system only produces distortion in the pre-amplifier stage. For this reason a simpler model structure is also considered by removing the pre-amplifier block and the second filter from the digital model which results in a Wiener–Hammerstein model topology.

When using a simpler model the optimization routine can also be simplified resulting in shorter optimization times and a more robust routine. The filters of the digital model can be measured. Only one grid search and one iterative parameter optimization step is needed to adjust the Wiener–Hammerstein model whereas one grid search and three iterative optimization steps are needed to adjust the cascaded digital model. To reduce uncertainty in the filter measurements the control parameters of the used high- and low-frequency shelving filters are used during parameter optimization.

The results from the Wiener–Hammerstein model are more convincing than of the cascaded model. The objective scores and the listening test prove that the simpler model outperforms the complex model when using an automated optimization routine. Even though the objective scores

are falling off in quality the more nonlinear the reference device gets, no amplifier is rated worse than ‘minor differences’ on average.

Considering the listening tests, the digital version of the JCM 900 with the cascaded model has an average rating around 40 – 50 (see Fig. 9.12b) which corresponds to ‘differences’ with a trend to ‘major differences’ on the listening test scale. The Wiener–Hammerstein model always has an average rating between 70 and 80 for the same amplifier with heavy distortion, which corresponds to ‘minor differences’ as shown in Fig. 9.15.

In the future a real-time implementation of the digital model would be desirable to test the performance when playing the digital model with a guitar. Switching between the reference and the model makes even the smallest differences in dynamic behavior audible.

---

## Summary

---

System identification or virtual analog modeling of nonlinear audio circuits for electric guitars is the aim and topic of this work. Virtual analog modeling describes the process of transforming an analog reference system into a mathematical model that can be computed on digital devices. Musicians like the sound of specific (vintage) equipment but are concerned about the limitations of analog devices like weight, age or price due to scarcity. This is the reason why digital solutions gain more and more popularity and there is ongoing research about different virtual analog modeling methods.

Contrary to already well researched methods, like circuit simulations, this work aims at using gray-box modeling techniques to recreate an analog device digitally. The main advantage of this method is that no knowledge about the reference system is needed. The proposed method is best characterized as semi-physical gray-box modeling. Knowledge about the physical structure of each type of system is used to construct a generalized block-oriented model. Afterwards black-box modeling methods are used to adapt the model by sending specific input signals, measuring the corresponding output and adapting the digital model to recreate the behavior of the analog reference device. The digital models only consist of linear time-invariant and nonlinear blocks which are connected according to the physical structure of the type of reference device.

The focus of this work lies on two groups of nonlinear audio effects. Dynamic range compression systems and distortion systems. A generic digital model is constructed for each type of analog reference and an optimization routine is developed to adjust it. Regardless of the type of reference system the best practice proved to be measuring every part of

the model that can be measured and use the measurement result directly in the simulation. Afterwards, iterative optimization techniques are used to adjust the remaining parameters of the model. Its output is compared to the recording of the reference device and a set of parameters is iteratively updated, according to the Levenberg–Marquardt method, to reduce the error between digital and analog system using a cost-function.

For black-box optimization problems the cost-function describes the error between the (digital) model and the (analog) reference and is usually the sum of squares of the time-domain error signal. But this metric is not suitable for highly nonlinear systems because it does not describe the perceptual similarity of two signals. Therefore, a time-frequency cost-function is proposed in this work which considers the human perception of frequency and amplitude when calculating the error between the output of the analog and digital system.

The first type of nonlinear system which is analyzed and modeled in this work is called dynamic range compression (DRC) system. It reduces the dynamic range of an audio signal by amplifying small signal levels and reducing larger signal levels according to a nonlinear characteristic. Several dynamic range compression systems are analyzed and modeled. Analog reference devices using different circuits and circuit elements to produce and apply the variable gain are chosen and analyzed like the ‘MXR Dynacomp’, the ‘Aguilar TLC’ or the ‘UREI 1176LN’. The static characteristic of each device is measured and used directly in the model and afterwards the parameters which control the dynamic behavior of the digital model are iteratively optimized in two steps to ensure a robust identification. The adjusted digital models are evaluated with objective error metrics which confirm that the proposed modeling procedure produces good results.

One popular dynamic range compression system, the UREI 1176LN, is modeled in detail. With black-box modeling techniques it is only possible to model the reference device in one particular setting. The controls of the 1176LN are included in the digital model by creating one model for each possible combination of user controls and interpolating between the parameter sets of the digital models according to the position of the knobs on the reference device. The subjective quality of this model is evaluated by a listening test which confirms the good results of the modeling process.

The second category of analyzed nonlinear systems are distortion systems which enrich the input signal with harmonic overtones. These systems are analyzed in varying complexity. Very simple circuits are recreated with basic models, consisting of one linear and one nonlinear block. Then more complex standalone devices are emulated with a Wiener–Hammerstein model which employs two linear and one nonlinear block. Afterwards guitar amplifiers are reproduced with two different models. A cascaded model, using two different nonlinear blocks and three linear blocks is investigated. But the results for highly nonlinear amplifiers are not convincing enough.

The simpler Wiener–Hammerstein model is used which proves to be suited better for gray-box optimization. In both cases the linear blocks are measured first and afterwards the parameters for the nonlinear blocks are iteratively adapted.

In general a multi-step optimization is useful to ensure that the cost-function converges into its global optimum where the error between the model and the reference device has its lowest value. Which cost-function is used depends on the type of reference device and is crucial for the success of the modeling process. The best results for highly nonlinear systems could only be obtained with a psycho-acoustically motivated cost-function considering both time- and frequency-domain as well as the human perception of frequency and amplitudes.

After the optimization is done the quality of the adapted digital model has to be evaluated. This can be done with conventional objective scores like the root-mean-square error (RMS), the error to signal ratio (ESR) or the Pearson correlation coefficient ( $\mathbf{r}_{y\hat{y}}$ ) but these metrics are prone to deviations like small phase-shifts between the signals which have to be compared. This means that a non-perceptible deviation leads to a high error value.

Unfortunately there is no objective metric describing the perceived difference of two signals that suits the needs for virtual analog modeling. To compare the quality of audio codecs, the perceptual evaluation of audio quality (PEAQ) exists, but this method is not suitable for virtual analog modeling quality assessment. Therefore, a score is developed describing the perceptual evaluation of audio similarity (PEAS). This score is based on the time-frequency cost-function which is used to adapt the digital model and is used throughout this work to objectively rate the adapted digital models. These scores show that the quality of a digital model gets worse if the reference system exhibits more nonlinear behavior.

To rate the quality of the adapted model from a subjective point of view a listening test is indispensable. MUSHRA or ABC/HR style listening tests are used in this work. The test subject has to rate the similarity of several test items to a reference item on a scale from 0 to 100. 0 represents ‘annoying differences’ and 100 stands for ‘imperceptible’.

The system identification would be perfect if no test subject could hear a difference between the digital model and the recording from the analog reference device, which is the case for some of the adapted models presented in this work. When the reference system is not highly nonlinear, the differences between the digital model and the reference device are rated as ‘imperceptible’. For highly nonlinear systems the model is rated with ‘minor differences’ on average, which means that a trained listener is able to distinguish between model and reference but the differences are only audible for certain conditions, like e.g. transients or high/low frequency content.

Each guitarist has his or her own unique style of playing the electric guitar. To ensure that a virtual analog model is accepted by many guitar players a test should be performed where different guitarists play their guitar via a real-time virtual analog model of a certain reference device with the possibility to switch instantly between reference and model. If they would feel comfortable with the digital solution the modeling process could be considered successful. Since no real-time model is yet available and due to the logistic complications such a test would implicate it has been left for future research.

---

**List of selected Symbols**

---

---

**Chapter 2: State of the Art**

---

$x(t)$	Continuous-time input signal
$T_n(x)$	Polynomial coefficient for the n-th branch of a multi-branch Hammerstein model
$H(f)$	Filter
$f$	Frequency
$y(t)$	Continuous-time output signal

---

**Chapter 3: Block-Oriented Models**

---

$n$	Discrete time index
$x(n)$	Discrete-time input signal
$y(n)$	Discrete-time output signal
$z^{-1}$	Unit-sample delay
$b_i$	Transfer-function numerator coefficient
$a_j$	Transfer-function denominator coefficient
$H(z)$	Transfer-function
$g(x)$	Nonlinear function for input amplitude $x$
$I_s$	Saturation current of a diode
$n_i$	Ideality factor of a diode
$v$	Thermal voltage
$k$	Boltzmann constant
$q$	Elementary charge
$T$	Temperature in Kelvin

---

 Chapter 3: Block-Oriented Models
 

---

$a$	Gain parameter for nonlinear equations
$f_0$	Fundamental frequency of a sinusoidal signal
$m(x)$	Piecewise defined mapping function based on hyperbolic tangent
$k_p, k_n$	Connection points for piecewise defined function based on hyperbolic tangent
$g_p, g_n$	Gains for piecewise defined function based on hyperbolic tangent
$G_p, G_n$	Gains for piecewise defined function based on hyperbolic tangent in dB
$L$	Oversampling factor
$y(t)$	Continuous-time output signal (approximation)
$\tilde{x}(t)$	Approximated continuous-time input signal
$M(x)$	Antiderivative of $m(x)$
$O(\cdot)$	Order of the approximation error
$f_s$	Sampling frequency
$\delta$	Passband ripple in dB
$A_s$	Stopband attenuation in dB
$f_{\text{pass}}$	Passband edge frequency
$f_{\text{stop}}$	Stopband edge frequency

---

 Chapter 4: System Identification
 

---

$\mathbf{p}$	Parameter vector of a digital model
$y(n)$	Discretized output of a reference system
$\hat{y}(n, \mathbf{p})$	Output of a digital model
$C(\mathbf{p})$	Cost-function describing the error between $y(n)$ and $\hat{y}(n, \mathbf{p})$
$M$	Dimensionality of the optimization problem
$\Delta \mathbf{p}$	Change of $\mathbf{p}$ per iteration
$\mathbf{g}$	Gradient
$\mathbf{H}$	Hessian matrix
$\mathbf{J}$	Jacobian matrix
$\mathbf{r}$	Residual
$\alpha$	Parameter for gradient-descent method
$\mu$	Damping factor for Levenberg–Marquardt method
$\epsilon$	Machine accuracy

## Chapter 4: System Identification

---

$b$	Block index of STFT
$k$	Frequency index of STFT
$N_{\text{FFT}}$	Block size of Fourier transform
$N_{\text{HopSize}}$	Hop size of STFT
$S(\cdot)$	Logarithmic magnitude scaling function
$R(b, k, \mathbf{p})$	STFT residual
$\tilde{k}$	Frequency index of semitone spectrum
$x_{\text{sweep}}(n)$	Discrete-time sine sweep
$A$	Amplitude of $x_{\text{sweep}}(n)$
$f_{\text{start}}, f_{\text{stop}}$	Start and stop frequencies of $x_{\text{sweep}}(n)$
$\omega_1, \omega_2$	Angular start and stop frequencies of $x_{\text{sweep}}(n)$
$L_s$	Length of $x_{\text{sweep}}(n)$ in samples
$f_s$	Sampling frequency
$x_{\text{inv}}(n)$	Inverse sweep
$\Delta n_0$	Initial time shift of measured impulse response
$c$	Amplitude correction factor of measured impulse response
$\hat{h}(n)$	Impulse response of a nonlinear system
$H_c(f)$	Correction filter to reduce effects of measurement equipment
$H_{\text{lb}}(f)$	Frequency response of the measurement equipment
$\gamma(f)$	Frequency-dependent regularization parameter

## Chapter 5: Evaluation: Error Metrics

---

$e_{\text{RMS}}$	Root mean square error
$e_{\text{ESR}}$	Error to signal ratio
$r_{y\hat{y}}$	Pearson correlation coefficient
ODG	Objective difference grade
PEAQ	Perceptual evaluation of audio quality
PEAS	Perceptual evaluation of audio similarity score

---

 Chapter 6: Dynamic Range Compression Systems
 

---

$I_{LED}$	Steady-state current through a light emitting diode
$R_{LDR}$	Resistance of a light dependent resistor
$g$	Gain of a non-inverting amplifier
$g_{inv}$	Gain of an inverting amplifier
$V_{inv}$	Output voltage of an inverting amplifier
$V_+$	Supply voltage
$i_{LED}$	Time-varying current through a light emitting diode
$V_{GS}$	Gate-source voltage of a JFET transistor
$V_{th}$	Threshold voltage of a JFET transistor
$x_{PEAK}(n)$	Output of a peak level detector
$c_{LVL}$	Filter coefficient of peak level detector
$AT\pm$	Value of $c_{LVL}$ in attack case (signal level increasing)
$RT\pm$	Value of $c_{LVL}$ in release case (signal level decreasing)
$c_{si}$	Coefficients for the $i^{th}$ first order low-pass filter of a smoothing filter
$\alpha_1, \alpha_2$	Weighting factors for smoothing filter outputs
$g(n)$	Time-varying gain signal
$G_{PRE}$	Pre-gain
$G_{POST}$	Post-gain
$\mathbf{p}_{comp}$	Parameter vector of the digital model
$x_{sin}(n)$	Multi-frequency sine wave input signal
$a(n)$	Amplitude weighting sequence for $x_{sin}(n)$
$t_c$	Time constant for low-pass filter
$p_i(\theta, a_i)$	Value of the $i^{th}$ model parameter for the user controls $\theta$ and polynomial coefficients $a_i$
$N_e$	Number of optimization routines
$S_j$	Number of subdivisions of the $j^{th}$ user control range
$m_c$	Number of user controls

---

 Chapter 7: Basic Distortion Circuits
 

---

$V_{out}$	Output voltage
$V_{in}$	Input voltage
$f_c$	Cut-off frequency of a filter
$V_{NI}$	Output voltage of a non-inverting amplifier
$\mathbf{p}_{NL}$	Parameter vector for the nonlinear block
$g_{MIX}$	Dry/Wet mixing parameter
$x_{nl}(n)$	Single frequency sine wave input signal
$x_{real}(n)$	Dry recordings of real-world guitar signals

## Chapter 8: Distortion Guitar Effect Pedals

---

$\mathbf{p}_{\text{LTI}}$	Parameter vector for one parametric filter block
$x_{\text{ms}}(n)$	Multi-frequency sine wave with modulated frequencies
$G$	Gain of a filter
$Q$	Q factor of a filter
$f_c$	Center- or Cut-off frequency of a filter

## Chapter 9: Guitar Amplifiers

---

$\mathbf{p}_{\text{H},i}$	Parameter vector for the $i^{\text{th}}$ filter block
$\mathbf{p}_{\text{nl}}$	Parameter vector of the nonlinear block
$x_{\text{ms}}(n)$	Multi-frequency sine wave with modulated frequencies
$p(x)$	polynomial function for input amplitude $x$
$a_n$	$n^{\text{th}}$ polynomial coefficient
$\omega$	Angular frequency $\omega = 2\pi f$
$k_n$	Spectral coefficient for the $(n + 1)^{\text{th}}$ harmonic
$H_i(z)$	Transfer function of the $i^{\text{th}}$ filter
$l_{\text{FIR}}$	Number of FIR filter coefficients
$n_{\text{Sec}}$	Number of second order IIR filter sections



---

## Bibliography

---

- [1] Timothy Stilson and Julius O. Smith III. Analyzing the Moog VCF with considerations for digital implementation. In *Proceedings of the International Computer Music Conference*, Hong Kong, China, 1996.
- [2] Timothy Stilson and Julius O. Smith III. Alias-free digital synthesis of classic analog waveforms. In *Proceedings of the International Computer Music Conference*, Hong Kong, China, 1996.
- [3] Toshinori Araya and Akio Suyama. Sound effector capable of imparting plural sound effects like distortion and other effects, Oct. 1996. US Patent 5,570,424.
- [4] Amar Chaudhary. Band-limited simulation of analog synthesizer modules by additive synthesis. In *Audio Engineering Society Convention 105*, San Francisco, CA, 1998.
- [5] Markus Sapp, Jörg Becker, and Claas Brouër. Simulation of vacuum-tube amplifiers. *Journal of the Acoustical Society of America*, 105(2):1331, 1999.
- [6] Jörn Schattschneider and Udo Zölzer. Discrete-time models for nonlinear audio systems. In *Proceedings of the 2nd International Conference on Digital Audio Effects (DAFx-99)*, pages 45–48, Trondheim, Norway, 1999.
- [7] Dale Vernon Curtis, Keith Lance Chapman, and Charles Clifford Adams. Simulated tone stack for electric guitar, Apr. 2001. US Patent 6,222,110.
- [8] Jiri Schimmel. Non-linear dynamics processing. In *Audio Engineering Society Convention 114*, Amsterdam, The Netherlands, 2003. Audio Engineering Society.
- [9] Antti Huovilainen. Non-linear digital implementation of the Moog ladder filter. In *Proceedings of the 7th International Conference on Digital Audio Effects (DAFx-04)*, pages 61–64, Naples, Italy, 2004.
- [10] Vesa Välimäki. Discrete-time synthesis of the sawtooth waveform with reduced aliasing. *IEEE Signal Processing Letters*, 12(3):214–217, 2005.
- [11] Antti Huovilainen. Enhanced digital models for digital modulation effects. In *Proceedings of the 8th International Conference on Digital Audio Effects (DAFx-05)*, Madrid, Spain, 2005.

- [12] Antti Huovilainen and Vesa Välimäki. New approaches to digital subtractive synthesis. In *Proceedings of the International Computer Music Conference*, pages 399–402, Barcelona, Spain, 2005.
- [13] Timothy Stilson. *Efficiently-variable non-oversampled algorithms in virtual-analog music synthesis: A root-locus perspective*. Dissertation, Stanford University, Stanford, 2005.
- [14] David Te-Mao Yeh and Julius O. Smith III. Discretization of the ‘59 fender bassman tone stack. In *Proceedings of the 9th International Conference on Digital Audio Effects (DAFx-06)*, pages 18–20, Montreal, Quebec, Canada, 2006.
- [15] Matti Karjalainen and Jyri Pakarinen. Wave digital simulation of a vacuum-tube amplifier. In *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, volume 5, pages V 153 – V 156. IEEE, 2006.
- [16] Thomas Hélie. On the use of Volterra series for real-time simulations of weakly nonlinear analog audio devices: Application to the Moog ladder filter. In *Proceedings of the 9th International Conference on Digital Audio Effects (DAFx-06)*, pages 7–12, Montreal, Quebec, Canada, 2006.
- [17] Vesa Välimäki and Antti Huovilainen. Oscillator and filter algorithms for virtual analog synthesis. *Computer Music Journal*, 30(2):19–31, 2006.
- [18] František Kadlec and Antonín Novák. Comparison of measurements method for analysis electro-acoustic systems with non-linearities. In *Proceedings of the 19th International Congress on Acoustics*, Madrid, Spain, Sep. 2007.
- [19] Antonín Novák. Identification of nonlinearity of electro-acoustic systems using a direct path miso method. In *Proceedings of the 19th International Congress on Acoustics*, Madrid, Spain, Sep. 2007.
- [20] Antonín Novák. Identification of nonlinear systems: Volterra series simplification. *Acta Polytechnica*, 47(4-5), 2007.
- [21] Vesa Välimäki and Antti Huovilainen. Antialiasing oscillators in subtractive synthesis. *IEEE Signal Processing Magazine*, 24(2):116–125, 2007.
- [22] David Te-Mao Yeh, Jonathan S. Abel, and Julius O. Smith III. Simulation of the diode limiter in guitar distortion circuits by numerical solution of ordinary differential equations. In *Proceedings of the 10th International Conference on Digital Audio Effects (DAFx-07)*, pages 197–204, Bordeaux, France, 2007.
- [23] Federico Fontana. Preserving the structure of the Moog VCF in the digital domain. In *Proceedings of the International Computer Music Conference*, Copenhagen, Denmark, 2007.
- [24] David Te-Mao Yeh and Julius O. Smith III. Simulating guitar distortion circuits using wave digital and nonlinear state-space formulations. In *Proceedings of the 11th International Conference on Digital Audio Effects (DAFx-08)*, pages 19–26, Espoo, Finland, 2008.
- [25] David Te-Mao Yeh, Jonathan S. Abel, Andrei Vladimirescu, and Julius O. Smith III. Numerical methods for simulation of guitar distortion circuits. *Computer Music Journal*, 32(2):23–42, 2008.
- [26] Jyri Pakarinen. *Modeling of nonlinear and time-varying phenomena in the guitar*. Dissertation, Helsinki University of Technology, 2008-03-04.

- [27] Giovanni De Sanctis and Augusto Sarti. Virtual analog modeling in the wave-digital domain. *IEEE transactions on audio, speech, and language processing*, 18(4):715–727, 2009.
- [28] Joseph Timoney, Victor Lazzarini, Brian Carty, and Jussi Pekonen. Phase and amplitude distortion methods for digital synthesis of classic analog waveforms. In *Audio Engineering Society Convention 126*, Munich, Germany, 2009. Audio Engineering Society.
- [29] Stefan Bilbao and Julian D. Parker. A virtual model of spring reverberation. *IEEE transactions on audio, speech, and language processing*, 18(4):799–808, 2009.
- [30] Juhan Nam, Vesa Välimäki, Jonathan S. Abel, and Julius O. Smith III. Alias-free virtual analog oscillators using a feedback delay loop. In *Proceedings of the 12th International Conference on Digital Audio Effects (DAFx-09)*, pages 1–6, Como, Italy, 2009.
- [31] Juhan Nam, Vesa Välimäki, Jonathan S. Abel, and Julius O. Smith III. Efficient antialiasing oscillator algorithms using low-order fractional delay filters. *IEEE transactions on audio, speech, and language processing*, 18(4):773–785, 2009.
- [32] Jussi Pekonen, Vesa Välimäki, Juhan Nam, Jonathan S. Abel, and Julius O. Smith III. Variable fractional delay filters in bandlimited oscillator algorithms for music synthesis. In *the 2010 International Conference on Green Circuits and Systems (ICGCS2010)*, Shanghai, China, Jun. 2010.
- [33] Glen Deslauriers and Colby Leider. A bandlimited oscillator by frequency-domain synthesis for virtual analog applications. In *Audio Engineering Society Convention 127*, New York, NY, 2009.
- [34] Thomas Hélie. Volterra series and state transformation for real-time simulations of audio circuits including saturations: Application to the moog ladder filter. *IEEE transactions on audio, speech, and language processing*, 18(4):747–759, 2009.
- [35] Antonín Novák, Laurent Simon, František Kadlec, and Pierrick Lotton. Nonlinear system identification using exponential swept-sine signal. *IEEE Transactions on Instrumentation and Measurement*, 59(8):2220–2229, 2009.
- [36] Antonín Novák, Laurent Simon, Pierrick Lotton, and František Kadlec. Modeling of nonlinear audio systems using swept-sine signals: Application to audio effects. In *Proceedings of the 12th International Conference on Digital Audio Effects (DAFx-09)*, Como, Italy, Sep. 2009.
- [37] Antonín Novák, Laurent Simon, and Pierrick Lotton. Analysis, synthesis, and classification of nonlinear systems using synchronized swept-sine method for audio effects. *EURASIP Journal on Advances in Signal Processing*, 2010(1):793816, 2010.
- [38] Vesa Välimäki, Julian D. Parker, and Jonathan S. Abel. Parametric spring reverberation effect. *Journal of the Audio Engineering Society*, 58(7/8):547–562, 2010.
- [39] Victor Lazzarini and Joseph Timoney. New perspectives on distortion synthesis for virtual analog oscillators. *Computer Music Journal*, 34(1):28–40, 2010.

- [40] Frederico Fontana and Marco Civolani. Modeling of the EMS VCS3 voltage-controlled filter as a nonlinear filter network. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(4):760–772, May 2010.
- [41] Jussi Pekonen and Vesa Välimäki. The brief history of virtual analog synthesis. In *Proceedings of the 6th Forum Acusticum*, pages 461–466, Aalborg, Denmark, 2011. European Acoustics Association.
- [42] Kristjan Dempwolf, Martin Holters, and Udo Zölzer. Discretization of parametric analog circuits for real-time simulations. In *Proceedings of the 13th international conference on digital audio effects (DAFx-10)*, Graz, Austria, 2010.
- [43] Kristjan Dempwolf, Martin Holters, and Udo Zölzer. A triode model for guitar amplifier simulation with individual parameter fitting. In *Audio Engineering Society Convention 131*, New York, NY, Oct. 2011.
- [44] Jaromír Mačák and Jiri Schimmel. Simulation of a vacuum-tube push-pull guitar power amplifier. Paris, France, Sep. 2011.
- [45] Martin Holters and Udo Zölzer. Physical modelling of a wah-wah effect pedal as a case study for application of the nodal DK method to circuits with variable parts. In *Proceedings of the 14th international conference on Digital Audio Effects (DAFx-11)*, pages 31–35, Paris, France, 2011.
- [46] Oliver Kröning, Kristjan Dempwolf, and Udo Zölzer. Analysis and simulation of an analog guitar compressor. In *Proceedings of the 14th International Conference on Digital Audio Effects (DAFx-11)*, pages 205–208, Paris, France, Sep. 2011.
- [47] Marco Fink and Rudolf Rabenstein. A Csound opcode for a triode stage of a vacuum tube amplifier. In *Proceedings of the 14th international conference on Digital Audio Effects (DAFx-11)*, pages 365–370, Paris, France, 2011.
- [48] Joseph Timoney and Victor Lazzarini. Exponential frequency modulation bandwidth criterion for virtual analog applications. In *Proceedings of the 14th International Conference on Digital Audio Effects (DAFx-11)*, Sep. 2011.
- [49] Jyri Pakarinen, Vesa Välimäki, Federico Fontana, Victor Lazzarini, and Jonathan S. Abel. Recent advances in real-time musical effects, synthesis, and virtual analog models. *EURASIP Journal on Advances in Signal Processing*, 2011(1):940784, 2011.
- [50] Julian D. Parker. A simple digital model of the diode-based ring-modulator. In *Proceedings of the 14th International Conference on Digital Audio Effects (DAFx-11)*, volume 14, pages 163–166, Paris, France, 2011.
- [51] Sami Oksanen and Vesa Välimäki. Modeling of the carbon microphone nonlinearity for a vintage telephone sound effect. In *Proceedings of the 14th International Conference on Digital Audio Effects (DAFx-11)*, pages 27–30, Paris, France, 2011.
- [52] Sami Oksanen and Vesa Välimäki. Digital modeling of the vintage telephone sound. In *Proceedings of the International Computer Music Conference*, pages 636 – 639, Huddersfield, United Kingdom, Aug. 2011.
- [53] Rafael Cauduro Dias de Paiva, Stefano D’Angelo, Jyri Pakarinen, and Vesa Valimaki. Emulation of operational amplifiers and diodes in audio distortion circuits. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 59(10):688–692, 2012.

- [54] Fredrik Gustaffsson, Per Connman, Oscar Öberg, Niklas Odelholm, and Martin Enqvist. System and method for simulation of non-linear audio equipment, Apr. 2012. US Patent 8,165,309.
- [55] Jaromír Mačák, Jiri Schimmel, and Martin Holters. Simulation of Fender type guitar preamp using approximation and state space model. In *Proceedings of the 10th International Conference on Digital Audio Effects (DAFx-12)*, York, United Kingdom, Sep. 2012.
- [56] François G. Germain, Jonathan S. Abel, Philippe Depalle, and Marcelo M. Wanderley. Uniform noise sequences for nonlinear system identification. In *Proceedings of the 15th International Conference on Digital Audio Effects (DAFx-12)*, pages 241–244, York, United Kingdom, Sep. 2012.
- [57] Thomaz Chaves de A. Oliveira, Gilmar Barreto, and Alexander Mattioli Pasqual. Review of digital emulation of vacuum-tube audio amplifiers and recent advances in related virtual analog models. *INFOCOMP*, 12(1):10–23, 2013.
- [58] Julian D. Parker and Stephano D’Angelo. A digital model of the Buchla lowpass-gate. In *Proceedings of the 16th International Conference on Digital Audio Effects (DAFx-13)*, pages 278–285, Maynooth, Ireland, 2013.
- [59] Antoine Falaize and Thomas Hélie. Simulation of an analog circuit of a wah pedal: a port-hamiltonian approach. In *Audio Engineering Society Convention 135*. Audio Engineering Society, 2013.
- [60] Felix Eichas, Marco Fink, Martin Holters, and Udo Zölzer. Physical modeling of the MXR Phase 90 guitar effect pedal. In *Proceedings of the 17th International Conference on Digital Audio Effects (DAFx-14)*, pages 153–158, Erlangen, Germany, 2014.
- [61] Kurt James Werner, Jonathan S. Abel, and Julius O. Smith III. More Cowbell: a Physically-Informed, Circuit-Bendable, Digital Model of the TR-808 Cowbell. In *137th International Audio Engineering Society (AES) Conference*, Los Angeles, CA, Oct. 2014.
- [62] Kurt James Werner, Jonathan S. Abel, and Julius O. Smith III. A physically-informed, circuit-bendable, digital model of the Roland TR-808 bass drum circuit. In *Proceedings of the 17th International Conference on Digital Audio Effects (DAFx-14)*, Erlangen, Germany, 2014.
- [63] Kurt James Werner, Jonathan S. Abel, and Julius O. Smith III. The TR-808 Cymbal: a Physically-Informed, Circuit-Bendable, Digital Model. In *40th International Computer Music Conference (ICMC) / 11th Sound and Music Computing Conference (SMC)*, Athens, Greece, Sep. 2014.
- [64] Antonín Novák, Pierrick Lotton, and Laurent Simon. Synchronized swept-sine: Theory, application, and implementation. *Journal of the Audio Engineering Society*, 63(10):786–798, 2015.
- [65] Martin Holters and Udo Zölzer. A generalized method for the derivation of non-linear state-space models from circuit schematics. In *Proceedings of the 23rd European Signal Processing Conference (EUSIPCO)*, pages 1073–1077, Nice, France, 2015. IEEE.
- [66] Antonín Novák, Laurent Simon, and Pierrick Lotton. Extension of generalized hammerstein model to non-polynomial inputs. In *Proceedings of the 24th European Signal Processing Conference (EUSIPCO)*, pages 21–25. IEEE, 2016.

- [67] Felix Eichas and Udo Zölzer. Modeling of an optocoupler-based audio dynamic range control circuit. In *SPIE Optical Engineering+ Applications*, San Diego (CA), U.S.A., Aug. 2016. International Society for Optics and Photonics.
- [68] Roope Kiiski, Fabián Esqueda, and Vesa Välimäki. Time-variant gray-box modeling of a phaser pedal. In *Proceedings of the 19th International Conference on Digital Audio Effects (DAFx-16)*, Brno, Czech Republic, 2016.
- [69] Ben Holmes and Maarten Van Walstijn. Physical model parameter optimisation for calibrated emulation of the Dallas Rangemaster treble booster guitar pedal. pages 47–54, Sep. 2016.
- [70] Antonín Novák, Léo Guadagnin, Bertrand Lihoreau, Pierrick Lotton, and Laurent Simon. Non-linear identification of an electric guitar pickup. In *Proceedings of the 19th International Conference on Digital Audio Effects (DAFx-16)*, Brno, Czech Republic, Sep. 2016.
- [71] Martin Holters and Udo Zölzer. Circuit simulation with inductors and transformers based on the Jiles-Atherton model of magnetization. In *Proceedings of the 19th International Conference on Digital Audio Effects (DAFx-16)*, pages 55–60, Brno, Czech Republic, Sep. 2016.
- [72] Kurt James Werner, W. Ross Dunkel, and François G. Germain. A computational model of the Hammond organ vibrato/chorus using wave digital filters. In *Proceedings of the 19th International Conference on Digital Audio Effects (DAFx-16)*, pages 271–278, Brno, Czech Republic, Sep. 2016.
- [73] Damien Bouvier, Thomas Hélie, and David Roze. Nonlinear homogeneous order separation for Volterra series identification. pages 3–10, Edinburgh, UK, Sep. 2017.
- [74] Fabián Esqueda, Henri Pöntynen, Vesa Välimäki, and Julian D. Parker. Virtual analog Buchla 259 wavefolder. In *Proceedings of the 20th International Conference on Digital Audio Effects (DAFx-17)*, pages 192–199, Edinburgh, UK, Sep. 2017.
- [75] Maximilian Rest, Julian D. Parker, and Kurt James Werner. WDF modeling of a Korg MS-50 based non-linear diode bridge VCF. In *Proceedings of the 20th International Conference on Digital Audio Effects (DAFx-17)*, pages 145–151, Edinburgh, United Kingdom, Sep. 2017.
- [76] Martin Holters and Julian D. Parker. A combined model for a bucket brigade device and its input and output filters. In *Proceedings of the 21st International Conference on Digital Audio Effects (DAFx-18)*, Aveiro, Portugal, Sep. 2018.
- [77] Stefano D’Ángelo. Lightweight virtual analog modeling. In *XXII Colloquio di Informatica Musicale 22nd Colloquium on Music Informatics Atti della Conferenza Conference Proceedings*, page 59, 2018.
- [78] Thomas Schmitz. *Nonlinear modeling of the guitar signal chain enabling its real-time emulation*. Dissertation, University of Liège, Liège, Belgium, 2019.
- [79] Adalberto Schuck, Luiz F. Ferreira, Ronaldo Husemann, and Bardo EJ Bodmann. A simple non-linear transfer function for a Wiener – Hammerstein model to simulate guitar distortion and overdrive effects. In *Integral Methods in Science and Engineering*, pages 425–436. Springer, 2019.

- [80] Rafael Cauduro Dias de Paiva, Jyri Pakarinen, Vesa Välimäki, and Miikka Tikander. Real-time audio transformer emulation for virtual tube amplifiers. *EURASIP Journal on Advances in Signal Processing*, 2011(1):347645, 2011.
- [81] Stefano D'Angelo, Jyri Pakarinen, and Vesa Välimäki. New family of wave-digital triode models. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(2):313–321, 2013.
- [82] Kurt James Werner, Vaibhav Nangia, Julius O. Smith III, and Jonathan S. Abel. Resolving wave digital filters with multiple/multiport nonlinearities. In *Proceedings of the 18th International Conference on Digital Audio Effects (DAFx-15)*, pages 387–394, Trondheim, Norway, Nov. 2015.
- [83] Kurt James Werner, Julius O. Smith III, and Jonathan S. Abel. Wave digital filter adaptors for arbitrary topologies and multiport linear elements. In *Proceedings of the 18th International Conference on Digital Audio Effects (DAFx-15)*, pages 379–386, Trondheim, Norway, Nov. 2015.
- [84] W. Ross Dunkel, Maximilian Rest, Kurt James Werner, Michael J. Olsen, and Julius O. Smith III. The Fender Bassman 5F6-A family of preamplifier circuits—a wave digital filter case study. In *Proceedings of the 19th International Conference on Digital Audio Effects (DAFx-16)*, Brno, Czech Republic, Sep. 2016.
- [85] Ben Holmes, Martin Holters, and Maarten van Walstijn. Comparison of germanium bipolar junction transistor models for real-time circuit simulation. In *Proceedings of the 20th International Conference on Digital Audio Effects (DAFx-17)*, pages 152–159, Edinburgh, UK, Sep. 2017.
- [86] Karl-Johan Åström and Bohlin Torsten. Numerical identification of linear dynamic systems from normal operating records. *IFAC Proceedings Volumes*, 2(2):96–111, 1965.
- [87] Karl Johan Åström and Peter Eykhoff. System identification – a survey. *Automatica*, 7(2):123–162, 1971.
- [88] Stephen A. Billings. Identification of nonlinear systems – a survey. In *IEE Proceedings D (Control Theory and Applications)*, volume 127, pages 272–285. IET, 1980.
- [89] Jonas Sjöberg, Qinghua Zhang, Lennart Ljung, Albert Benveniste, Bernard Delyon, Pierre Yves Glorennec, Håkan Hjalmarsson, and Anatoli Juditsky. Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, 31(12):1691–1724, 1995.
- [90] Lennart Ljung and Keith Glover. Frequency domain versus time domain methods in system identification. *Automatica*, 17(1):71–86, 1981.
- [91] Lennart Ljung. System identification. *Wiley Encyclopedia of Electrical and Electronics Engineering*, 2001.
- [92] Lennart Ljung. *Nonlinear black-box modeling in system identification*. Linköping University, 1995.
- [93] Torsten Söderström and Petre Stoica. *System identification*. Prentice-Hall, Inc., 1988.
- [94] Ronald K. Pearson and Martin Pottmann. Gray-box identification of block-oriented nonlinear models. *Journal of Process Control*, 10(4):301–315, 2000.

- [95] Emil Cauer, Wolfgang Mathis, and Rainer Pauli. Life and work of Wilhelm Cauer (1900 – 1945). In *Proceedings of the Fourteenth International Symposium of Mathematical Theory of Networks and Systems (MTNS2000)*, Jun. 2000.
- [96] T.P. Bohlin. *Practical Grey-box Process Identification: Theory and Applications*. Advances in Industrial Control. Springer London, 2006.
- [97] Felix Eichas, Marco Fink, and Udo Zölzer. Feature design for the classification of audio effect units by input/output measurements. In *Proceedings of the 18th International Conference on Digital Audio Effects (DAFx-15)*, pages 27–33, Trondheim, Norway, Nov. 2015.
- [98] Udo Zölzer. *Digital Audio Effects (DAFx)*. John Wiley and Sons Ltd., Chichester, 2. edition, 2011.
- [99] Guy W. McNally. Dynamic range control of digital audio signals. *Journal of the Audio Engineering Society*, 32(5):316–327, 1984.
- [100] Eric Barbour. The cool sound of tubes [vacuum tube musical applications]. *IEEE Spectrum*, 35(8):24–35, 1998.
- [101] W. Stephen Bussey and Robert Haigler. Tubes versus transistors in electric guitar amplifiers. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '81*, pages 800–803, Atlanta, Georgia, USA, Apr. 1981.
- [102] David Te-Mao Yeh. *Digital implementation of musical distortion circuits by analysis and simulation*. Dissertation, Stanford University, Stanford, 2009.
- [103] Kristjan Dempwolf. *Modellierung analoger Gitarrenverstärker mit digitaler Signalverarbeitung (Modeling of analog guitar amplifiers with digital signal processing)*. Dissertation, Helmut Schmidt University, Hamburg, 2012.
- [104] Jaromír Mačák. *Real-time digital simulation of guitar amplifiers as audio effects*. Dissertation, Brno University of Technology, Brno, 2012.
- [105] Martin Holters and Udo Zölzer. Automatic decomposition of non-linear equation systems in audio effect circuit simulation. In *Proceedings of the 20th International Conference on Digital Audio Effects (DAFx-17)*, pages 138–144, Edinburgh, UK, 2017.
- [106] Jaromír Mačák and Jiri Schimmel. Real-time guitar tube amplifier simulation using an approximation of differential equations. In *Proceedings of the 13th International Conference on Digital Audio Effects (DAFx-10)*, Graz, Austria, Sep. 2010.
- [107] Alfred Fettweis. Wave digital filters: Theory and practice. *Proceedings of the IEEE*, 74(2):270–327, 1986.
- [108] Rafael Cauduro Dias de Paiva. *Circuit modeling studies related to guitars and audio processing*. Dissertation, Aalto University, Helsinki, 2013.
- [109] Stefano D’Angelo. *Virtual analog modeling of nonlinear musical circuits*. Dissertation, Aalto University, Helsinki, 2014.
- [110] Kurt James Werner. *Virtual Analog Modeling of Audio Circuitry Using Wave Digital Filters*. Dissertation, Stanford University, Stanford, 2016.
- [111] Stefano D’Angelo and Vesa Välimäki. Wave-digital polarity and current inverters and their application to virtual analog audio processing. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 469–472. IEEE, 2012.

- [112] Antoine Falaize, Nicolas Lopes, Thomas Hélie, Denis Matignon, and Bernhard Maschke. Energy-balanced models for acoustic and audio systems: A port-Hamiltonian approach. In *Unfold Mechanics for Sounds and Music*, pages 1–1, Sep. 2014.
- [113] Antoine Falaize and Thomas Hélie. Passive guaranteed simulation of analog audio circuits: A port-hamiltonian approach. *Applied Sciences*, 6(10):273, 2016.
- [114] Antoine Falaize and Thomas Hélie. Passive simulation of electrodynamic loudspeakers for guitar amplifiers: a port-hamiltonian approach. In *International Symposium on Musical Acoustics*, pages 1–5, 2014.
- [115] Chet Gnegy and Kurt James Werner. Digitizing the Ibanez Weeping Demon wah pedal. In *Proceedings of the 18th International Conference on Digital Audio Effects (DAFx-15)*, Trondheim, Norway, Nov. 2015.
- [116] Ben Holmes and Maarten van Walstijn. Improving the robustness of the iterative solver in state-space modelling of guitar distortion circuitry. In *Proceedings of the 18th International Conference on Digital Audio Effects (DAFx-15)*, Trondheim, Norway, Nov. 2015.
- [117] Jyri Pakarinen and Matti Karjalainen. Enhanced wave digital triode model for real-time tube amplifier emulation. *IEEE Transactions on audio, speech and language Processing*, 18(4):738–746, 2010.
- [118] Antonín Novák. *Identification of nonlinear systems in acoustics*. Dissertation, Université du Maine, Le Mans, 2009.
- [119] Antonín Novák, Laurent Simon, Pierrick Lotton, and Joël Gilbert. Chebyshev model and synchronized swept sine method in nonlinear audio effect modeling. In *Proceedings of the 13th International Conference on Digital Audio Effects (DAFx-10)*, 2010.
- [120] Angelo Farina. Advancements in impulse response measurements by sine sweeps. In *Audio Engineering Society Convention 122*, Vienna, Austria, 2007. Audio Engineering Society.
- [121] Rafael Cauduro Dias de Paiva, Jyri Pakarinen, and Vesa Välimäki. Reduced-complexity modeling of high-order nonlinear audio systems using swept-sine and principal component analysis. In *Audio Engineering Society Conference: 45th International Conference: Applications of Time-Frequency Processing in Audio*. Audio Engineering Society, 2012.
- [122] François G. Germain. A nonlinear analysis framework for electronic synthesizer effects. Master’s thesis, McGill University, Montreal, Canada, 2011.
- [123] Adalberto Schuck Jr and Bardo Ernst Josef Bodmann. Audio nonlinear modeling through hyperbolic tangent functionals. In *Proceedings of the 19th International Conference on Digital Audio Effects (DAFx-16)*, Brno, Czech Republic, Sep. 2016.
- [124] Andrea Primavera, Michele Gasparini, Stefania Cecchi, Wataru Hariya, Shogo Murai, Koji Oishi, and Francesco Piazza. A novel measurement procedure for Wiener/Hammerstein classification of nonlinear audio systems. In *Audio Engineering Society Convention 144*. Audio Engineering Society, 2018.
- [125] Maarten Schoukens and Koen Tiels. Identification of block-oriented nonlinear systems starting from linear approximations: A survey. *Automatica*, 85:272–292, 2017.

- [126] Felix Eichas, Stephan Möller, and Udo Zölzer. Block-oriented modeling of distortion audio effects using iterative minimization. In *Proceedings of the 18th International Conference on Digital Audio Effects (DAFx-15)*, pages 243–248, Trondheim, Norway, Nov. 2015.
- [127] Felix Eichas and Udo Zölzer. Black-box modeling of distortion circuits with block-oriented models. In *Proceedings of the 19th International Conference on Digital Audio Effects (DAFx-16)*, pages 39–45, Brno, Czech Republic, 2016.
- [128] Felix Eichas and Udo Zölzer. Block-oriented gray box modeling of guitar amplifiers. In *Proceedings of the 20th International Conference on Digital Audio Effects (DAFx-17)*, pages 184–191, Edinburgh, United Kingdom, 2017.
- [129] Fractal Audio Systems. Multipoint iterative matching and impedance correction technology (MIMIC™). Technical report, Fractal Audio Systems, Apr. 2013.
- [130] Christoph Kemper. Musical instrument with acoustic transducer, Jun. 2008. US Patent App. 11/881,818.
- [131] Georg Holzmann. Reservoir computing: a powerful black-box framework for nonlinear audio processing. In *International Conference on Digital Audio Effects (DAFx)*, 2009.
- [132] John Covert and David L. Livingston. A vacuum-tube guitar amplifier model using a recurrent neural network. In *2013 Proceedings of IEEE Southeastcon*, pages 1–5. IEEE, 2013.
- [133] Zhichen Zhang, Edward Olbrych, Joseph Bruchalski, Thomas J. McCormick, and David L. Livingston. A vacuum-tube guitar amplifier model using long/short-term memory networks. In *SoutheastCon 2018*, pages 1–5. IEEE, 2018.
- [134] Thomas Schmitz and Jean-Jacques Embrechts. Real time emulation of parametric guitar tube amplifier with long short term memory neural network. In *Proceedings of the 4th International Conference on Image Processing and Pattern Recognition (IPPR 2018)*, Copenhagen, Denmark, Apr. 2018.
- [135] Thomas Schmitz and Jean-Jacques Embrechts. Nonlinear real-time emulation of a tube amplifier with a long short time memory neural-network. In *Audio Engineering Society Convention 144*. Audio Engineering Society, 2018.
- [136] Alec Wright, Eero-Pekka Damskågg, and Vesa Välimäki. Real-time black-box modelling with recurrent neural networks. In *Proceedings of the 22nd International Conference on Digital Audio Effects (DAFx-19)*, Birmingham, United Kingdom, Sep. 2019.
- [137] Aaron v.d. Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu. Wavenet: A generative model for raw audio. *arXiv preprint arXiv:1609.03499*, Sep. 2016.
- [138] Eero-Pekka Damskågg, Lauri Juvela, Etienne Thuillier, and Vesa Välimäki. Deep learning for tube amplifier emulation. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 471–475. IEEE, 2019.

- [139] Eero-Pekka Damskögg, Lauri Juvela, and Vesa Välimäki. Real-time modeling of audio distortion circuits with deep learning. In *Proceedings of the 16th Sound & Music Computing Conference SMC 2019*, pages 332 – 339, 2019.
- [140] Marco A Martínez Ramírez and Joshua D. Reiss. Modeling nonlinear audio effects with end-to-end deep neural networks. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 171–175. IEEE, 2019.
- [141] Martin Holters and Udo Zölzer. Graphic equalizer design using higher-order recursive filters. In *Proceedings of the 9th International Conference on Digital Audio Effects (DAFx-06)*, pages 37–40, 2006.
- [142] Balázs Bank, Jose A. Belloch, and Vesa Välimäki. Efficient design of a parallel graphic equalizer. *Journal of the Audio Engineering Society*, 65(10):817–825, 2017.
- [143] Vesa Välimäki and Juho Liski. Accurate cascade graphic equalizer. *IEEE Signal Processing Letters*, 24(2):176–180, 2017.
- [144] Robert Bristow-Johnson. The equivalence of various methods of computing biquad coefficients for audio parametric equalizers. In *Audio Engineering Society Convention 97*. Audio Engineering Society, 1994.
- [145] Ralf Kories and Heinz Schmidt-Walter. *Taschenbuch der Elektrotechnik: Grundlagen und Elektronik*. Harri Deutsch, 2006.
- [146] Julen Kahles, Fabián Esqueda, and Vesa Välimäki. Oversampling for nonlinear waveshaping: Choosing the right filters. *Journal of the Audio Engineering Society*, 67(6):440–449, 2019.
- [147] Julian D Parker, Vadim Zavalishin, and Efflam Le Bivic. Reducing the aliasing of nonlinear waveshaping using continuous-time convolution. In *Proceedings of the 19th International Conference on Digital Audio Effects (DAFx-16)*, pages 137–144, Brno, Czech Republic, Sep. 2016.
- [148] Stefan Bilbao, Fabián Esqueda, Julian D Parker, and Vesa Välimäki. Antiderivative antialiasing for memoryless nonlinearities. *IEEE Signal Processing Letters*, 24(7):1049–1053, 2017.
- [149] Robert Haber. Structural identification of quadratic block-oriented models based on estimated Volterra kernels. *International journal of systems science*, 20(8):1355–1380, 1989.
- [150] Alain Y. Kibangou and Gérard Favier. Wiener – Hammerstein systems modeling using diagonal Volterra kernels coefficients. *IEEE signal processing letters*, 13(6):381–384, 2006.
- [151] František Kadlec, Pierrick Lotton, Antonín Novák, and Laurent Simon. A new method for identification of nonlinear systems using miso model with swept-sine technique: Application to loudspeaker analysis. In *Audio Engineering Society Convention 124*. Audio Engineering Society, 2008.
- [152] Jari Turunen, Pekka Loula, and Juha Tantt. Block-oriented-model in speech coding. 11 2019.
- [153] Tilo Strutz. *Data Fitting and Uncertainty: A Practical Introduction to Weighted Least Squares and Beyond*. Vieweg and Teubner, Germany, 2010.
- [154] Eric Walter and Luc Pronzato. *Identification of parametric models from experimental data*. Springer Verlag, 1997.

- [155] Kenneth Levenberg. A method for the solution of certain problems in least squares. *Quarterly of applied mathematics*, 2:164–168, 1944.
- [156] Donald W. Marquardt. An algorithm for least-squares estimation of non-linear parameters. *Journal of the Society for Industrial & Applied Mathematics*, 11(2):431–441, 1963.
- [157] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C (2nd Ed.): The Art of Scientific Computing*. Cambridge University Press, New York, NY, USA, 1992.
- [158] Manfred Zollner. *Physik der Elektrogitarre (physics of the electric guitar)*. Zollner, 2014.
- [159] Eberhard Zwicker and Hugo Fastl. *Psychoacoustics: Facts and models*, volume 22. Springer Science & Business Media, 2013.
- [160] Martin Holters, Tobias Corbach, and Udo Zölzer. Impulse response measurement techniques and their applicability in the real world. In *Proceedings of the 12th International Conference on Digital Audio Effects (DAFx-09)*, Como, Italy, Sep. 2009.
- [161] Ole Kirkeby, Philip A. Nelson, and Hareo Hamada. The ‘stereo dipole’: A virtual source imaging system using two closely spaced loudspeakers. *Journal of the Audio Engineering Society*, 46(5):387–395, 1998.
- [162] RME Intellignet Audio Solutions. *Fireface UC – User’s Guide*, Mar. 2016. Version 1.9.
- [163] Manfred Zollner. Die Dummy-Load als Lautsprecher-Ersatz (the dummy-load as speaker replacement). In *GITEC Forum*, 2016.
- [164] ITU Radiocommunication Assembly. ITU-R BS. 1387-1: Method for objective measurements of perceived audio quality, 2015.
- [165] Thilo Thiede, William C Treurniet, Roland Bitto, Christian Schmidmer, Thomas Sporer, John G Beerends, and Catherine Colomes. PEAQ-The ITU standard for objective measurement of perceived audio quality. *Journal of the Audio Engineering Society*, 48(1/2):3–29, 2000.
- [166] Peter Kabal. An examination and interpretation of ITU-R BS. 1387: Perceptual evaluation of audio quality. *TSP Lab Technical Report, Dept. Electrical & Computer Engineering, McGill University*, pages 1–89, 2002.
- [167] ITU Radiocommunication Assembly. ITU-R BS. 1116-3: Methods for the subjective assessment of small impairments in audio systems, 2015.
- [168] ITU Radiocommunication Assembly. ITU-R BS. 1534-3: Method for the subjective assessment of intermediate quality level of audio systems, 1998–2001.
- [169] Dimitrios Giannoulis, Michael Massberg, and Joshua D. Reiss. Digital dynamic range compressor design – a tutorial and analysis. *Journal of the Audio Engineering Society*, 60(6):399–408, 2012.
- [170] John Hollis. Assorted circuit designs - <http://www.hollis.co.uk/john/circuits.html>, 2001.
- [171] Perkin Elmer. Photoconductive cells and analog optoisolators. Technical report, Perkin Elmer, 2001.
- [172] KEITHLEY. *Series 2600 System SourceMeter ® Manual*, 2008.

- [173] Etienne Gerat. Virtual analog modeling of a UREI 1176LN analog dynamic range control system. Master's thesis, Helmut Schmidt University, Hamburg, Germany, 2016.
- [174] UREI. *Model 1176LN Peak Limiter User Manual*, 2009.
- [175] Etienne Gerat, Felix Eichas, and Udo Zölzer. Virtual analog modeling of a UREI 1176LN dynamic range control system. In *Audio Engineering Society Convention 143*. Audio Engineering Society, 2017.
- [176] Manfred Schroeder. Synthesis of low-peak-factor signals and binary sequences with low autocorrelation (corresp.). *IEEE Transactions on Information Theory*, 16(1):85–89, 1970.
- [177] Udo Zölzer. *Digitale Audiosignalverarbeitung*. B.G. Teubner, Stuttgart, 3., überarbeitete und erweiterte edition, 2005.
- [178] Sebastian Kraft and Udo Zölzer. BeagleJS: HTML5 and JavaScript based framework for the subjective evaluation of audio quality. In *Linux Audio Conference, Karlsruhe, DE*, 2014.
- [179] Kit Rae. Big muff pi page - <http://www.bigmuffpage.com/>, 2012.
- [180] R. G. Keen. The technology of the tube screamer - [http://www.geofex.com/article\\_folders/tstech/tsxtech.htm](http://www.geofex.com/article_folders/tstech/tsxtech.htm), 1998.
- [181] Jyri Pakarinen and David Te-Mao Yeh. A review of digital techniques for modeling vacuum-tube guitar amplifiers. *Computer Music Journal*, 33(2):85–100, 2009.
- [182] Musikding. The face 69er fuzz - <https://www.musikding.de/The-Face-69er-Fuzz-kit>, 2015.
- [183] Hughes and Kettner. *Tube Factor – User's Manual*, 1998.
- [184] Musikding. The screamer overdrive - <https://www.musikding.de/The-Screamer-Overdrive-kit>, 2015.
- [185] Kristjan Dempwolf, Martin Holters, Stephan Möller, and Udo Zölzer. The influence of small variations in a simplified guitar amplifier model. In *Proceedings of the 12th International Conference on Digital Audio Effects (DAFx-09)*, Como, Italy, Sep. 2009.
- [186] Herwig Behrends, Adrian von dem Knesebeck, Werner Bradinal, Peter Neumann, and Udo Zölzer. Automatic equalization using parametric IIR filters. *Journal of the Audio Engineering Society*, 59(3):102–109, 2011.
- [187] Fraunhofer Institute for Digital Media Technology IDMT. IDMT-SMT-guitar database, [http://www.idmt.fraunhofer.de/en/business\\_units/m2d/smt/guitar.html](http://www.idmt.fraunhofer.de/en/business_units/m2d/smt/guitar.html), 2014.