10.8.5 Distortion devices

In communication engineering, two types of distortion are specified: linear and non-linear distortions. Linear distortion is created in systems with a frequency-dependent transfer-characteristic, i.e. for example in sound filters♣ (also called tone controls, or EQ). Non-linear distortion occurs in a non-linear system. For a system to operate linearly, it needs to have proportionality, lack of sources, and the possibility for superpositioning. If only one of these conditions is missing, the respective system is non-linear.

- “Proportionality” means that an n-fold increase of the input signal will result in an n-fold increase of the output signal (doubling the input results in doubling of the output).
- “lack of sources” means that without an input signal, the output signal will be zero.
- “possibility for superpositioning” means that a transformation of a sum of signals effected by the system will correspond to the sum of the transformed individual signals :$T(x+y) = T(x)+T(y)$.

The description will be simpler if linear and non-linear systems occur strictly separately, i.e. if every sub-system will perform only linear or only non-linear mapping. A non-linear system that does not cause any linear distortion includes no memory – this is because only devices including memory (inductances, capacitances) generate frequency-dependent resistances and thus create a frequency-dependent transmission. In a **memory-free system**, the output signal will therefore not depend in any way on the past input signal but exclusively on the input signal occurring at the very same moment. The transmission characteristics can be described via a characteristic $y(x)$. For non-linear behavior to be present, we need to have a curved characteristic (strictly speaking, an offset also introduces non-linearity).

The amplifying elements (tube, transistor) used in guitar amplifiers all feature a curved characteristic, and therefore every guitar amplifier operates as a non-linear device. According to the rules of classical amplifier technology, these non-linearities are supposed to be as small as possible, and therefore negative-feedback circuits reduce the gain and at the same time perform a linearization. Many guitar players were satisfied with the resulting so-called “clean” sound, but some forced non-linear distortion by overdriving their amplifiers, creating “crunch”, “distortion”, or “fuzz”. In many amps this required using their full power and thus very high loudness – but sometimes even at maximum gain, the resulting non-linear distortion was not pronounced enough. This is the reason why additional devices for the generation of non-linear distortion were created under various monikers: fuzz-box, distortion-pedal, overdrive … Before long, the effect was not limited to additional devices: with an increasing number of tubes, guitar amplifiers themselves soon offered possibilities to control the desired degree of distortion.

The distortion devices described in the following are systems that add **non-linear distortion** to the guitar sound. Whether this happens in the amplifier itself or in a separate device is not distinguished to begin with. With the distortion, the guitar sound becomes fuller, sustaining, more shrill, more aggressive, buzzing, more alive – this is always depending on the chosen settings. There is not “the” distortion sound. The distortion also changes the **dynamics** of the signal in the sense that sustain is extended. Since practically all distortion devices have a degressive, limiting characteristic curve, any level-differences in the guitar signal are reduced and differences between loud and soft are evened out. The originally percussive guitar sound becomes steadier, and takes on some sound-characteristics of horns (saxophone, trumpet) or strings (cello).

♣ M. Zollner: Signalverarbeitung. Hochschule Regensburg, 2009.

There are two types of drive-situations for customary tube- and transistor-circuits: low drive-levels with weak distortion, and overdrive with strong distortion. In the left section of **Fig. 10.8.20**, we see the characteristic curve of an ECC83, and in the right-hand section the time-function of the sine-shaped input voltage and the distorted output voltage. For small drive-levels (up to about 1 V_S), there are only small differences between input- and output-voltage. For high drive-levels, we find a strong – and in this case asymmetrical – limiting of the signal ("clipping").

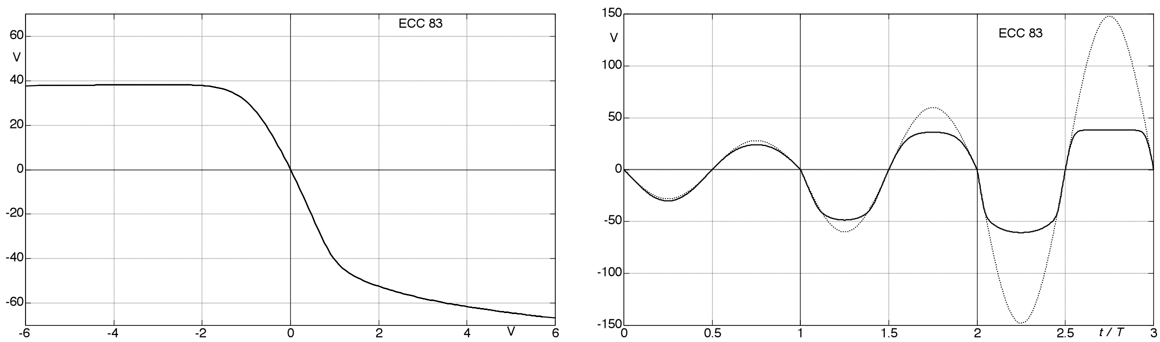

Fig. 10.8.20: Non-linear distortion for an ECC-83 (compare to Chapter 10.1.4).

Now, the human auditory system is not a sensor for directly analyzing the time-function. Rather, it detects in a first step the time-variant short-term spectrum (Chapters 8.2.4 & 8.6) to determine the sound color. If a sine-tone (e.g. 1 kHz) undergoes non-linear distortion, new spectral lines are created at the integer multiples of the fundamental frequency, i.e. at 2 kHz, 3 kHz, 4 kHz, etc. Distortion of a signal composed of several partials will generate sum- and difference-tones as multiples of the largest common divisor of all partials♣. If the sound of a single guitar string were of strictly harmonic content, distorting it would still result in a harmonic spectrum. The level and the phases of the partials would change, and so would the sound color, but the frequency of the partials would remain unchanged. However, the spectrum of every real string-vibration is spread **in-harmonically**, and it is here where we find the key to understanding the impact of a distortion device.

If, for example, a complex tone consisting of a 100-Hz- and a 202-Hz-partial undergoes 2^{nd}-order distortion, additional partials at 0 Hz, 102 Hz, 200 Hz, 302 Hz, 404 Hz are created. The 0-Hz-component may be ignored because it is DC that the circuit will not transmit further. The partials at 302 and 404 Hz will brighten up the sound if they are strong enough, but the main effect will be close to the primaries: the partials at 100 and 102 Hz will beat against each other, and so will the 200- and 202-Hz-partials. The non-linear distortion will, on one hand, enlarge the spectrum towards high frequencies (i.e. emphasize the treble more), and on the other hand the amplitudes of the primaries will start to fluctuate due to the beat effects: the sound becomes more lively. If not only 2^{nd}-order distortion occurs but higher-order distortion as well, a large number of additional partials is created and correspondingly many and possibly strong fluctuations. These fluctuations bring a kind of noise-character to the sound; we get an effect as if additional **noise** would be superimposed. Strictly speaking, noise in its usual definition belongs to the group of stochastic (random) signals, while the fluctuation of partials generated by non-linearity is not stochastic but determined. Since this special noise has a periodicity, it is called **pseudo-noise**.

♣ M. Zollner: Frequenzanalyse. Hochschule Regensburg, 2009.

Fig. 10.8.21 depicts the zero-symmetric characteristic curves that will be used for distortion in the following. They are odd-order functions, i.e. functions that can be expanded into a series containing only members of odd-order power (x, x^3, x^5, ...).

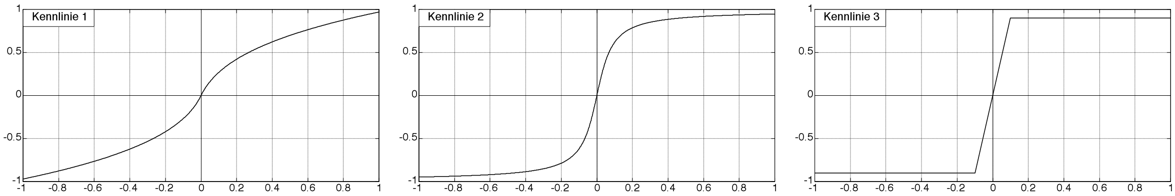

Fig. 10.8.21: Characteristic curves of the systems used for distortion.

Simple periodic functions are distorted in **Fig. 10.8.22** using characteristic curve 2 (an arctan-function). Distortion of the sine-signal (uppermost row in the figure) results in a spectrum with only odd-numbered harmonics. However, merely adding a second partial to the primary signal (2[nd] row in the figure) may generate even-numbered harmonics – although this is not generally the case, as demonstrated by the third row in the figure. Only a signal of half-wave anti-symmetry will, in its spectrum, contain exclusively odd-numbered harmonics. Such a half-wave anti-symmetric signal, if distorted via an odd-order characteristic curve, remain half-wave anti-symmetric, and will not gain any even-numbered harmonics. In the lowermost line of the figure, a signal of three partials is distorted. Due to the 2[nd] harmonic, this signal cannot be half-wave anti-symmetric, and therefore the spectrum of the distorted signal contains even-order harmonics as well.

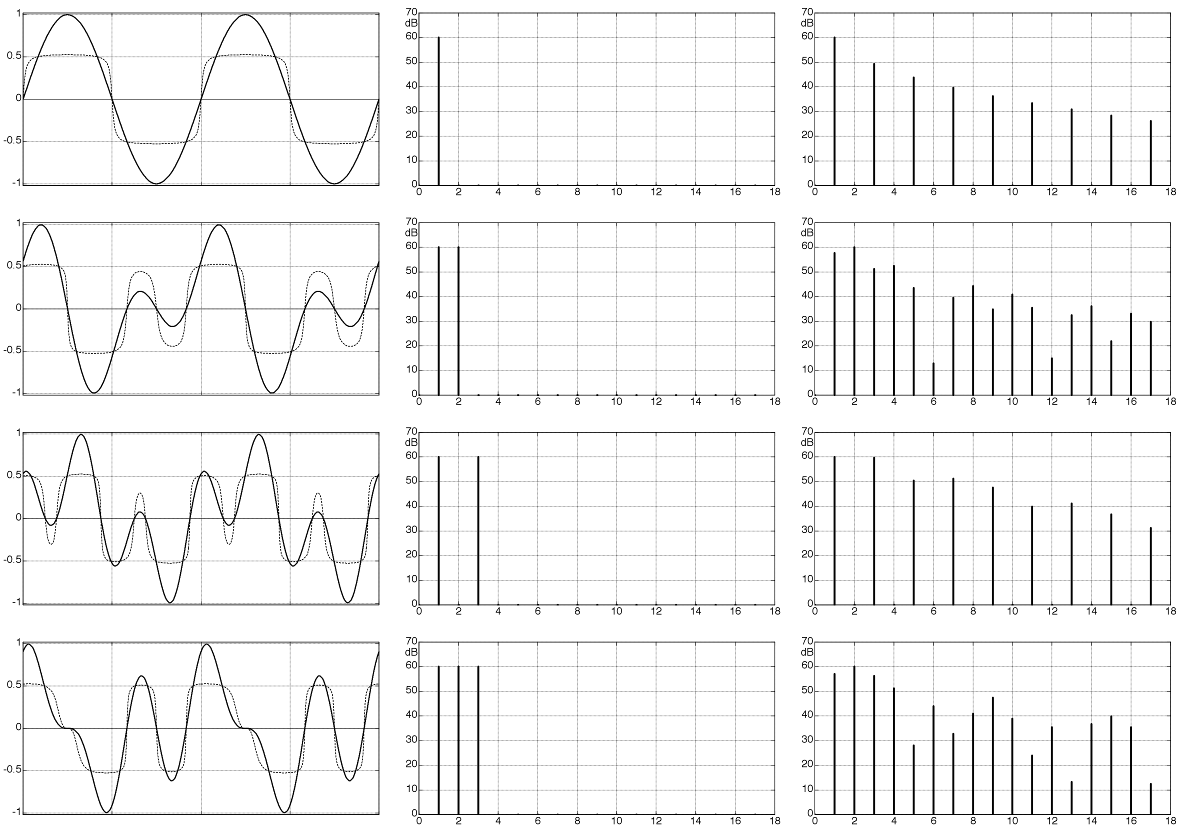

Fig. 10.8.22: Time-function of undistorted and distorted signal (left), spectrum of undistorted signal, spectrum of distorted signal (right). Time-functions and spectra individually scaled. Characteristic curve 2.

For the analyses discussed in the following, synthetic guitar sounds were used; their rules of construction were nonetheless derived from a real guitar. Synthetic sounds were used because all parameters of these sounds are known – which is not the case for real sounds. For the **harmonic** signal, 45 partials were added. The frequencies of these partials all had an integer relationship to the fundamental of 82 Hz, and their levels and decay time constants were taken from a real guitar tone (open E_2-string). The **inharmonic** signal was synthesized with the same levels and time constants, but the frequencies of the partials were slightly spread out according to the formula outlined in Chapter 1 (dispersion due to the bending stiffness of the strings). The spectrograms of the undistorted and distorted synthetic guitar sound are shown in **Fig. 10.8.23**. The spectrum of the harmonic sound ends at 3,7 kHz in view of the usage of 45 partials; the spectrum of the inharmonic sound goes up to 4.1 kHz due to the frequency spreading (again 45 partial were used).

Distorting the **harmonic** sound results in additional partials that however are positioned exactly within the harmonic grid. On one hand, the new partials fill up the frequency range above 3.7 kHz, and on the other hand they change the level of the primary partials. The degressive curvature of the distortion characteristic has the effect that the partials decay more slowly, for some there is even an initial growth. The changes of the partial-levels are slow, with change speeds similar to those of the primary levels. For the **inharmonic** sound, the distortion generates many new partials positioned closely to the primaries such that the DFT-analysis (and the auditory system) cannot recognize them individually anymore. The spectral pooling of these undistinguishable lines results in fast signal modulations bearing some resemblance to a stochastic noise process but being (strictly speaking) determined (pseudo-noise).

The distortion has three effects on the sound: the treble-content grows (a more brilliant sound), the dynamics are compressed (longer sustain), and the partials are pseudo-stochastically modulated (creating a “buzzing” and “raspy” character). The pseudo-stochastic modulation happens only for inharmonic sounds and is dependent on the **string-parameters**. **The thicker the string, the more noise is created**. Maybe we should explicitly mention that this holds for single tones, because for chords, the spectrum is not harmonic in a simple fashion anymore, anyway.

The **level evolutions** shown in the picture below indicate that the pseudo-stochastic modulations increase if the characteristic is more strongly curved. Analyses for characteristic 3 are not included; a similar picture as for characteristic 2 would emerge, as long there is strong overdrive. Larger differences become apparent for less overdrive: the change from “undistorted” to “distorted” happens abruptly for characteristic 3, and more gradually for the other characteristics. Moreover, the spectrum becomes more treble-heavy if abrupt signal-limitation (clipping) occurs.

It should be expressly mentioned here again that despite a zero-symmetric characteristic (“odd-order function”), distortion products of even-order do occur. The assumption, that an odd-order-characteristic would generate only odd-order distortion products, only holds for half-wave anti-symmetric signals (e.g. for a sine tone). For real guitar sounds, even-order distortion products can very well result from odd-order characteristics. In the same manner, the assumption that tubes would generate predominantly even-order distortion products is wrong as a general statement. Tubes do not generate “better” distortion than transistors – otherwise nobody would ever have used a Range-Master ahead of the amp. In the Range-Master, pure transistor distortion is generated (Chapter 10.8.5.3).

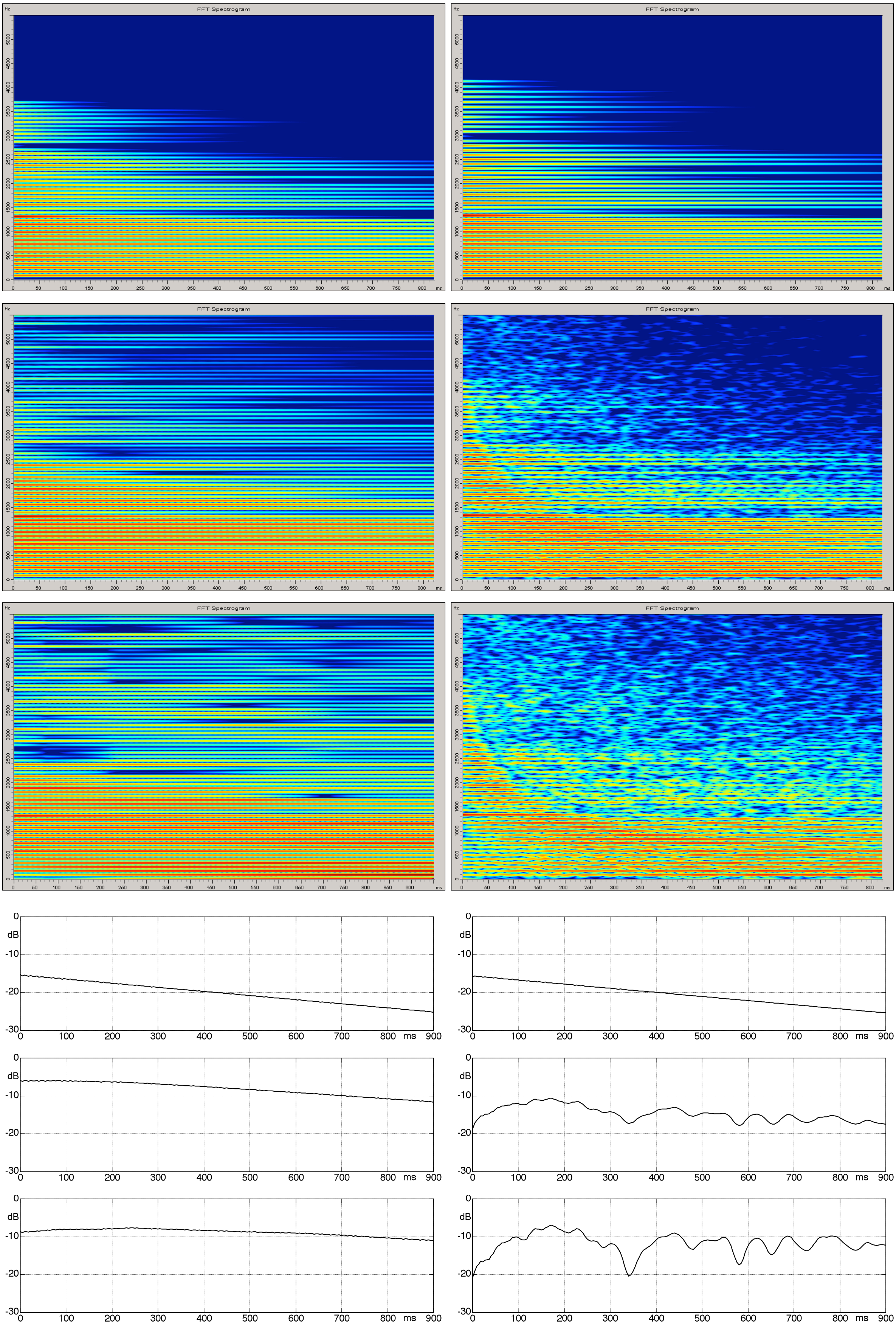

Fig. 10.8.23: Spectrograms, harmonic (left) und inharmonic signal, 0 - 5.5 kHz, ΔL = 40dB.
1st row: no distortion, 2nd row: characteristic 1, 3rd row: characteristic 2. Bottom: level-evolution of 15th partial.

Digital emulations of analog distortion devices merit some special consideration. It is generally known that an anti-aliasing filter needs to be included ahead of an A/D-converter in any time-discrete system. This is in order to avoid spectral back-convolutions (sampling theorem). If the signal is already appropriately bandwidth-limited, this filter may of course be dispensed of. For example, there is no need to include a filter if a 5-kHz-tone is sampled at 44.1 kHz. If the A/D-converted 5-kHz-tone is, however, distorted in the digital realm with a digital distortion characteristic, new frequency lines are generated – above half of the sampling frequency, as well. Since time-discretization has the effect of spectral periodization, these new lines appear around all multiples of the sampling frequency. If the distortion characteristic is point-symmetrical, new partials are generated at 15 kHz, 25 kHz, 35 kHz and further odd-numbered multiples. The distortion products are mirrored with respect to the sampling frequency (e.g. 35 kHz re. 44.1 kHz), a new distortion line appears at 9.1 kHz that would not be generated by an analog distortion device. At 0.9 kHz, and at many other frequencies, further partials appear, sounding rather unpleasant (as a rule, if their level is high enough). It is therefore insufficient to digitally emulate an analog distortion characteristic in order to create a digital equivalent. The higher the frequency of the signal to be distorted, and the more angularly shaped the distortion, the more disappointing the emulation will be.

To avoid such back-convolutions, the sampling frequency needs to be increased. Whether a ten-fold increase is adequate or whether even much higher sampling rates are necessary, depends on the signal, the distortion characteristic, and the quality requirements. Here is a simple estimate: if a sine signal undergoes hard clipping with a symmetric rectangular characteristic, new partials are generated following an si-envelope. The level of the 11th partial is 21 dB below the level of the primary, the 99th partial is 40 dB below the level of the primary. If the sampling frequency is 100 times of the frequency of the tone to be distorted, the back-convolution creates an interfering tone which is 40 dB down relative to the primary-level (**Fig. 10.8.24**). Of course, not only this one interfering tone is back-convoluted, and back convolution does not happen only at the sampling frequency. The figure shows merely one back-convolution such the lines can still be associated properly.

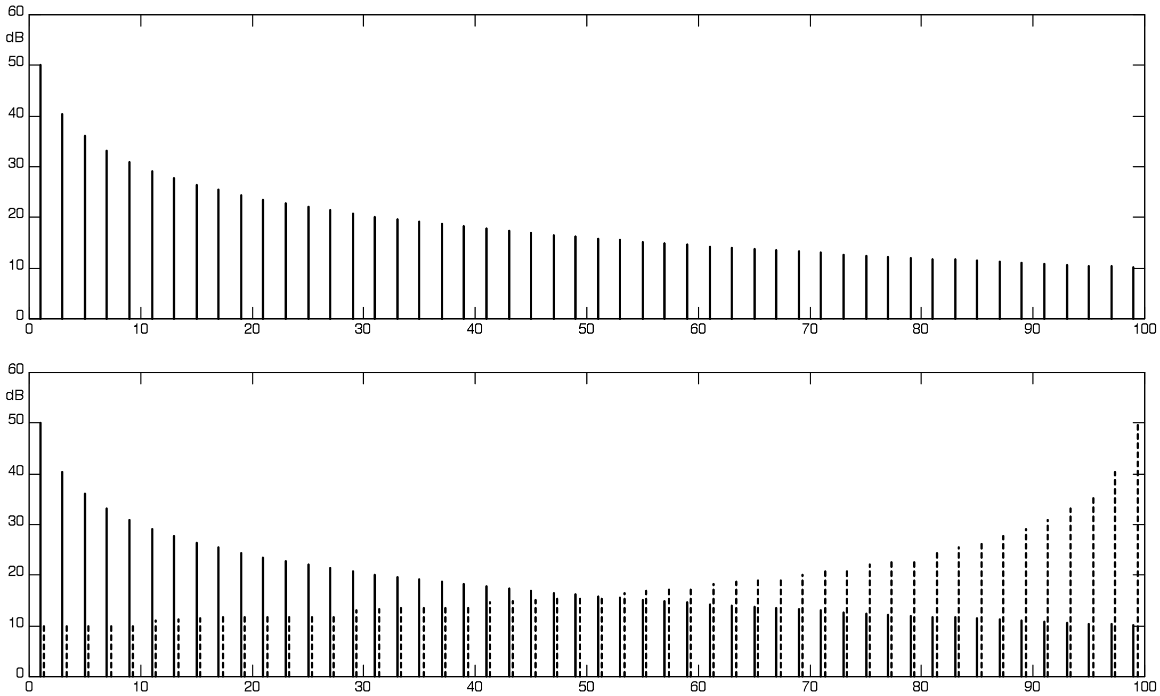

Fig. 10.8.24: Spectrum of a strongly distorted sine tone; time-continuous (top); time-discrete with one back-convolution (bottom). The frequency of the back-convoluted lines is strongly dependent on the relation f / f_a .

