4

Nonlinear processing

P. Dutilleux, K. Dempwolf, M. Holters and U. Zölzer

4.1 Introduction

Audio effect algorithms for dynamics processing, valve simulation, overdrive and distortion for guitar and recording applications, psychoacoustic enhancers and exciters fall into the category of nonlinear processing. They create intentional or unintentional harmonic or inharmonic frequency components which are not present in the input signal. Harmonic distortion is caused by nonlinearities within the effect device. For most of these signal processing devices, a lot of listening and recording experience is necessary to obtain sound results which are preferred by most listeners. The effect parameters have to be adjusted carefully by ear while monitoring the output signal by a level meter. The application of these signal processing devices is an art of its own and of course one of the main tools for recording engineers and musicians.

The terms nonlinear processing or nonlinear processors are used for all signal processing algorithms or signal processing devices in the analog or digital domains which do *not* satisfy the condition of linearity. A system with input $x(n)$ and output $y(n)$ is called linear if the property

$$x(n) = Ax_1(n) + Bx_2(n) \rightarrow y(n) = Ay_1(n) + By_2(n) \tag{4.1}$$

is fullfilled. In all other cases it is called nonlinear. A difference can be made between static or memoryless nonlinearities, that can be described by a static input-to-output mapping, and dynamic nonlinear systems.

If a sinusoid of known amplitude and frequency according to $x(n) = A\sin(2\pi f_1 Tn)$ is fed to a linear system the output signal can be written as $y(n) = A_{\text{out}}\sin(2\pi f_1 Tn + \varphi_{\text{out}})$ which again is a sinusoid where the amplitude is modified by the magnitude response $|H(f_1)|$ of the transfer function according to $A_{\text{out}} = |H(f_1)| \cdot A$ and the phase response $\varphi_{\text{out}} = \varphi_{\text{in}} + \angle H(f_1)$. In contrast, a

DAFX: Digital Audio Effects, Second Edition. Edited by Udo Zölzer.
© 2011 John Wiley & Sons, Ltd. Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-66599-2

nonlinear system will deliver a sum of sinusoids $y(n) = A_0 + A_1 \sin(2\pi f_1 Tn) + A_2 \sin(2 \cdot 2\pi f_1 Tn) + \ldots + A_N \sin(N \cdot 2\pi f_1 Tn)$. Block diagrams in Figure 4.1 showing both input and output signals of a linear and a nonlinear system illustrate the difference between both systems.

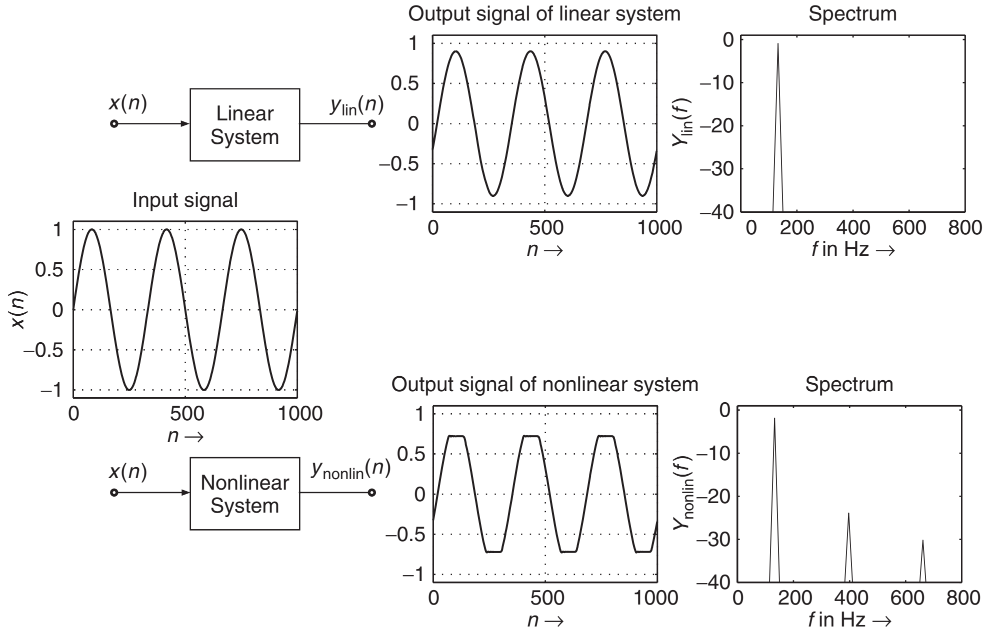

Figure 4.1 Input and output signals of a linear and nonlinear system. The output signal of the linear system is changed in amplitude and phase. The output signal of the nonlinear system is strongly shaped by the nonlinearity and consists of a sum of harmonics, as shown by the spectrum.

A measurement of the total harmonic distortion gives an indication of the nonlinearity of the system. Total harmonic distortion is defined by

$$\mathrm{THD} = \sqrt{\frac{A_2^2 + A_3^2 + \cdots + A_N^2}{A_1^2 + A_2^2 + \cdots + A_N^2}}, \qquad (4.2)$$

which is the square root of the ratio of the sum of powers of all harmonic frequencies above the fundamental frequency to the power of all harmonic frequencies, including the fundamental frequency.

The example of a sine wave has been used as an introduction to non linear distortion, but musical signals are usually comprised of several partials which induce many distortion products. If the input signal $x(n)$ is a sum of harmonics of a fundamental frequency f_0, the individual distortion products are also harmonics of f_0. The resulting signal $y(n)$ hence contains only harmonics, the relative amplitudes of which are non monotonously related to the amplitude of the input signal. If the input signal $x(n)$ is a sum of non harmonically related partials, the non linear system generates a series of harmonics for each input partial, as well as new partials at the frequencies of the sum and difference of the input partials and of their harmonics [CVAT01]. The non harmonic components produced by the non linear system are called intermodulation products because they can also be interpreted as the result of an amplitude modulation of one partial by another. In digital systems, foldover of the distortion products around $f_s/2$ are additional distortion products which have to be dealt with (refer to Section 4.1.1 and Figure 4.6).

In most usual musical applications, the intermodulation products are not welcome because they blur the input signal, impeding the identification of the individual musical sources or even introducing new musical parts which are out of tune. However, these effects can sometimes be used in a controlled and creative way, such as the amplification of the difference tones $f_2 - f_1$ [Dut96]. A general recommendation when applying non linear processing is to carefully control, before processing, the input signal $x(n)$ as to its harmonic content, amplitude and bandwidth, because filtering out the unpleasant or unwanted distortion products is usually challenging.

We will discuss nonlinear processing in three main musical categories. The first category consists of dynamic range controllers where the main purpose is the control of the signal envelope according to some control parameters. The amount of harmonic distortion introduced by these control algorithms should be kept as low as possible. Dynamics processing algorithms will be introduced in Section 4.2. The second class of nonlinear processing is designed for the creation of strong harmonic distortion such as guitar amplifiers, guitar effect processors, etc. and will be introduced in Section 4.3. The third category can be described by the same theoretical approach and is represented by signal-processing devices called exciters and enhancers. Their main field of application is the creation of additional harmonics for a subtle improvement of the main sound characteristic. The amount of harmonic distortion is usually kept small to avoid a pronounced effect. Exciters and enhancers are based on psycho-acoustic fundamentals and will be discussed in Section 4.4.

4.1.1 Basics of nonlinear modeling

Digital signal processing is mainly based on linear time-invariant systems. The assumption of linearity and time invariance is certainly valid for a variety of technical systems, especially for systems where input and output signals are bounded to a specified amplitude range. However, several analog audio processing devices have nonlinearities like valve amplifiers, analog effect devices, analog tape recorders, loudspeakers and at the end of the chain the human hearing mechanism. A compensation and the simulation of these nonlinearities need nonlinear signal processing and of course a theory of nonlinear systems. From several models for different nonlinear systems discussed in the literature, the Volterra series expansion is a suitable approach, because it is an extension of the linear systems theory. Not all technical and physical systems can be described by the Volterra series expansion, especially systems with extreme nonlinearities. If the inner structure of a nonlinear system is unknown, a typical measurement set-up, as shown in Figure 4.2, with a pseudo-random signal as the input signal and recording the output signal is used. Input and output signals allow the calculation of the linear impulse response $h_1(n)$ by cross-correlation and *kernels* (impulse responses) of higher order $h_2(n_1, n_2), h_3(n_1, n_2, n_3), \cdots, h_N(n_1, \cdots, n_N)$ by higher order cross-correlations. The linear impulse response $h_1(n)$ is a one-dimensional, $h_2(n_1, n_2)$ is a two-dimensional and $h_N(n_1, \cdots, n_N)$ is an N-dimensional kernel. An exhaustive treatment of these techniques can be found in [Sch80]. These N kernels can be used for an N-order Volterra system model given by

$$
\begin{aligned}
y(n) = \sum_{i=1}^{N} y_i(n) &= \sum_{\nu_1=0}^{\infty} h_1(\nu_1)x(n-\nu_1) && (4.3)\\
&+ \sum_{\nu_1=0}^{\infty}\sum_{\nu_2=0}^{\infty} h_2(\nu_1, \nu_2)x(n-\nu_1)x(n-\nu_2)\\
&+ \sum_{\nu_1=0}^{\infty}\sum_{\nu_2=0}^{\infty}\sum_{\nu_3=0}^{\infty} h_3(\nu_1, \nu_2, \nu_3)x(n-\nu_1)x(n-\nu_2)x(n-\nu_3) + \cdots\\
&+ \sum_{\nu_1=0}^{\infty}\cdots\sum_{\nu_N=0}^{\infty} h_N(\nu_1, \ldots, \nu_N)x(n-\nu_1)\cdots x(n-\nu_N). && (4.4)
\end{aligned}
$$

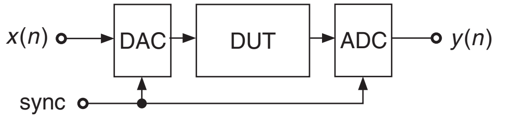

Figure 4.2 Measurement of nonlinear systems.

Figure 4.3 shows the block diagram representing (4.4).

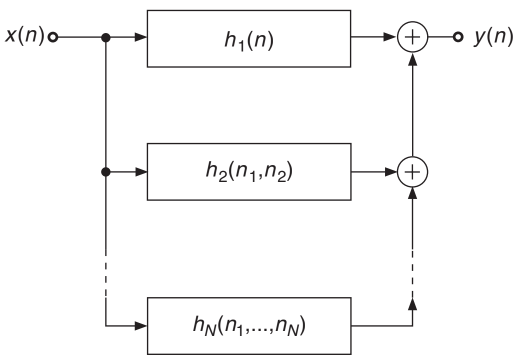

Figure 4.3 Simulation of nonlinear systems by an N-order Volterra system model.

A further simplification [Fra97] is possible if the kernels can be factored according to

$$h_i(n_1, n_2, \cdots, n_i) = h_i^f(n_1)h_i^f(n_2)\cdots h_i^f(n_i). \tag{4.5}$$

Then (4.4) can be written as

$$y(n) = \sum_{\nu=0}^{\infty} h_1^f(\nu)x(n-\nu) + \left(\sum_{\nu=0}^{\infty} h_2^f(\nu)x(n-\nu)\right)^2 + \cdots + \left(\sum_{\nu=0}^{\infty} h_N^f(\nu)x(n-\nu)\right)^N, \tag{4.6}$$

which is shown in block diagram representation in Figure 4.4. This representation shows several advantages, especially from the implementation point of view, because every subsystem can be realized by a one-dimensional impulse response. At the output of each subsystem we have to perform the $()^i$ operation on the corresponding output signals. The discussion so far can be applied to nonlinear systems with memory, which means that besides nonlinearities linear filtering operations are also included. Further material on nonlinear audio systems can be found in [Kai87, RH96, Kli98, FUB$^+$98].

A simulation of a nonlinear system without memory, namely static nonlinear curves, can be done by a Taylor series expansion given by

$$y(n) = f[x(n)] = \sum_{i=0}^{N} b_i x^i(n). \tag{4.7}$$

Static nonlinear curves can be applied directly to the input signal, where each input amplitude is mapped to an output amplitude according to the nonlinear function $y = f[x(n)]$. If one applies

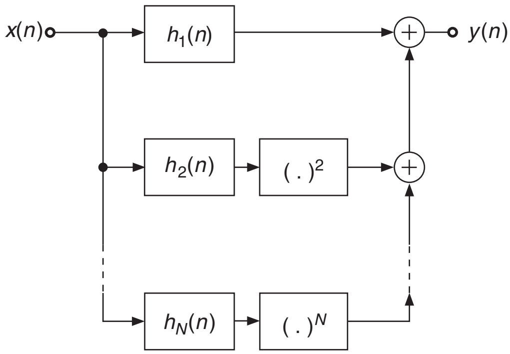

Figure 4.4 Simulation of nonlinear systems by an N-order Volterra system model with factored kernels.

a squarer operation to the input signal of a given bandwidth, the output signal $y(n) = x^2(n)$ will double its bandwidth. As soon as the highest frequency after passing a nonlinear operation exceeds half the sampling frequency $f_s/2$, aliasing will fold this frequency back to the base band. This means that for digital signals we first have to perform over-sampling of the input signal before applying any nonlinear operation to the input signal in order to avoid any aliasing distortions. This over-sampling is shown in Figure 4.5 where first up-sampling is performed and then an interpolation filter H_I is used to suppress images up to the new sampling frequency Lf_s. Then the nonlinear operation can be applied followed by a band-limiting filter to $f_s/2$ and down-sampling.

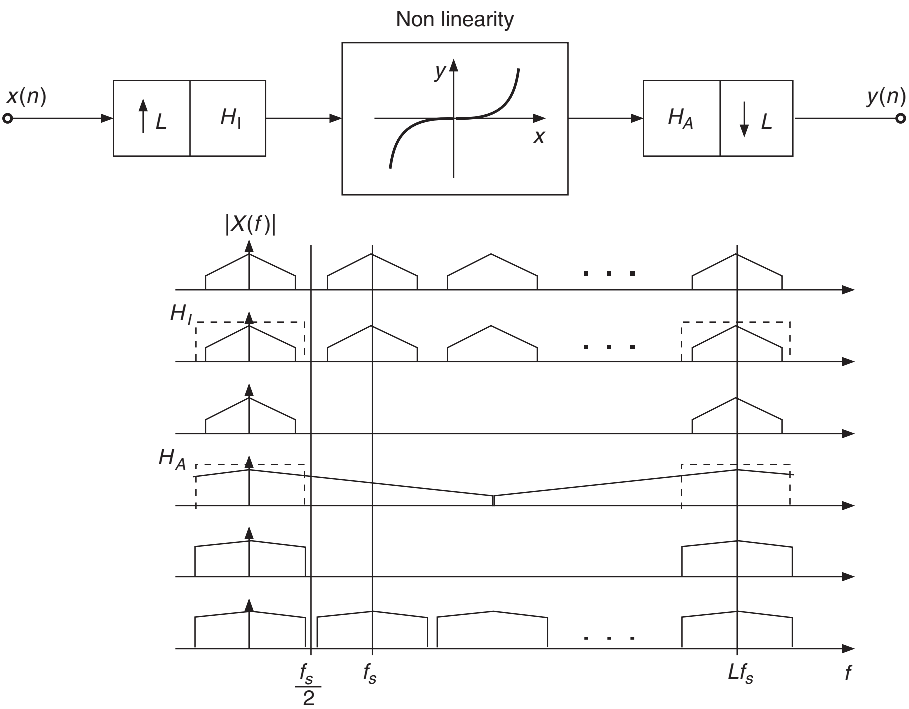

Figure 4.5 Nonlinear processing by over-sampling techniques.

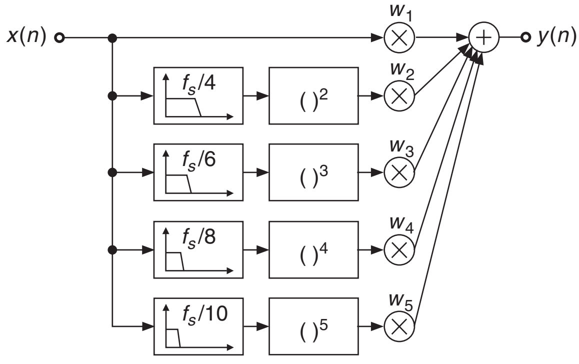

Figure 4.6 Nonlinear processing by band-limiting input range [SZ99].

As can be noticed, the output spectrum only contains frequencies up to $f_s/2$. Based on this fact the entire nonlinear processing can be performed without over-sampling and down-sampling by the system shown in Figure 4.6 [SZ99]. The input signal is split into several lowpass versions which are forwarded to an individual nonlinearity. The output signal is a weighted combination of the individual output signals after passing a nonlinearity. With this approach the problem of aliasing is avoided and an approximation to the over-sampling technique is achieved. A comparison with our previous discussion on factored Volterra kernels shows also a close connection. As a conclusion for a special static nonlinearity applied to an input signal, the input signal has to be filtered by a lowpass of cut-off frequency $f_s/(2 \cdot \text{order of the Taylor series})$, otherwise aliasing distortions will occur.

4.2 Dynamic range control

Dynamics processing is performed by amplifying devices where the gain is automatically controlled by the level of the input signal. We will discuss limiters, compressors, expanders and noise gates. A good introduction to the parameters of dynamics processing can be found in [Ear76, pp. 242–248].

Dynamics processing is based on an amplitude/level detection scheme sometimes called an envelope follower, a static curve to derive a gain factor from the result of the envelope follower, a smoothing filter to prevent too abrupt gain changes and a multiplier to weight the input signal (see Figure 4.7). Optionally, the input signal is delayed to compensate for any delay in the *side chain*, the lower path in Figure 4.7. Normally, the gain factor is derived from the input signal, but the side chain path can also be connected to another signal for controlling the gain factor of the input signal.

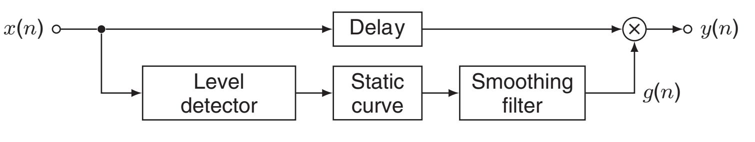

Figure 4.7 Block diagram of a dynamic range controller [Zöl05].

An example of a dynamic range controller's impact on a signal is shown in Figure 4.8. In this example, the dynamic range controller is a *compressor*: during periods of low signal level (first half of the shown input signal $x(n)$), the gain $g(n)$ is set to 1, while for higher input levels (second half), the gain is decreased to reduce the signal's amplitude. Thus the difference in signal levels along time is reduced, compressing the signal's dynamic range.

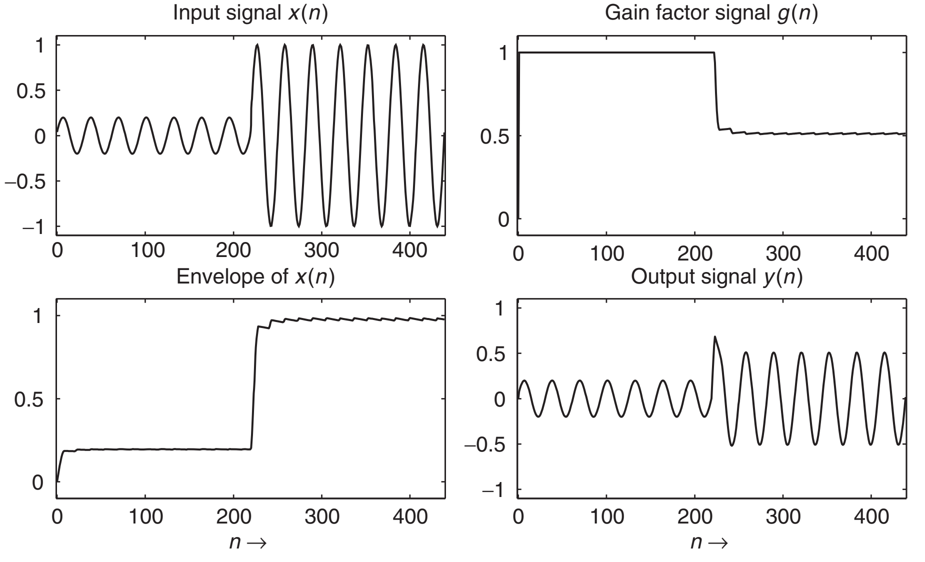

Figure 4.8 The input signal $x(n)$, the envelope of $x(n)$, the derived gain factor $g(n)$ and the output signal $y(n)$.

Signal processing

The level detector follows one of the approaches described in Section 4.3. Two commonly used choices are:

- A full-wave rectifier in combination with an AR-averager (see Figure 4.9) with very short attack time may be used to track the peak value of the signal.
- A squarer averaged with only a single time-constant (see Figure 4.10) provides the RMS value, measuring the signal's power.

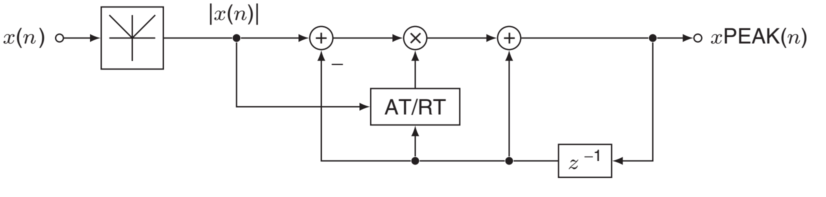

Figure 4.9 Peak measurement (envelope detector/follower) for a dynamic range controller.

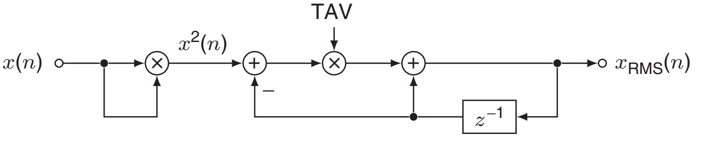

Figure 4.10 RMS measurement (envelope detector/follower) for a dynamic range controller [McN84, Zöl05].

Typical dynamic range controllers contain one of these level detectors, or even both, to be able to react to sudden level increases with the peak detector, while usually operating based on the RMS measurement.

The static curve decides which function (limiter, compressor, expander, ...) the dynamic range controller assumes. It is best described on a dB scale, so we introduce X, G and Y for the levels of the respective signals in dB, i.e., $X = 20 \cdot \log_{10}(x_{\text{PEAK}}(n))$ or $X = 10 \cdot \log_{10}(x_{\text{RMS}}(n))$. The multiplication at the dynamic range controller's output then becomes $Y = X + G$.

In this logarithmic domain, the gain is typically derived from the input level by a piecewise linear function. Figure 4.11 shows the characteristic curve of a typical compressor, mapping input to output level (on the left) and input level to gain (on the right). Up to a certain input level, the compressor threshold CT (-30 dB in this example), the gain is left at 0 dB. Above the threshold, the characteristic curve is defined by the ratio R, specifying the relation between input level change ΔX to output level change ΔY by $\Delta Y = \frac{1}{R} \cdot \Delta X$. Alternatively, the slope factor $S = 1 - \frac{1}{R}$ may be specified instead of the ratio R.

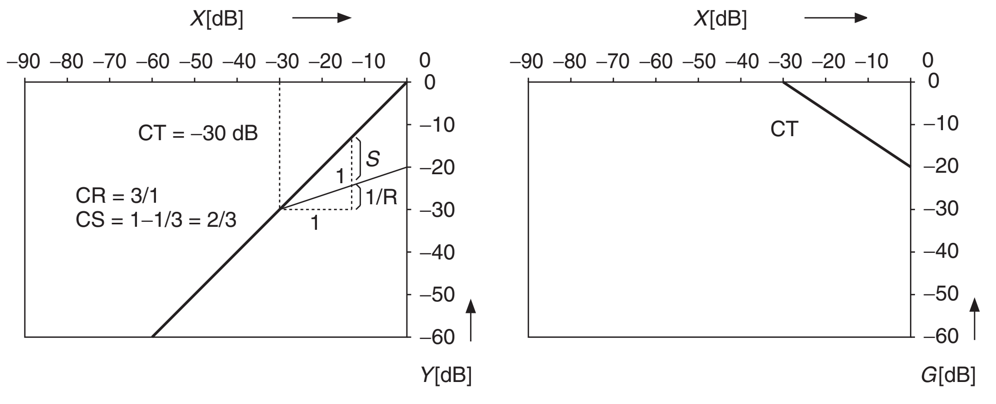

Figure 4.11 Static characteristic: definition of slope S and ratio R [Zöl05] exemplified for a compressor.

The *dynamic behavior* of a dynamic range controller is influenced by the level measurement approach (with attack AT and release time RT for peak measurement and averaging time TAV for RMS measurement) and further adjusted with a smoothing filter, as shown in Figure 4.12,

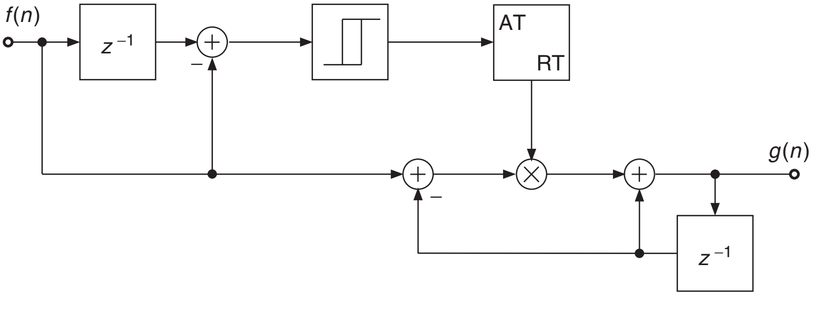

Figure 4.12 Dynamic filter: attack and release time adjustment for a dynamic range controller [McN84, Zöl05].

which also employs special attack/release times. The calculation of the attack time parameter is carried out by

$$\mathrm{AT} = 1 - \mathrm{e}^{-2.2T/t_{\mathrm{AT}}},$$

where t_{AT} is the time parameter in seconds and T is the sampling period. The release time parameter RT and the averaging parameter TAV can be computed by the same formula by replacing t_{AT} by t_{RT} or t_{TAV}, respectively. Further details and derivations can be found in [McN84, Zöl05]. The output of the static function is used as the input signal to the dynamic filter with attack and release times in Figure 4.12. The output signal $g(n)$ is the gain factor for weighting the delayed input signal $x(n-D)$ as shown in Figure 4.7. In the following sections some special dynamic range controllers are discussed in detail.

4.2.1 Limiter

The purpose of a limiter is to provide control over the highest peaks in the signal, but to otherwise change the dynamics of the signal as little as possible. This is achieved by employing a characteristic curve with an infinite ratio $R = \infty$ above a limiter threshold LT, i.e.,

$$G = \begin{cases} 0 \text{ dB} & \text{if} \quad X < LT \\ LT - X & \text{else.} \end{cases} \tag{4.8}$$

Thus, the output level $Y = X + G$ should never exceed the threshold LT. By lowering the peaks, the overall signal can be boosted. Beside limiting single instrument signals, limiting is also often performed on the final mix of a multichannel application.

A limiter makes use of peak level measurement and should react very quickly when the input signal exceeds the limiter threshold. Typical parameters for a limiter are $t_{\mathrm{AT}} = 0.02 \ldots 10$ ms and $t_{\mathrm{RT}} = 1 \ldots 5000$ ms, for both the peak measurement and the smoothing filter.

An actual implementation, like the one in Figure 4.13, may perform the gain computation in linear values by

$$g(n) = \min\left(1, \frac{lt}{x_{\mathrm{PEAK}}(n)}\right) \tag{4.9}$$

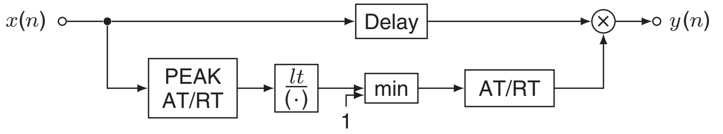

Figure 4.13 Block diagram of a limiter.

where $lt = 10^{\frac{LT}{20}}$ is the threshold on the linear scale. This approach is also used in M-file 4.1, a very simple limiter implementation with the same filter coefficients in the level detector and the smoothing filter. Figure 4.14 demonstrates the behavior of the time signals inside a limiter configuration of Figure 4.13.

M-file 4.1 (limiter.m)

```
function y = limiter(x, lt)
% function y = limiter(x, lt)
% Author: M. Holters
```

```
at = 0.3;
rt = 0.01;
delay = 5;

xpeak = 0;
g = 1;
buffer = zeros(1,delay);

for n = 1:length(x)
  a = abs(x(n));
  if a > xpeak
    coeff = at;
  else
    coeff = rt;
  end;
  xpeak = (1-coeff) * xpeak + coeff * a;
  f = min(1, lt/xpeak);
  if f < g
    coeff = at;
  else
    coeff = rt;
  end;
  g = (1-coeff) * g + coeff * f;
  y(n) = g * buffer(end);
  buffer = [x(n) buffer(1:end-1)];
end;
```

4.2.2 Compressor and expander

While a limiter completely eliminates any dynamics above a certain threshold by keeping the output level constant, a compressor only reduces the dynamics, compressing the dynamic range. The reduced dynamics can be exploited to increase the overall level, thereby boosting the loudness, without exceeding the allowed amplitude range. An expander does just the opposite of the compressor, increasing dynamics by mapping small level changes of the input to larger changes of the output level. Applying an expander to low-level signals gives a lively sound characteristic. The corresponding ratios and slopes of the characteristic curve are $CR > 1$ and $0 < CS < 1$ for the compressor and $0 < ER < 1$ and $ES < 0$ for the expander.

Compressors and expanders typically employ RMS level detectors with an averaging time in the range $t = 5 \ldots 130$ ms and a smoothing filter with $t_{\mathrm{AT}} = 0.1 \ldots 2600$ ms and $t_{\mathrm{RT}} = 1 \ldots 5000$ ms.

Combining a compressor for high signal levels with an expander for low signal levels leads to the gain computation

$$G = \begin{cases} CS \cdot (CT - X) & \text{if} \quad X > CT \\ 0\ \mathrm{dB} & \text{if} \quad ET \le X \le CT \\ ES \cdot (ET - X) & \text{if} \quad X < ET \end{cases} \tag{4.10}$$

$$= \min\big(0, CS \cdot (CT - X), ES \cdot (ET - X)\big). \tag{4.11}$$

Here CT and ET denote the thresholds above which the compressor and below which the expander affect the signal, respectively. The resulting combined system is depicted in Figure 4.15. Again, the

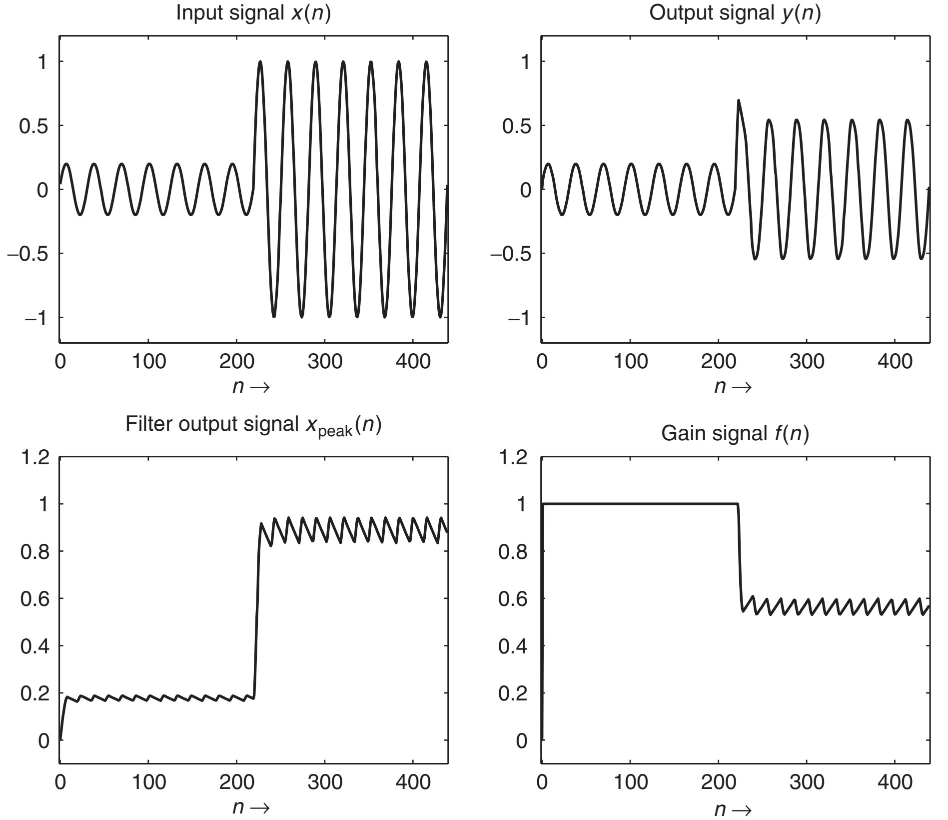

Figure 4.14 Signals for limiter.

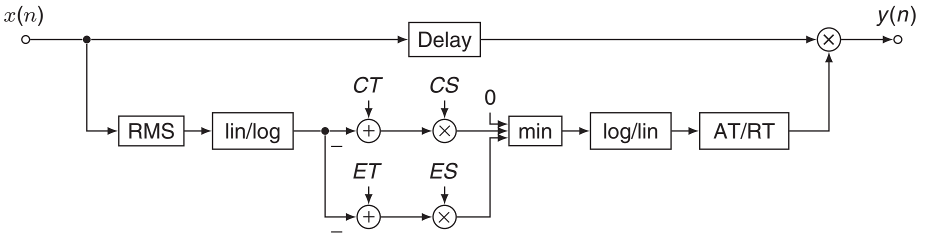

Figure 4.15 Block diagram of a compressor/expander [Zöl05].

gain can alternatively be calculated without using logarithmic values by

$$g(n) = \min\left(1, \left(\frac{x_{\text{RMS}}(n)}{ct^2}\right)^{-CS/2}, \left(\frac{x_{\text{RMS}}(n)}{et^2}\right)^{-ES/2}\right), \tag{4.12}$$

where the square root operation of the RMS measurement is moved to the exponenents by halving the respective slopes. This approach makes conversion to and from the logarithmic domain unnecessary, but requires exponentiation. An implementation following the approach of Figure 4.15 is given in M-file 4.2.

M-file 4.2 (compexp.m)

```
function y = compexp(x, CT, CS, ET, ES)
% function y = compexp(x, CT, CS, ET, ES)
% Author: M. Holters

tav = 0.01;
at = 0.03;
rt = 0.003;
delay = 150;

xrms = 0;
g = 1;
buffer = zeros(1,delay);

for n = 1:length(x)
  xrms = (1-tav) * xrms + tav * x(n)^2;
  X = 10*log10(xrms);
  G = min([0, CS*(CT-X), ES*(ET-X)]);
  f = 10^(G/20);
  if f < g
    coeff = at;
  else
    coeff = rt;
  end;
  g = (1-coeff) * g + coeff * f;
  y(n) = g * buffer(end);
  buffer = [x(n) buffer(1:end-1)];
end;
```

Parallel compression

The usual application of compressors is the reduction of the amplitude variations above a specified threshold. In some circumstances however, e.g., for percussion instruments or an acoustic guitar, it is desired to leave unaffected the high-amplitude transients while enhancing the weaker parts of the signal. For such applications, parallel compression is recommended, where a heavily compressed version of the signal is added to the unaffected signal. The soft parts of the sound are heavily compressed using typically a low threshold, a high ratio and a short attack time. The compressed part is then amplified and mixed to the original signal. The resulting sound retains the transparency of the original sound because the transients are unaffected. This type of processing is sometimes referred to as New-York compression or side-chain compression [Bev77, Hul97, Kat07, Izh10].

Multiband compression

Depending on the spectral content of the sound to process, compression might be required only on a selected frequency band. Let us consider the musical example of a piece for tuba and live electronics where the fingering and the breathing noises had to be processed irrespective of the overwhelming low frequency tones of the instrument. The solution was the insertion of a high-pass filter in the main path as well as the side-chain of a compressor so that the device left the lower part of the spectrum unaffected by the compression.

For the general case where several frequency bands of the spectrum have to be processed individually, the multiband compressor has been developed. A filter bank splits the input signal in several bands (usually three to five) which are then individually processed. The side-chain signal of each compressor is also derived from a band-limited version of the input signal. The filter-bank for the side-chain is usually the same as the for the input signal, but it is sometimes implemented as an independent processing unit so that intricate relations can be set up where the content of the input signal in a limited frequency band drives the compression in another frequency band.

Multiband dynamics processors are often used during mastering. At this stage, no direct correction of the individual musical parts is possible since the mix is finished and the music is usually available as a stereo track. Under the assumption that each musical part or instrument mainly occupies a particular frequency band, minor corrections are still possible if they are band limited.

One advantage of multiband compression is that the unwanted side-effects of compression, such as alteration of the tonal balance or distortions, can be limited to a given frequency band. With the objective of dynamic range compression, the reduction of the ratio of peak to RMS amplitude can be more effectively implemented in individual frequency bands than at the full audio bandwidth. This allows the increase of the overall loudness of the track. Considering that the casual listener often gives preference to the louder of two musical works, a trend has developed towards always higher average levels. Moderation is, however, necessary because the “loudness war” can induce severe degradations of the audio and musical quality [Kat07, Lun07].

4.2.3 Noise gate

A noise gate can be considered as an extreme expander with a slope of $-\infty$. This results in the complete muting of signals below the chosen threshold NT. As the name implies, the noise gate is typically used to gate out noise by setting the threshold just above the level of the background noise, such that the gate only opens when a desired signal with a level above the threshold is present. A particular application is found when recording a drum set. Each element of the drum set has a different decay time. When they are not manually damped, their sounds mix together and the result is no longer distinguishable. When each element is processed by a noise gate, every sound can automatically be faded out after the attack part of the sound. This results in an overall cleaner sound.

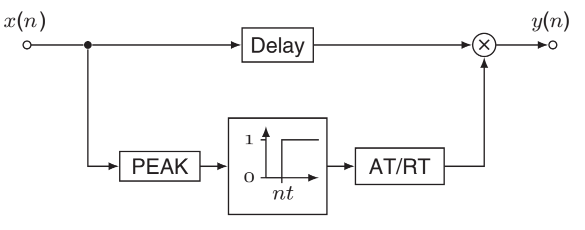

Figure 4.16 Block diagram of a noise gate [Zöl05].

The functional units of a noise gate are shown in Figure 4.16. The decision to activate the gate is typically based on a peak measurement which leads to a fade in/fade out of the gain factor $g(n)$ with appropriate attack and release times. Further possible refinements include the use of two thresholds to realize a hysteresis and a hold time to avoid unpleasant effects when the input level fluctuates around the threshold. These are demonstrated in the implementation given in M-file 4.3 [Ben97].

M-file 4.3 (noisegt.m)

```
function y=noisegt(x,holdtime,ltrhold,utrhold,release,attack,a,Fs)
% function y=noisegt(x,holdtime,ltrhold,utrhold,release,attack,a,Fs)
% Author: R. Bendiksen
```

```
% noise gate with hysteresis
% holdtime   - time in seconds the sound level has to be below the
%                  threshhold value before the gate is activated
% ltrhold    - threshold value for activating the gate
% utrhold    - threshold value for deactivating the gate > ltrhold
% release    - time in seconds before the sound level reaches zero
% attack     - time in seconds before the output sound level is the
%                  same as the input level after deactivating the gate
% a          - pole placement of the envelope detecting filter <1
% Fs         - sampling frequency
rel=round(release*Fs);  %number of samples for fade
att=round(attack*Fs);   %number of samples for fade
g=zeros(size(x));
lthcnt=0;
uthcnt=0;
ht=round(holdtime*Fs);
h=filter([(1-a)^2],[1.0000 -2*a a^2],abs(x));%envelope detection
h=h/max(h);
for i=1:length(h)
  if (h(i)<=ltrhold) | ((h(i)<utrhold) & (lthcnt>0))
% Value below the lower threshold?
    lthcnt=lthcnt+1;
    uthcnt=0;
    if lthcnt>ht
% Time below the lower threshold longer than the hold time?
      if lthcnt>(rel+ht)
        g(i)=0;
      else
        g(i)=1-(lthcnt-ht)/rel;    % fades the signal to zero
      end;
    elseif ((i<ht) & (lthcnt==i))
      g(i)=0;
    else
      g(i)=1;
    end;
  elseif (h(i)>=utrhold) | ((h(i)>ltrhold) & (uthcnt>0))
% Value above the upper threshold or is the signal being faded in?
    uthcnt=uthcnt+1;
    if (g(i-1)<1)
% Has the gate been activated or isn't the signal faded in yet?
      g(i)=max(uthcnt/att,g(i-1));
    else
      g(i)=1;
    end;
    lthcnt=0;
  else
    g(i)=g(i-1);
    lthcnt=0;
    uthcnt=0;
  end;
end;
y=x.*g;
y=y*max(abs(x))/max(abs(y));
```

Further implementations of limiters, compressors, expanders and noise gates can be found in [Orf96] and in [Zöl05], where special combined dynamic range controllers are also discussed.

4.2.4 De-esser

A de-esser is a signal processing device for processing speech and vocals. It consists of a bandpass filter tuned to the frequency range between 2 and 6 kHz to detect the level of the signal in this frequency band. If a certain threshold is exceeded, the gain factor of a peak or notch filter tuned to the same frequency band is controlled to realize a compressor for that band only, as shown in Figure 4.17. De-essers are mainly applied to speech or vocal signals to avoid high frequency sibilance.

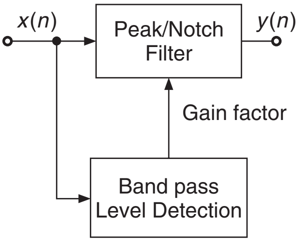

Figure 4.17 Block diagram of a de-esser.

As an alternative to the bandpass/notch filters, highpass and shelving filters are used with good results. In order to make the de-esser more robust against input level changes, the threshold should depend on the overall level of the signal – that is, a relative threshold [Nie00].

4.2.5 Infinite limiters

In order to catch overshoots from a compressor and limiter, a clipper or *infinite limiter* may be used [Nie00]. In contrast to the previously considered dynamic range controllers, the infinite limiter is a nonlinear operation working directly on the waveform by flattening the signal above a threshold. The simplest one is hard clipping, which generates lots of high-order harmonics. A gentler infinite limiter is the soft clipper, which rounds the signal shape before the absolute clipping threshold. Although infinite limiting should usually be avoided during mix down of multi channel recordings and recording sessions, several CDs make use of infinite limiting or saturation (see wave file in Figure 4.18 and listen to Carlos Santana/Smooth [m-San99]). A more detailed description of clipping effects and their artistic use is given in the following section.

4.3 Musical distortion and saturation effects

4.3.1 Valve simulation

Introduction

Valve or tube devices dominated electronic signal processing circuits during the first part of the last century and have experienced a revival in audio processing every decade since their introduction [Bar98, Ham73]. One of the most commonly used effects for electric guitars is the amplifier and especially the valve amplifier. The typical behavior of the amplifier and the connected loudspeaker

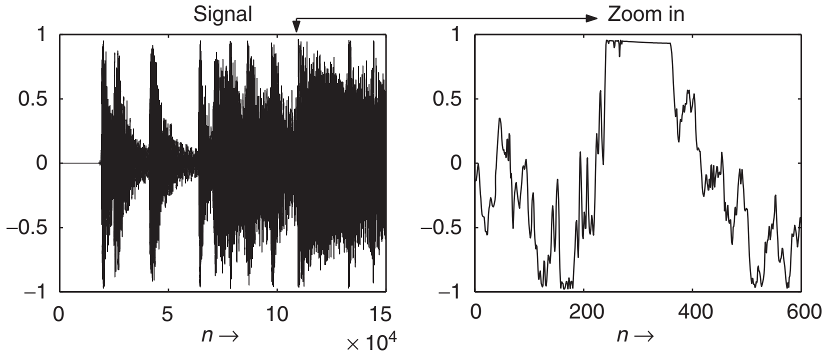

Figure 4.18 Infinite limiting (Santana – "Smooth").

cabinet have demonstrated their influence on the sound of rock music over the past decades. Besides the two most important guitars, namely the Fender Stratocaster and the Gibson Les Paul, several valve amplifiers have helped in creating exciting sounds from these classic guitars.

Valve microphones, preamplifiers and effect devices such as compressors, limiters and equalizers are also used for vocal recordings where the warm and subtle effect of valve compression is applied. A lot of vocalists prefer recording with valve condenser microphones because of their warm low end and smooth top end frequency response. Also the recording of acoustical instruments such as acoustic guitars, brass instruments and drums benefit from being processed with valve outboard devices. Valve processors also assist the mixing process for individual track enhancing and on the mix buses. The demand for valve outboard effects and classic mixing consoles used in combination with digital audio workstations has led back to entire valve technology mixing consoles. For the variety of digital audio workstations a lot of plug-in software modules for valve processing are available.

Vintage valve amplifiers

An introduction to valve amplifiers and their history can be found in [Fli93, Bar98], where several comments on sound characteristics are published. We will concentrate on the most important amplifier manufacturers over the years and point out some characteristic features.

- **Fender:** The Fender series of guitar amplifiers goes back to the year 1946 when the first devices were introduced. These were based on standard tube schematics supplied by the manufacturers of tubes. Over the years modifications of the standard design approach were integrated in response to musicians' needs and proposals. The range of Fender amplifiers is still expanding, but also reissues of the originals are very popular with musicians. The sound of Fender amplifiers is the "classic tube sound" and was made famous by blues musicians like Buddy Guy and Stevie Ray Vaughan. A detailed analysis of a Fender amplifier and a discussion on its similarity to other amplifier designs is presented in [Kue05].

- **Vox:** The manufacturer Vox is always associated with its most famous amplifier, the Vox AC30/4. Its sound is best characterized by guitar player Brian May in [PD93] where he states "the quality at low levels is broad and crisp and unmistakably *valve like*, and as the volume is turned up it slides into a pleasant, creamy compression and distortion." There is always a ringing treble quality through all level settings of the amp. The real "soul of the amp" comes out if you play it at full volume and control your sound with the volume knob of your guitar. The heart of the sound characteristic of the Vox AC30/4 is claimed to be the use of EL84 pentodes, negative feedback and cathode biasing in Class A configuration. The four small EL84s should sound more lively than the bigger EL34s. The sound of the Vox AC30 can be found on recordings by Brian May, Status Quo, Tom Petty and Bryan Adams.

- **Marshall:** The Fender Bassman 5F6 was the basis for the Marshall JTM 45. The differences between both are discussed in [Doy93] and [Kue05], and are claimed to be the output transformers, speakers, input valve and feedback circuit, although the main circuit diagrams are nearly identical. The sound of Marshall is characterized by an aggressive and "crunchy" tone with brilliant harmonics, as Eric Clapton says, "I was probably playing full volume to get that sound" [Doy93]. Typical representatives of the early Marshall sound are Jimi Hendrix, Eric Clapton, Jimmy Page and Ritchie Blackmore. The JCM800 series established the second generation of Marshall amplifiers featuring the typical hardrock and heavy metal sound as played by Zakk Wylde or Slash.

- **Mesa-Boogie:** The "creamy" tone and the high gain of Mesa-Boogie amplifiers has its seeds in a wrongly connected test arrangement of two preamp stages. Founder Randall Smith stated in an interview [Sal02]: "Then when we plugged it in right, Lee hit a big power chord and practically blew both our bodies right through the back wall! (...) It was just HUGE sounding. And it would sustain forever. That was the beginning of cascading high-gain pre-amp architecture". Through this assembly Mesa-Boogie amplifiers featured 50 to 80 times more gain compared to amplifier designs of that time, leading to a long sustaining tone. An ambassador for Mesa-Boogie amplifiers is Carlos Santana.

Signal processing

The sound of a valve amplifier is based on a combination of several important factors. First of all the main processing features of valves or tubes are important [Bar98, Ham73]. Then the amplifier circuit has its influence on the sound and, last but not least, the chassis and loudspeaker combination play an important role in sound shaping. We will discuss all three factors now.

Valve basics. Valves or vacuum tubes are active electronic components used for amplifying, rectifying, switching, modulating or generating electrical signals. Prior to the invention of transistors, valves were the main active components in electronic equipment. In today's electronics, valves are replaced completely by semiconductors, except for some special applications. In the following we will discuss the reasons why these components are still very popular in hi-fi and guitar amplifiers.

Triode valves [Rat95, RCA59] consist of three electrodes, namely anode (or plate), cathode (or filament) and grid. Applying a voltage between anode and cathode, electrons are emitted by the heated cathode and flow to the positively charged anode. The grid is placed between these electrodes and can be used to modulate the rate of electron flow. A negative charge on the grid electrode affects the electron flow: the larger the charge, the smaller the current from anode to cathode. Thus, the triode can be used as voltage-controlled amplifier. The corresponding transfer function relates the anode current I_A to the input grid voltage V_G, as depicted in Figure 4.19. This nonlinear curve has a quadratic shape. An input signal represented by the grid voltage V_G delivers an anode output current $I_A = f(V_G)$ representing the output signal. The corresponding output spectrum shows a second harmonic in addition to the input frequency. This second harmonic can be lowered in amplitude when the operating point of the nonlinear curve is shifted right and the input voltage is applied to the more linear region of the quadratic curve. As a consequence of this, triodes are considered to provide a warm and soft sound coloration when used in preamplifiers.

The dc component in the output signal can be suppressed by a subsequent highpass filter. Note also the asymmetrical soft clipping of the negative halves of the output sinusoid, which is the result of the quadratic curve of the triode. Input stages of valve amplifiers make use of these triode valves. A design parameter is the operating point which controls the amplitude of the second harmonic.

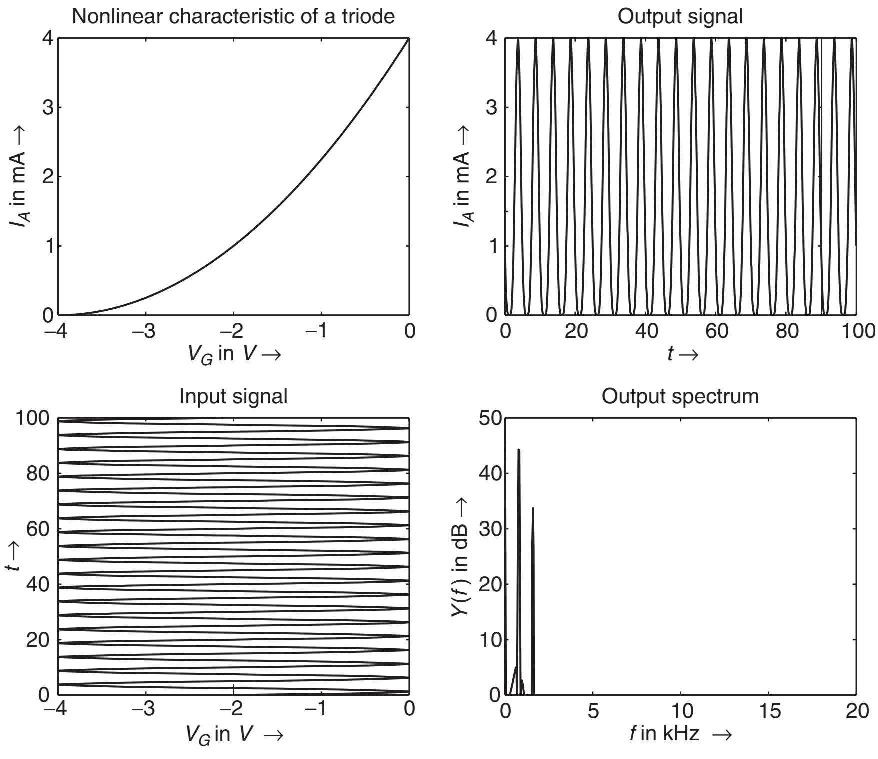

Figure 4.19 Triode: nonlinear characteristic curve $I_A = f(V_G)$ and nonlinear effect on input signal. The output spectrum consists of the fundamental input frequency and a second harmonic generated by the quadratic curve of the triode.

Pentode valves feature two additional electrodes, the screen and the suppressor grid. With this arrangement oscillations can be suppressed which can arise in triodes. When driving a load resistance the characteristic curve $I_A = f(V_G)$ is shaped like a S-curve, as shown in Figure 4.20. Through this, the output signal is compressed for higher-input amplitudes, leading to a symmetrical soft clipping. The corresponding output spectrum shows the creation of odd order harmonics. For lower input amplitudes the static characteristic curve operates in a nearly linear region, which again shows the control of the nonlinear behavior by properly selecting the operating point.

The technical parameters of valves have wide variation, which leads to a wide variation of sound features, although selected valves (so-called "matched pairs") with limited deviations of parameters are of course available. All surrounding environmental parameters like humidity and temperature have their influence as well.

Valve amplifier circuits. Valve amplifier circuits are based on the block diagram in Figure 4.21. Several measured signals from a Vox AC30 at different stages of the signal flow path are also displayed. This will give an indication of typical signal distortions in valve amplifiers. The corresponding spectra of the signals for the Vox AC30/4 measurements are shown in Figure 4.22. The distortion characteristic can be visualized by the waterfall representation of short-time FFTs for a chirp input signal in Figure 4.23. The main stages of a valve amplifier are given below:

- The *input stage* consists of a triode circuit providing the input matching and preamplification followed by a volume control for the next stages.

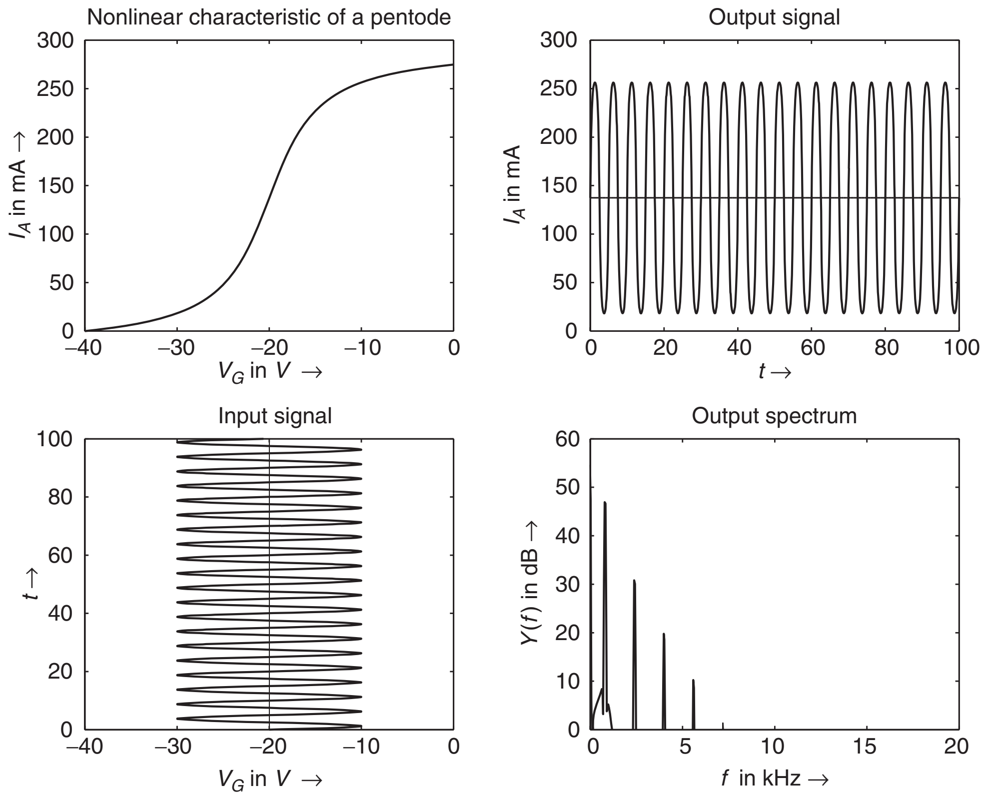

Figure 4.20 Pentode: nonlinear characteristic curve $I_A = f(V_G)$ and nonlinear effect on input signal.

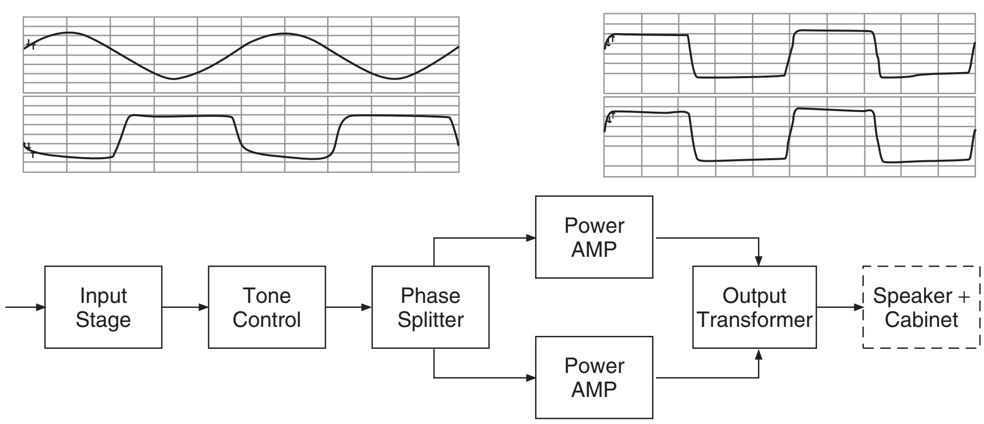

Figure 4.21 Main stages of a valve amplifier. Upper left plot shows signal after pre-amplifier, lower left plot shows signal after phase splitter, upper right plot shows signal after power amplifier and lower right plot shows signal after output transformer.

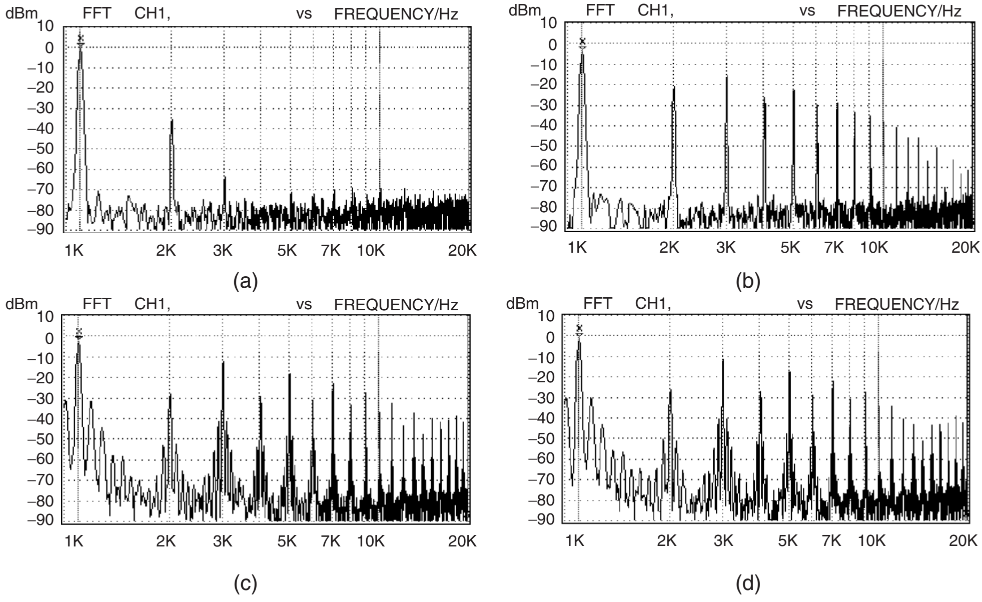

Figure 4.22 Vox AC30/4 spectra at different stages: (a) input stage, (b) output phase splitter, (c) output power amp and (d) output of transformer.

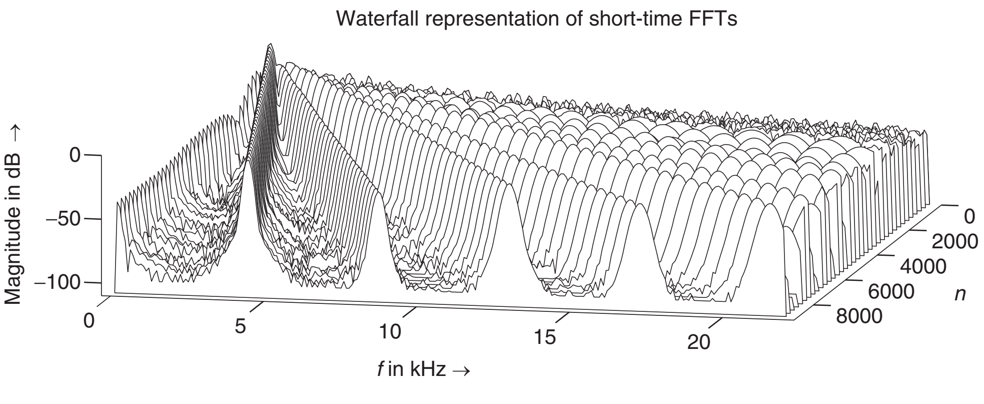

Figure 4.23 Short-time FFTs (waterfall representation) of Vox AC30 with a chirp input signal. The amplifier is operating with full volume setting.

- The *tone control* circuitry is based on passive filter networks, typically with three controls for bass, mid and treble.
- The *phase splitter* stage provides symmetrical power amp feeding. This phase splitter delivers the original input for the upper power amp and a phase inverted replica of the input for the lower power amp.
- The *power amp* stage in push-pull configuration performs individual amplification of the original and the phase inverted replica in a class A, class B or class AB configuration (see Figure 4.24). Class A is shown in the left plot, where the output signal is valid all the

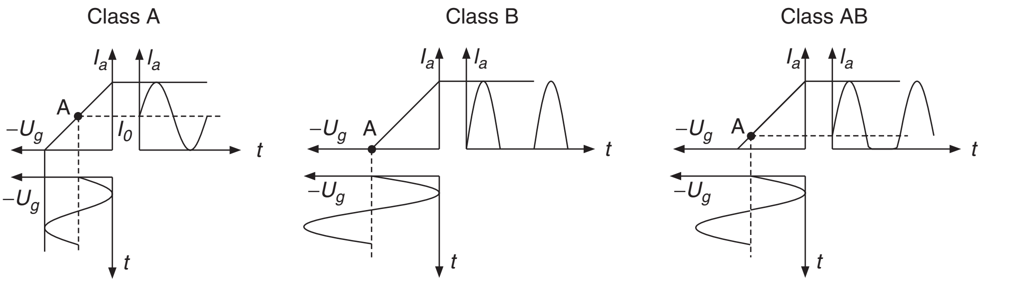

Figure 4.24 Power amplifier operation with idealized transfer characteristic (left class A, middle class B, right class AB operation).

time. Class B performs amplification only for one half wave, as depicted in the middle plot. The working point for class AB operation (right plot) lies in-between class A and class B, also amplifying a part of the negative half wave. Class A and class AB are the main configurations for guitar power amplifiers.

- The *output transformer* performs the subtraction of both waveforms delivered by the power amplifiers, which leads to a doubling of the output amplitude. Figure 4.25 shows the principle of push-pull amplification and the typical connection of the output transformer. The nonlinear behavior of transformers is beyond the scope of this discussion.

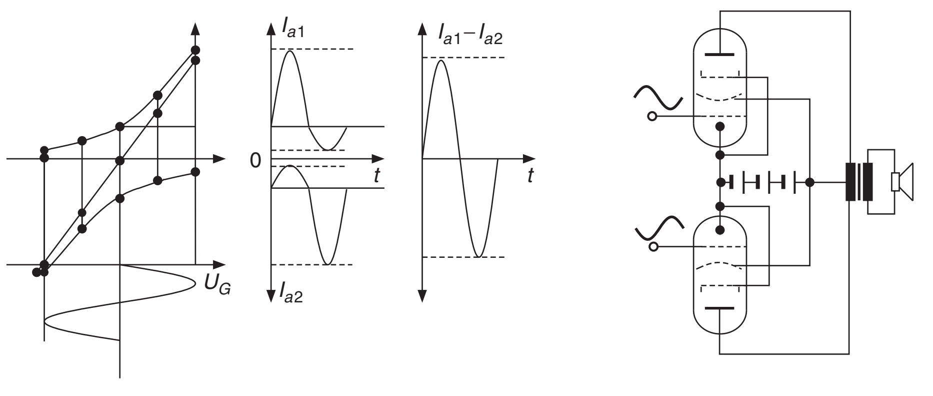

Figure 4.25 Power amplifier stage and output transformer. Left: principle and waveforms for class AB push-pull power amplifier. Right: simplified circuit with two pentodes, output transformer and loudspeaker.

Speaker and cabinet. Guitar amplifiers are built either as a *combo*, where amplifier chassis and one or more loudspeakers are combined in the same enclosure or as a *stack* with separated amplifier *head* and loudspeaker *cabinet*. The traditional guitar cabinet is closed-back and houses four 12″ speakers (4×12), but different combinations with loudspeakers in the range from 8″ to 15″ in closed or open enclosures are also available. The frequency responses of common guitar cabinets show an uneven bandpass characteristic with many resonances at mid frequencies. Simulations can be done by impulse response measurements of the loudspeaker and cabinet combination. The nonlinear behavior of a loudspeaker cabinet was analyzed and modeled in [YBK08].

As well as the discussed topics, the influence of the power supply with valve rectifier [zL97, pp. 51–54] is claimed to be of importance. A soft reduction of the power supply voltage occurs when in a high-power operation short transients need a high current. This power supply effect leads to a soft clipping of the audio signal.

A proposal for tube simulation by using asymmetrical clipping [Ben97] is given by

$$f(x) = \begin{cases} \frac{x-Q}{1-e^{-dist\cdot(x-Q)}} + \frac{Q}{1-e^{dist\cdot Q}}, & Q \neq 0, x \neq Q, \\ \frac{1}{dist} + \frac{Q}{1-e^{dist\cdot Q}}, & x = Q. \end{cases} \tag{4.13}$$

The underlying design parameters for the simulation of tube distortion are based on the mathematical model [Ben97] where no distortion should occur when the input level is low (the derivative of $f(x)$ has to be $f'(0) \approx 1$ and $f(0) = 0$). The static characteristic curve should perform clipping and limiting of large negative input values and be approximately linear for positive values. The result of Equation (4.13) is shown in Figure 4.26.

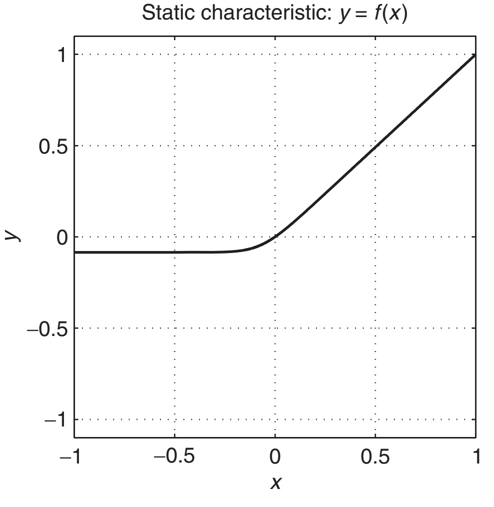

Figure 4.26 Static characteristic curve of asymmetric soft clipping for tube simulation $Q = -0.2$ and $dist = 8$.

The following M-file 4.4 performs Equation (4.13) from [Ben97]. To remove the dc component and to shape higher harmonics, additional lowpass and highpass filtering of the output signal is performed.

M-file 4.4 `(tube.m)`

```
function y=tube(x, gain, Q, dist, rh, rl, mix)
% function y=tube(x, gain, Q, dist, rh, rl, mix)
% Author: Bendiksen, Dutilleux, Zölzer

% y=tube(x, gain, Q, dist, rh, rl, mix)
% "Tube distortion" simulation, asymmetrical function
% x    - input
% gain - the amount of distortion, >0->
% Q    - work point. Controls the linearity of the transfer
%            function for low input levels, more negative=more linear
% dist - controls the distortion's character, a higher number gives
%            a harder distortion, >0
```

```
% rh   - abs(rh)<1, but close to 1. Placement of poles in the HP
%            filter which removes the DC component
% rl   - 0<rl<1. The pole placement in the LP filter used to
%            simulate capacitances in a tube amplifier
% mix  - mix of original and distorted sound, 1=only distorted
q=x*gain/max(abs(x));                      %Normalization
if Q==0
  z=q./(1-exp(-dist*q));
  for i=1:length(q)                        %Test because of the
    if q(i)==Q                             %transfer function's
      z(i)=1/dist;                         %0/0 value in Q
    end;
  end;
else
  z=(q-Q)./(1-exp(-dist*(q-Q)))+Q/(1-exp(dist*Q));
  for i=1:length(q)                        %Test because of the
    if q(i)==Q                             %transfer function's
      z(i)=1/dist+Q/(1-exp(dist*Q));    %0/0 value in Q
    end;
  end;
end;
y=mix*z*max(abs(x))/max(abs(z))+(1-mix)*x;
y=y*max(abs(x))/max(abs(y));
y=filter([1 -2 1],[1 -2*rh rh^2],y);    %HP filter
y=filter([1-rl],[1 -rl],y);                        %LP filter
```

Short-time FFTs (waterfall representation) of this algorithm applied to a 1 kHz sinusoid are shown in Figure 4.27. The waterfall representation shows strong even-order harmonics and also odd-order harmonics.

Digital amp modeling

New guitar amplifier designs with digital preamplifiers are based on digital modeling technology, featuring the simulation of classic valve amplifiers. Available are combos, amplifier heads or separated preamplifiers, mostly providing a wide variety of amplifier models in combination with additional effects. Besides these guitar amplifiers, software models are very popular which can be used as plug-ins together with a recording software. The principles of established modeling techniques will be introduced in Section 12.3.

Musical applications

Musical applications of valve amplifiers can be found on nearly every recording featuring guitar tracks. Ambassadors of innovative guitar players from the blues to the early rock period are B. B. King, Albert King and Chuck Berry, who mainly used valve amplifiers for their warm and soft sound. Representatives of the classic rock period are Jimi Hendrix [m-Hen67a, m-Hen67b, m-Hen68], Eric Clapton [m-Cla67], Jimmy Page [m-Pag69], Ritchie Blackmore [m-Bla70], Jeff Beck [m-Bec89] and Carlos Santana [m-San99]. All make extensive use of valve amplification and special guitar effect units. There are also players from the new classic period like Eddie van Halen, Steve Ray Vaughan and Steve Morse up to the new guitar heroes such as Steve Lukather, Joe Satriani, Gary Moore, Steve Vai and Paul Gilbert, who are using effect devices together with valve amplifiers.

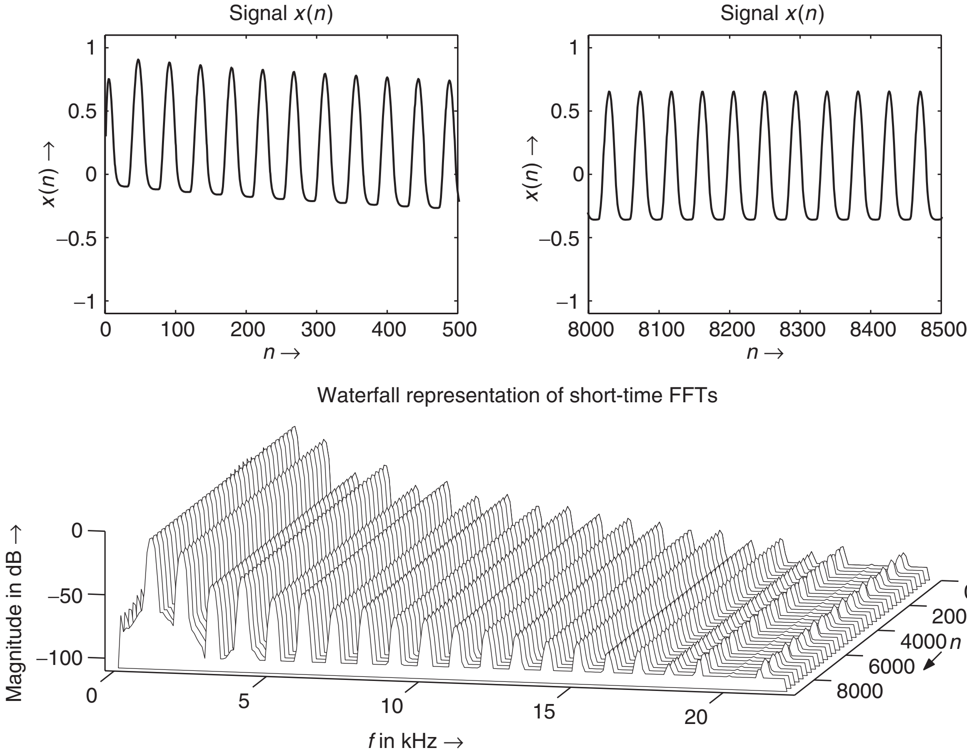

Figure 4.27 Short-time FFTs (waterfall representation) of asymmetrical soft clipping.

4.3.2 Overdrive, distortion and fuzz

Introduction

As pointed out in the section on valve simulation, the distorted electric guitar is a central part of rock music. In addition to the guitar amplifier as a major sound effect device, several stomp boxes (foot-operated pedals) have been used by guitar players for the creation of their typical guitar sound. Guitar heroes like Jimi Hendrix have made use of several small analog effect devices to achieve their unmistakable sound. Most of these effect devices have been used to create higher harmonics for the guitar sound in a faster way and at a much lower sound level compared to valve amplifiers. In this context terms like overdrive, distortion and fuzz are used. Several definitions of these terms for musical applications especially in the guitar player world are available [Kee00]. For our discussion we will define *overdrive* as a first state where a nearly linear audio effect device at low input levels is driven by higher input levels into the nonlinear region of its characteristic curve. The operating region is in the linear region as well as in the nonlinear region, with a smooth transition. The main sound characteristic is of course from the nonlinear part. Overdrive has a warm and smooth sound. The second state is termed *distortion*, where the effects device mainly operates in the nonlinear region of the characteristic curve and reaches the upper input level, where the output level is fixed to a maximum level. Distortion covers a wide tonal area starting beyond tube warmth to buzz saw effects. All metal and grunge sounds fall into this category. The operating status of *fuzz* is represented by a completely nonlinear behavior of the effect device with a sound characterized by the guitar player terms “harder” and “harsher” than distortion. The fuzz effect is generally used on single-note lead lines.

Signal processing

Overdrive. For overdrive simulations a soft clipping of the input values has to be performed. One possible approach for a soft saturation nonlinearity [Sch80] is given by

$$f(x) = \begin{cases} 2x & \text{for} \quad 0 \le x \le 1/3 \\ \frac{3-(2-3x)^2}{3} & \text{for} \quad 1/3 \le x \le 2/3 \\ 1 & \text{for} \quad 2/3 \le x \le 1. \end{cases} \tag{4.14}$$

The static input to output relation is shown in Figure 4.28. Up to the threshold of $1/3$ the input is multiplied by two and the characteristic curve is in its linear region. Between input values of $1/3$ up to $2/3$, the characteristic curve produces a soft compression described by the middle term of Equation (4.14). Above input values of $2/3$ the output value is set to one. The corresponding M-file 4.5 for overdrive with symmetrical soft clipping is shown next.

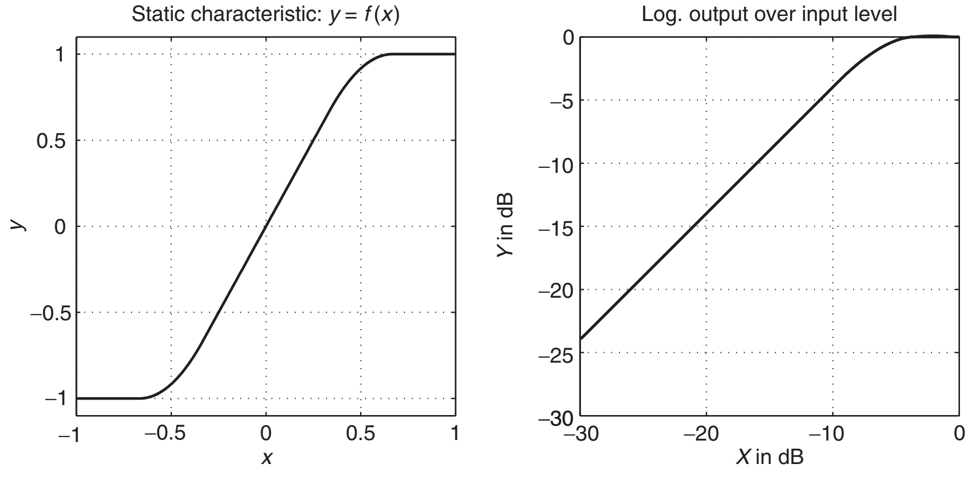

Figure 4.28 Static characteristic curve of symmetrical soft clipping (right part shows logarithmic output value versus input value).

M-file 4.5 (symclip.m)

```
function y=symclip(x)
% function y=symclip(x)
% Author: Dutilleux, Zölzer
% "Overdrive" simulation with symmetrical clipping
% x    - input
N=length(x);
th=1/3; % threshold for symmetrical soft clipping
        % by Schetzen Formula
for i=1:1:N,
   if abs(x(i))< th, y(i)=2*x(i);end;
   if abs(x(i))>=th,
      if x(i)> 0, y(i)=(3-(2-x(i)*3).^2)/3; end;
      if x(i)< 0, y(i)=-(3-(2-abs(x(i))*3).^2)/3; end;
   end;
   if abs(x(i))>2*th,
```

```
      if x(i)> 0, y(i)=1;end;
      if x(i)< 0, y(i)=-1;end;
   end;
end;
```

Figure 4.29 shows the waveforms of a simulation with the above-described characteristic curve and a decaying sinusoid of 1 kHz. In the upper left plot the first part of the output signal is shown with high signal levels, which corresponds to the saturated part of the characteristic curve. The tops and bottoms of the sinusoid run with a soft curve towards the saturated maximum values. The upper right plot shows the output signal where the maximum values are in the soft clipping region of the characteristic curve. Both the negative and the positive top of the sinusoid are rounded in their shape. The lower waterfall representation shows the entire decay of the sinusoid down to −12 dB. Notice the odd order harmonics produced by this nonlinear symmetrical characteristic curve, which appear in the nonlinear region of the characteristic curve and disappear as soon as the lower threshold of the soft compression is reached. The prominent harmonics are the third and the fifth harmonic. The slow increase or decrease of higher harmonics is the major property of symmetrical soft clipping. As soon as simple hard clipping without a soft compression is performed, higher harmonics appear with significantly higher levels (see Figure 4.30). The discussion of overdrive and distortion has so far only considered the creation of harmonics for a single sinusoid as the input signal. Since a single guitar tone itself consists of the fundamental frequency plus all odd- and even-order harmonics, it is always the sum of sinusoids that is processed by the nonlinearity. The nonlinearity also produces sum and difference frequencies.

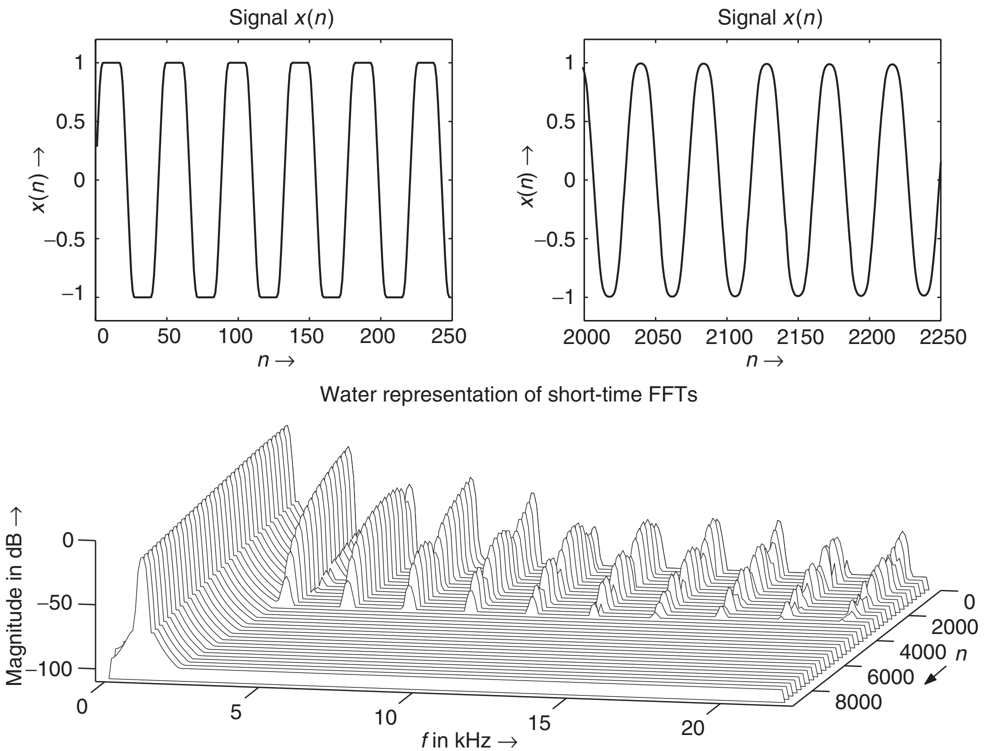

Figure 4.29 Short-time FFTs (waterfall representation) of symmetrical soft clipping for a decaying sinusoid of 1 kHz.

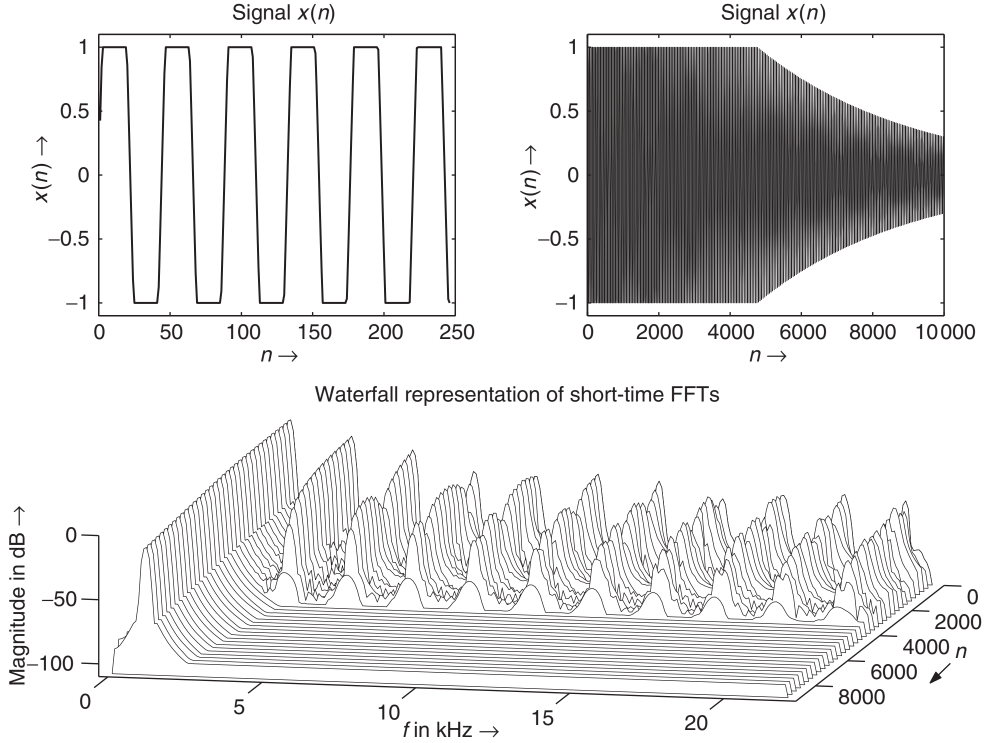

Figure 4.30 Short-time FFTs (waterfall representation) of symmetrical hard clipping for a decaying sinusoid of 1 kHz.

Distortion. A nonlinearity suitable for the simulation of distortion [Ben97] is given by

$$f(x) = \mathrm{sgn}(x)\left(1 - \mathrm{e}^{-|x|}\right). \tag{4.15}$$

The M-file 4.6 for performing Equation (4.15) is shown next.

M-file 4.6 (expdist.m)

```
function y=expdist(x, gain, mix)
% function y=expdist(x, gain, mix)
% Author: Bendiksen, Dutilleux, Zölzer, Dempwolf
% y=expdist(x, gain, mix)
% Distortion based on an exponential function
% x    - input
% gain - amount of distortion, >0
% mix  - mix of original and distorted sound, 1=only distorted
q=x*gain;
z=sign(q).*(1-exp(-abs(q)));
y=mix*z+(1-mix)*x;
```

The static characteristic curve is illustrated in Figure 4.31 and short-time FFTs of a decaying 1 kHz sinusoid are shown in Figure 4.32.

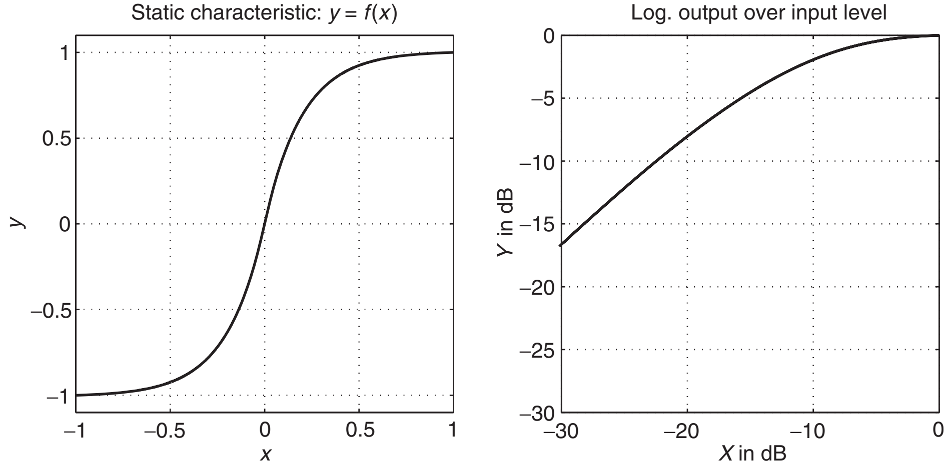

Figure 4.31 Static characteristic curve of exponential distortion.

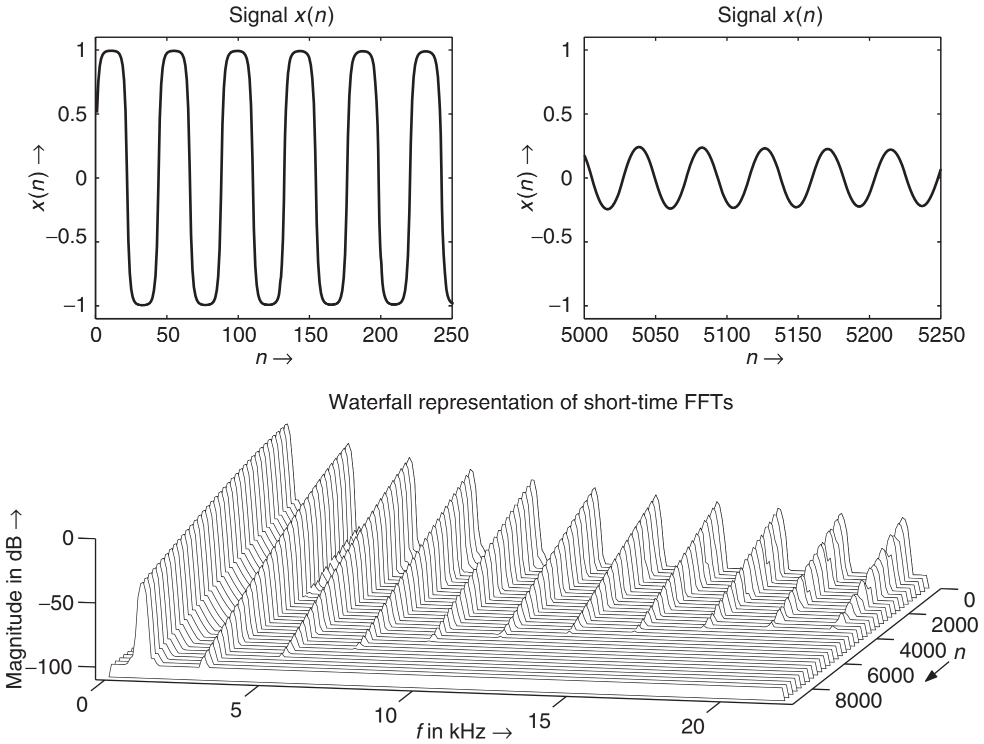

Figure 4.32 Short-time FFTs (waterfall representation) of exponential distortion.

Fuzz. We have already discussed the behavior of triode valves in the previous section which produce an asymmetrical overdrive. One famous representative of asymmetrical clipping is the *Fuzz Face* [Kee98a], which was used by Jimi Hendrix. The basic analog circuit is shown in Figure 4.33 and consists only of a few components with two transistors in a feedback arrangement.

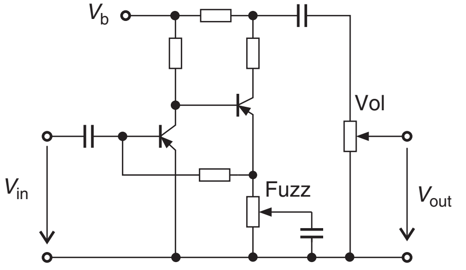

Figure 4.33 Analog circuit of Fuzz Face.

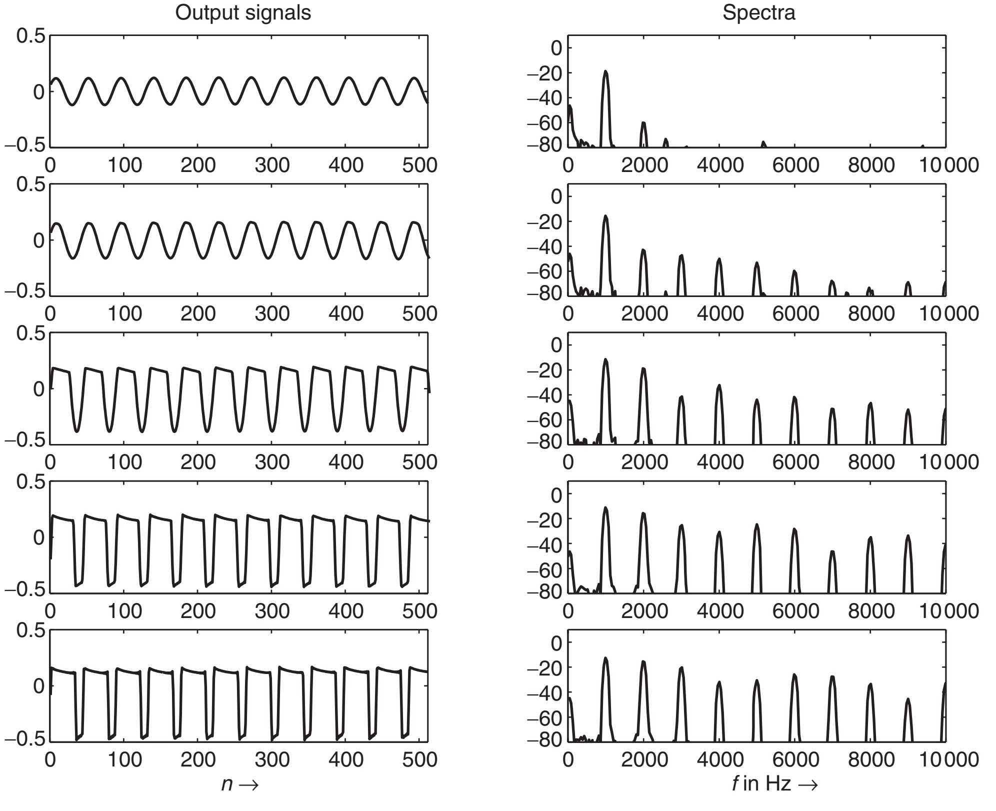

Figure 4.34 Signals and corresponding spectra of Fuzz Face.

The output signals for various input levels are presented in Figure 4.34 in conjunction with the corresponding spectra for a 1 kHz sinusoid. The upper plots down to the lower plots show an increasing input level. For low-level input signals the typical second harmonic of a triode valve can be noticed, although the time signal shows no distortion components. With increasing input level the second harmonic and all even-order harmonics as well as odd-order harmonics appear. The asymmetrical clipping produces enhanced even-order harmonics, as shown in the third row of Figure 4.34. Notice that only the top of the positive maximum values are clipped. As soon as the input level further increases, the negative part of the waveform is clipped. The negative clipping level is lower than the positive clipping value and so asymmetrical clipping is performed.

Short-time Fourier transforms (in waterfall representation) for an increasing 1 kHz sinusoid, together with two waveforms are shown in Figure 4.35.

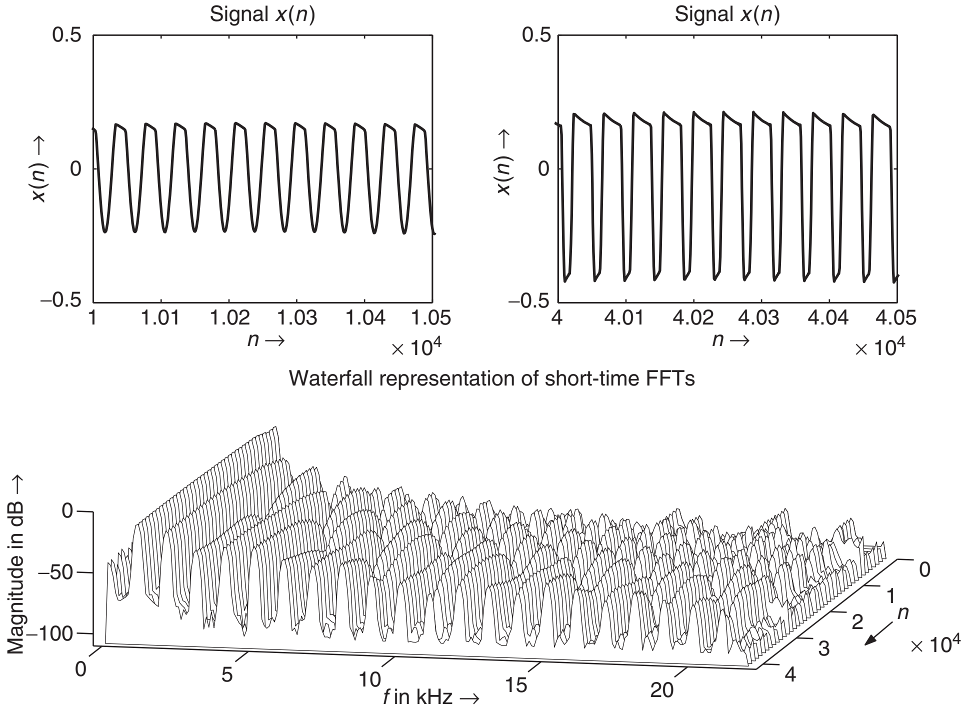

Figure 4.35 Short-time FFTs (waterfall representation) of Fuzz Face for an increasing 1 kHz sinusoid. The upper plots show segments of samples from the complete analyzed signal.

Musical applications

There are a lot of commercial stomp effects for guitarists on the market. Some of the most interesting distortion devices for guitars are the *Fuzz Face* which performs asymmetrical clipping towards symmetrical soft clipping and the *Tube Screamer* [Kee98b], which performs symmetrical soft clipping. The *Fuzz Face* was used by Jimi Hendrix and the *Tube Screamer* by Stevie Ray Vaughan. They both offer classical distortion and are well known because of their famous users. It is difficult to explain the sound of a distortion unit without listening personally to it. The technical specifications for the sound of distortion are missing, so the only way to choose a distortion effect is by a comparative listening test. Investigations about the sound coloration caused by typical distortion or overdrive effects can be found in [MM05, DHMZ09].

4.3.3 Harmonic and subharmonic generation

Introduction

Harmonic and subharmonic generation are performed by simple analog or digital effect devices, which should produce an octave above and/or an octave below a single note. Advanced techniques to achieve pitch shifting of instrument sounds will be introduced in Chapter 6. Here, we will focus on simple techniques, which lead to the generation of harmonics and subharmonics.

Signal processing

The signal-processing algorithms for harmonic and subharmonic generation are based on simple mathematical operations like absolute value computation and counting of zero crossings, as shown in Figure 4.36 when an input sinusoid has to be processed (first row shows time signal and corresponding spectrum).

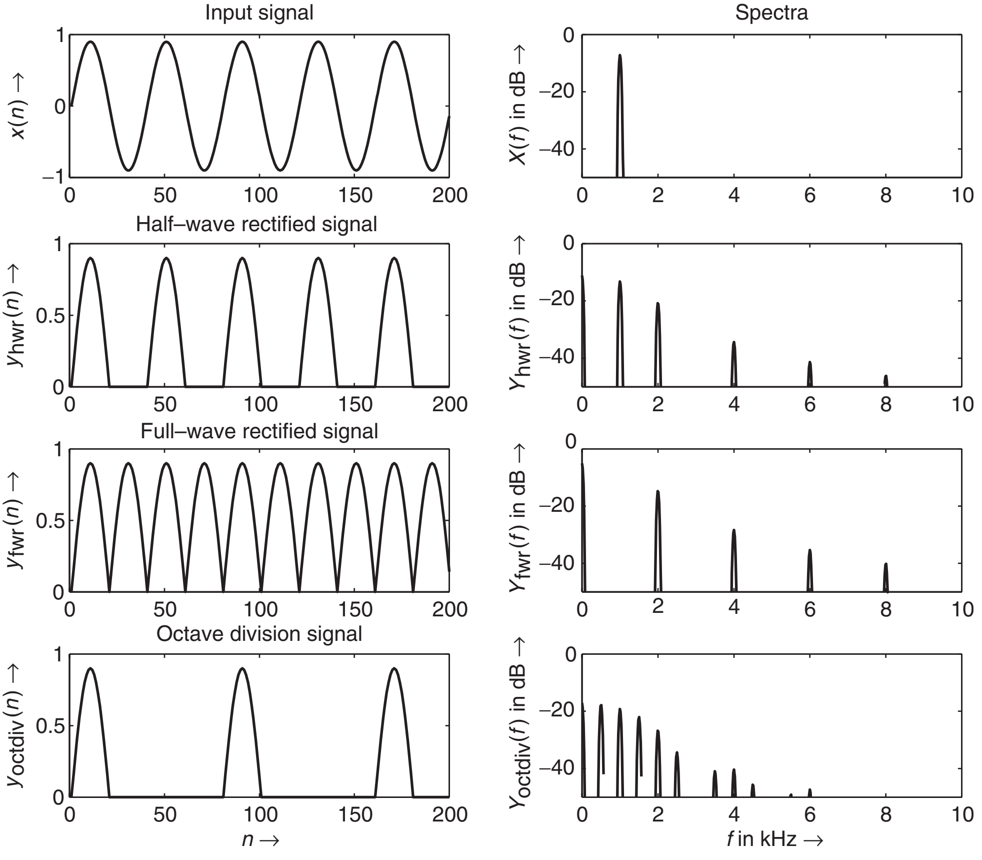

Figure 4.36 Signals and corresponding spectra of half-wave rectification, full-wave rectification and octave division.

The second row of Figure 4.36 demonstrates half-wave rectification, where positive values are kept and negative values are set to zero. This operation leads to the generation of even-order harmonics. Full-wave rectification, where the absolute value is taken from the input sequence, leads to even-order harmonics, as shown in the third row of Figure 4.36. Notice the absence of the fundamental frequency. If a zero crossing counter is applied to the half-wave or the full-wave rectified signal, a predefined number of positive wave parts can be set to zero to achieve the signal in the last row of Figure 4.36. This signal has a fundamental frequency which is one octave lower than the input frequency in the first row of the figure, but also shows harmonics of this new fundamental frequency. If appropriate lowpass filtering is applied to such a signal, only the fundamental frequency can be obtained, which is then added to the original input signal.

Musical applications

Harmonic and subharmonic generation is mostly used on single-note lead lines, where an additional harmonic or subharmonic frequency helps to enhance the octave effect. Harmonic generators can be found in stomp boxes for guitar or bass guitar and appear under the name *octaver*. Subharmonic generation is often used for solo and bass instruments to give them an extra bass boost or simply a fuzz bass character.

4.3.4 Tape saturation

Introduction and musical application

The special sound characteristic of analog tape recordings has been acknowledged by a variety of producers and musicians in the field of rock music. They prefer doing multi track recordings with analog tape-based machines and use the special physics of magnetic tape recording as an analog effects processor for sound design. One reason for their preference for analog recording is the fact that magnetic tape goes into distortion gradually [Ear76, pp. 216-218] and produces those kinds of harmonics which help special sound effects on drums, guitars and vocals.

Signal processing

Tape saturation can be simulated by the already introduced techniques for valve simulation. An input-level-derived weighting curve is used for generating a gain factor, which is used to compress the input signal. A variety of measurements of tape recordings can help in the design of such processing devices. An example of the input/output behavior is shown in Figure 4.37 and a short-time FFT of a sinusoid input signal in Figure 4.38 illustrates a tape saturation algorithm. For low-level inputs the transfer characteristic is linear without any distortions. A smooth soft compression simulates the gradually increasing distortion of magnetic tape.

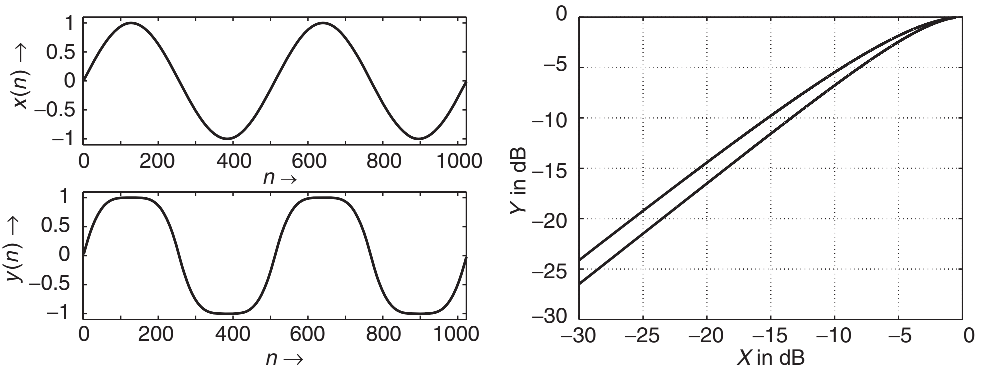

Figure 4.37 Tape saturation: input and output signal (left) and characteristic curve.

4.4 Exciters and enhancers

4.4.1 Exciters

Introduction

An exciter is a signal processor that emphasizes or de-emphasizes certain frequencies in order to change a signal's timbre. An exciter increases brightness without necessarily adding equalization.

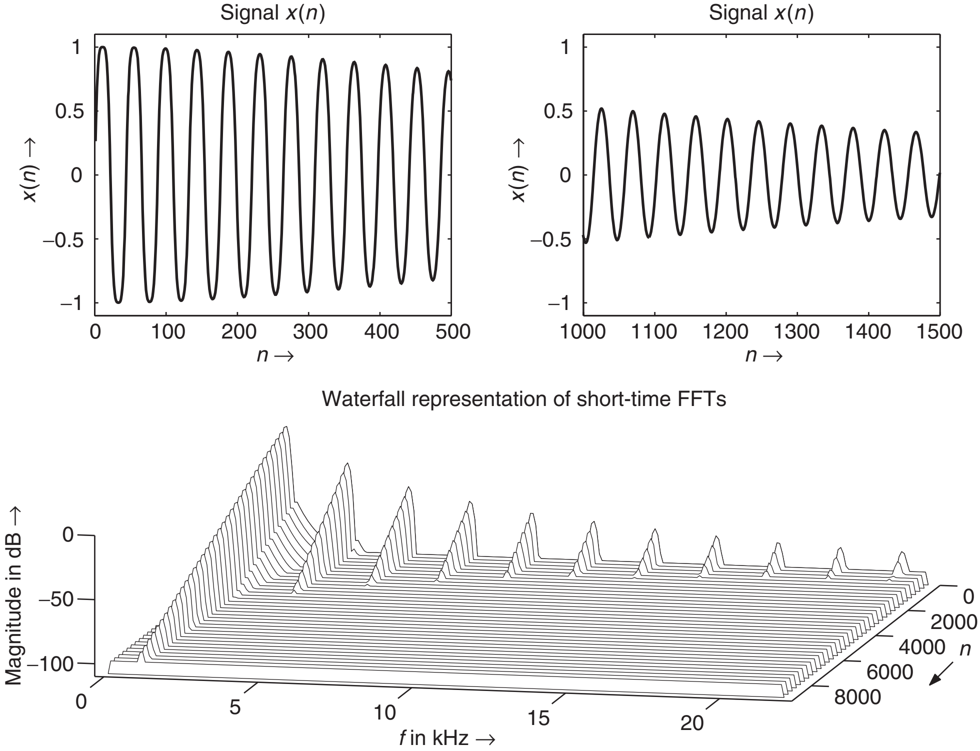

Figure 4.38 Tape saturation: short-time FFTs (waterfall representation) for decaying sinusoid of 1 kHz.

The result is a brighter, "airier" sound without the stridency that can sometimes occur by simply boosting the treble. This is often accomplished with subtle amounts of high-frequency distortion, and sometimes by phase shifting. Usually there will only be one or two parameters, such as exciter mix and exciter frequency. The former determines how much "excited" sound gets added to the straight sound, and the latter determines the frequency at which the exciter effect starts [Whi93, And95, Dic87, WG94].

This effect was discovered by the Aphex company and "Aural Exciter" is a trademark of this company. The medium and treble parts of the original signal are processed by a nonlinear circuit that generates higher overtones. These components are then mixed to some extent to the original signal. A compressor at the output of the nonlinear element makes the effect dependent on the input signal. The initial part of percussive sounds will be more enriched than the following part, when the compressor limits the effect depth. The enhanced imaging or spaciousness is probably the result of the phase rotation within the filter [Alt90].

Signal processing

Measurement results of the APHEX Aural Exciter are shown in Figures 4.39 and 4.40, where the generation of a second harmonic is clearly visible. The input signal is a chirp signal with an increasing frequency up to 5 kHz. Signal-processing techniques to achieve the effect have already been discussed in the previous sections. The effect is created in the side-chain path and is mixed with the input signal.

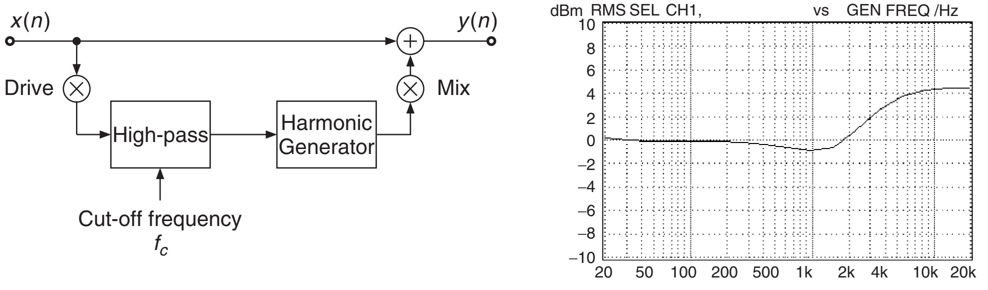

Figure 4.39 Block diagram of the psycho-acoustic equalizer APHEX Aural Exciter and frequency response.

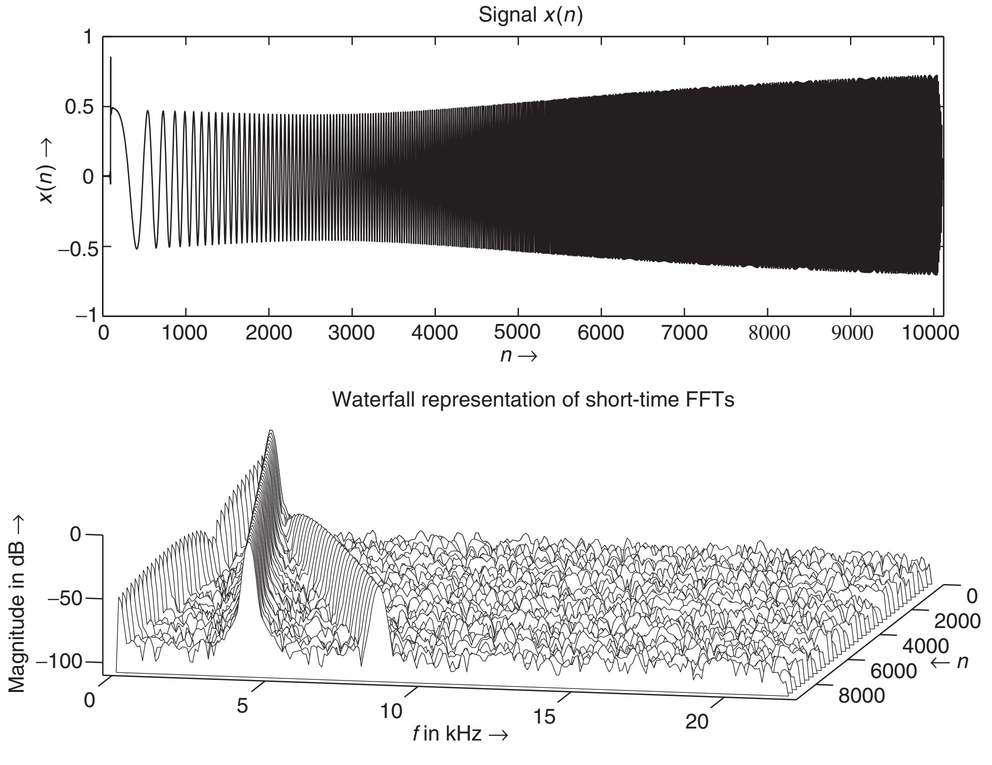

Figure 4.40 Short-time FFTs (waterfall representation) of a psycho-acoustic equalizer.

Musical applications

The applications of this effect are widespread and range from single-instrument enhancement to enhancement of mix buses and stereo signals. The effect increases the presence and clarity of a single instrument inside a mix and helps to add natural brightness to stereo signals. Applied to vocals and speech the effect increases intelligibility. Compared to equalizers the sound level is only increased slightly. The application of this effect only makes sense if the input signal lacks high-frequency content.

4.4.2 Enhancers

Introduction

Enhancers are signal processors which combine equalization together with nonlinear processing. They perform equalization according to the fundamentals of psychoacoustics [ZF90] and introduce a small amount of distortion in a just noticeable manner. An introduction to the ear's own nonlinear distortions, sharpness, sensory pleasantness and roughness can be also be found in [ZF90].

Signal processing

As an example of this class of devices the block diagram and the frequency response of the SPL vitalizer are shown in Figure 4.41. This effect processor has also a side-chain path which performs equalization with a strong bass enhancement, a mid-frequency cut and a high-frequency boost. The short-time FFT of the output signal when a chirp input signal is applied is shown in Figure 4.42. The resulting waterfall representation clearly shows higher harmonics generated by this effect processor.

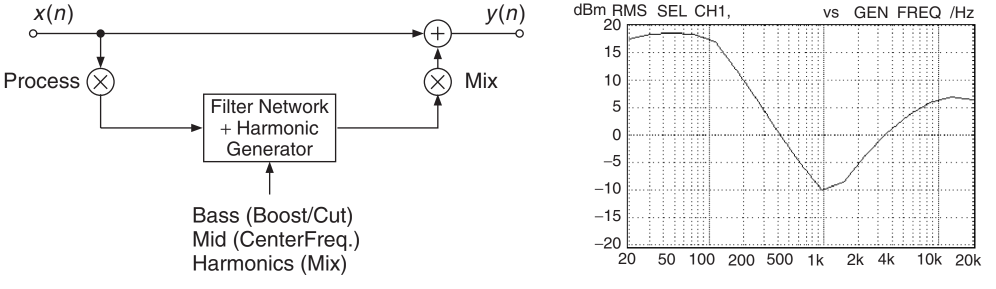

Figure 4.41 Block diagram of the psycho-acoustic equalizer SPL Vitalizer and frequency response.

Further refinements of enhancers can be achieved through multiband enhancers which split the input signal into several frequency bands. Inside each frequency band nonlinear processing plus filtering is performed. The output signals of each frequency band are weighted and summed up to form the output signal (see Figure 4.43).

Musical applications

The main applications of such effects are single-track processing as a substitute for the equalizers inside the input channels of mixing consoles and processing of final mixes. The side-chain processing allows the subtle mix of the effects signal together with the input signal. Further applications are stereo enhancement for broadcast stations and sound reinforcement.

4.5 Conclusion

The most challenging tools for musicians and sound engineers are nonlinear processors such as dynamics processors, valve simulators and exciters. The successful application of these audio processors depends on the appropriate control of these devices. A variety of interactive control parameters influences the resulting sound quality.

The primary purpose of this chapter is to enable the reader to attain a fundamental understanding of different types of nonlinear processors and the special properties of nonlinear operations applied

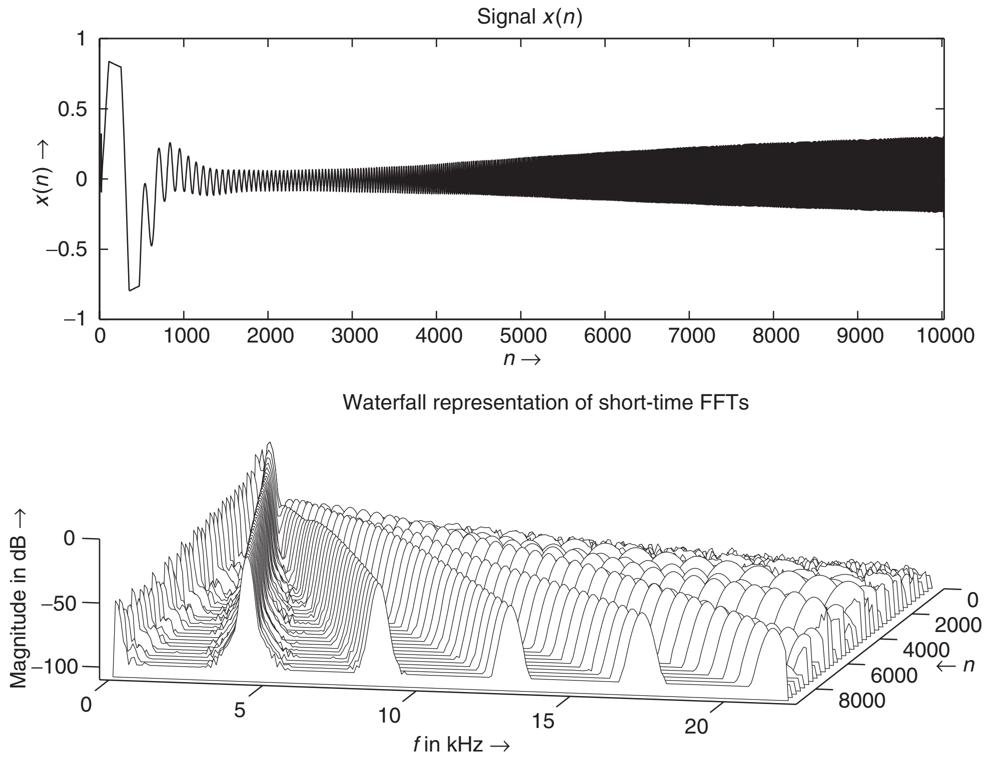

Figure 4.42 Short-time FFTs (waterfall representation) of psycho-acoustic equalizer SPL Vitalizer.

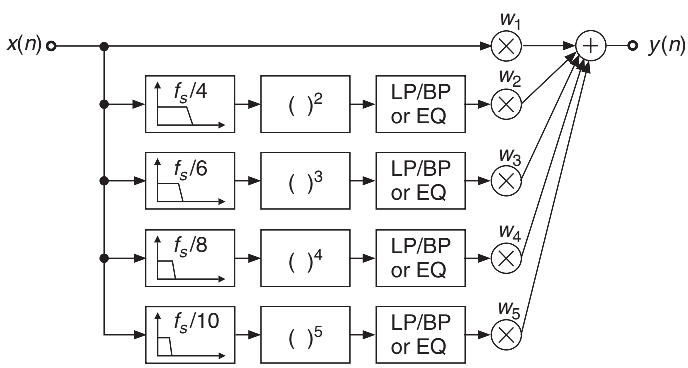

Figure 4.43 Multiband enhancer with nonlinear processing in frequency bands.

to the audio signal. Dynamics processors need a careful consideration of the interaction of thresholds and time constants to achieve sonic purity and avoid harmonic distortion. On the other hand, nonlinear processors are used for the simulation of valve amplifiers or nonlinear audio systems, where a special kind of nonlinearity provides a sound distortion with an accepted sound characteristic. We have presented the basics of nonlinear modeling and focused on the combination of filters and nonlinearities. Several applications demonstrate the importance of nonlinear processors. The basic building blocks of the previous chapters such as filters, delays and modulators/demodulators are particularly useful for exploring new, improved nonlinear processors.

Sound and music

[m-Bec89] Jeff Beck: *Guitar Shop*. 1989.
[m-Bla70] Deep Purple: *Deep Purple in Rock*. 1970.
[m-Cla67] Cream: *Disraeli Gears*. 1967.
[m-Hen67a] Jimi Hendrix: *Are You Experienced?* 1967.
[m-Hen67b] Jimi Hendrix: *Axis: Bold as Love*. 1967.
[m-Hen68] Jimi Hendrix: *Electric Ladyland*. 1968.
[m-Pag69] Led Zeppelin: *I/II/III/IV* . 1969-1971.
[m-San99] Santana: *Supernatural*. 1999.

References

[Alt90] S. R. Alten. *Audio in Media*. Wadsworth, 1990.
[And95] C. Anderton. *Multieffects for Musicians*. Amsco Publications, 1995.
[Bar98] E. Barbour. The cool sound of tubes. *IEEE Spectrum*, pp. 24–32, 1998.
[Ben97] R. Bendiksen. Digitale Lydeffekter. *Master's thesis*, Norwegian University of Science and Technology, 1997.
[Bev77] M. Bevelle. Compressors and limiters. *Studio Sound*, p. 32, 1977.
[CVAT01] E. Czerwinski, A. Voishvillo, S. Alexandrov and A. Terekhov. Multitone testing of sound system components - some results and conclusions. Part 1: History and theory. *J. Audio Eng. Soc*., 49(11), 2001. 1011–1048
[DHMZ09] K. Dempwolf, M. Holters, S. Möller and U. Zölzer. The influence of small variations in a simplified guitar amplifier model. In *Proc. of the 12th Int. Conf. on Digital Audio Effects (DAFx-09)*, Como, Italy, Sept. 1–4 2009.
[Dic87] M. Dickreiter. *Handbuch der Tonstudiotechnik, Band I und II*. K.G. Saur, 1987.
[Doy93] M. Doyle. *The History of Marshall*. Hal Leonhard Publishing Corporation, 1993.
[Dut96] P. Dutilleux. Verstärkung der Differenztöne (f2-f1). In *Bericht der 19. Tonmeistertagung Karlsruhe*, *Verlag K.G. Saur*, pp. 798–806, 1996.
[Ear76] J. Eargle. *Sound Recording*. Van Nostrand, 1976.
[Fli93] R. Fliegler. *AMPS! The Other Half of Rock"n'Roll*. Hal Leonhard Publishing Corporation, 1993.
[Fra97] W. Frank. *Aufwandsarme Modellierung und Kompensation nichtlinearer Systeme auf Basis von Volterra-Reihen*. PhD thesis, University of the Federal Armed Forces Munich, Germany, 1997.
[FUB$^+$98] A. Farina, E. Ugolotti, A. Bellini, G. Cibelli and C. Morandi. Inverse numerical filters for linearisation of loudspeaker's responses. In *Proc. DAFX-98 Digital Audio Effects Workshop*, pp. 12–16, Barcelona, November 1998.
[Ham73] R. O. Hamm. Tubes versus transistors – is there an audible difference ? *J. Audio Eng. Soc*., 21(4): 267–273, May 1973.
[Hul97] R. Hulse. A different way of looking at compression. *Studio Sound*, 1997.
[Izh10] R. Izhaki. *Mixing Audio: Concepts, Practices and Tools*. Focal Press, 2010.
[Kai87] A. J. M. Kaizer. Modeling of the nonlinear response of an electrodynamic loudspeaker by a volterra series expansion. *J. Audio Eng. Soc*., 35(6): 421–433, 1987.
[Kat07] B. Katz. *Mastering Audio*. Focal Press, 2007.
[Kee98a] R. G. Keen. The technology of the fuzz face. [online] www.geofex.com, 1998.
[Kee98b] R. G. Keen. The technology of the tube screamer. [online] www.geofex.com, 1998.
[Kee00] R. G. Keen. A musical distortion primer. [online] www.geofex.com, 1993–2000.
[Kli98] W. Klippel. Direct feedback linearization of nonlinear loudspeaker systems. *J. Audio Eng. Soc*., 46(6): 499–507, 1998.
[Kue05] R. Kuehnel. *Circuit Analysis of a Legendary Tube Amplifier: The Fender Bassman 5F6-A*, 2nd edition. Pentode Press, 2005.
[Lun07] T. Lund. Level and distortion in digital broadcasting. *EBU Technical Review*, April 2007.

[McN84] G. W. McNally. Dynamic range control of digital audio signals. *J. Audio Eng. Soc.*, 32(5): 316–327, May 1984.

[MM05] A. Marui and W. L. Martens. Timbre of nonlinear distortion effects: Perceptual attributes beyond sharpness. In *Proc. of Conf. Interdisc. Musicol. (CIM05)*, Montréal, Canada, March 2005.

[Nie00] S. Nielsen. Personal communication. TC Electronic A/S, 2000.

[Orf96] S. J. Orfanidis. *Introduction to Signal Processing*. Prentice-Hall, 1996.

[PD93] D. Peterson and D. Denney. *The VOX Story*. The Bold Strummer Ltd., 1993.

[Rat95] L. Ratheiser. *Das Große Röhrenhandbuch*. Reprint. Franzis-Verlag, 1995.

[RCA59] RCA. *Receiving Tube Manual. Technical Series RC-19*. Radio Corporation of America, 1959.

[RH96] M. J. Reed and M. O. Hawksford. Practical modeling of nonlinear audio systems using the volterra series. In *Proc. 100th AES Convention*, Preprint 4264, 1996.

[Sal02] T. Salter. *Talking Shop with Randall Smith of Mesa Boogie, Interview*. Musicians Hotline, May/June 2002.

[Sch80] M. Schetzen. *The Volterra and Wiener Theories of Nonlinear Systems*. Robert Krieger Publishing, 1980.

[SZ99] J. Schattschneider and U. Zölzer. Discrete-time models for nonlinear audio systems. In *Proc. DAFX-99 Digital Audio Effects Workshop*, pp. 45–48, Trondheim, December 1999.

[WG94] M. Warstat and T. Görne. *Studiotechnik – Hintergrund und Praxiswissen*. Elektor-Verlag, 1994.

[Whi93] P. White. *L'enregistrement Créatif, Effets et Processeurs, Tomes 1 et 2*. Les Cahiers de l'ACME, 1993.

[YBK08] D. Yeh, B. Bank and M. Karjalainen. Nonlinear modeling of a guitar amplifier cabinet. In *Proc. DAFX-08 Conf. Digital Audio Effects*, pp. 89–96, Espoo, Finland, Sept 1–4 2008.

[ZF90] E. Zwicker and H. Fastl. *Psychoacoustics*. Springer-Verlag, 1990.

[zL97] R. zur Linde. *Röhrenverstärker*. Elektor-Verlag, 1997.

[Zöl05] U. Zölzer. *Digital Audio Signal Processing*, 2nd edition. John Wiley & Sons, Ltd, 2005.