11.6 Non-linear Distortion

In communication engineering we very carefully distinguish between linear and non-linear signal distortions: a linear system generates linear distortion exclusively, and a non-linear system (as far as it is free of memory) generates only non-linear distortion. Generally, one seeks to avoid mixing up linear and non-linear effects by defining sub-systems that individually generate purely linear or purely non-linear distortion.

In a **linear system** (e.g. an amplifier) the principles of proportionality, superposition and absence of sources hold. The latter characteristic is easily explained: where there's no input signal, there's no output signal. For a (non-zero) output signal $\tilde{y}$, a matching (non-zero) input signal $\tilde{x}$ must exist. If $\tilde{x}$ is doubled, $\tilde{y}$ must double, as well – that's proportionality. We quickly realize that the "linear function", as used in mathematics and defined by the linear equation $y = k{\cdot}x + m$, will fulfill the requirements of proportionality and freedom of sources only if m is zero. The law of superposition requires that the mapping of a sum must equal the sum of the mapped summands. Thus: $y = T\{x\}$ represents the mapping of the input signal x onto the output signal y. If the sum of two signals is fed to the system input, the following must hold in a linear system:

$$y = T\{x_1 + x_2\} = T\{x_1\} + T\{x_2\}$$ Superposition in the linear system

Proportionality and absence of sources alone do not suffice to specify linear behavior, as shown by the example of an **ideal full-wave rectifier** (that reverses the sign of a negative input signal): this device is source-free, and an n-fold input signal is matched with an n-fold output signal – but as soon as a further signal (e.g. a DC voltage) is added at the input, the waveform of the output changes … the rectifier is non-linear.

It is tempting to go and reduce the linear system to the matching-formula $y = k{\cdot}x$; however, this would unduly exclude the group of differential equations. A system that maps the speed of a mass onto its acceleration is a (time-related) **differentiator**♣. This system meets the requirements of absence of sources [$d/dt(0) = 0$], of proportionality [$d/dt(k{\cdot}x) = k{\cdot}dx/dt$], and of superposition: $d/dt(\xi + \mu) = d\xi/dt + d\mu/dt$. The differentiator is a linear system in spite of the fact that its sinus-transfer characteristic is not a straight line but an ellipse. Typically, the **linear distortion** generated by a linear system is specified for sinusoidal drive-signals as amplitude- and phase-distortion (or delay-time distortion), and is graphically represented as amplitude-frequency-response and phase-frequency-response. The bass-cut generated by an RC-highpass is a linear signal distortion, as is the presence boost of an equalizer (that of course must not be overdriven). Reacting to an impulse-like excitation, the resonant circuit of an equalizer will ring (theoretically for an indefinite time). Without a doubt, this is a signal distortion – but a linear one. Unfortunately, there is often a lack of distinction between linear and non-linear distortion, especially when it comes to loudspeaker characteristics discussed in popular "science". **Non-linear distortion** results if a system fails to fulfill one of the above mentioned linearity criteria – this system is then non-linear. Whenever possible, we try to separate linear and non-linear distortion into subsystem (possibly only existing as a model): a linear subsystem described by its "straight" characteristic, and a (memory-free) non-linear subsystem defined by its "curved" transmission characteristic.

♣ The formula-representation is meant to save space: it may not meet the expectations of all mathematicians.

It is, however, in many cases not possible to divide a real system into *one* linear and *one* non-linear subsystem: since in non-linear systems the (commutative) exchangeability is not there anymore, it may be that a plurality of subsystems is required, and the corresponding description may become highly complicated. The dynamic **loudspeaker**, as well, includes several non-linearities that do not allow themselves to be modeled in one and the same subsystem: the displacement-dependent force-factor (aka. transducer coefficient *Bl*), the displacement-dependent stiffness of the membrane-suspension, and the inductance that is also displacement-dependent. If the loudspeaker is mounted in an airtight enclosure, the non-liner stiffness of the air-suspension caused by the enclosure weighs in, as well. With the speaker mounted in a ported enclosure, the airflow within the port-tunnel introduces non-linearity. All these non-linearities generate a non-linear transmission characteristic but also cause a reaction on the electrical side and make for a strongly non-linear loudspeaker-impedance. In addition, we have non-linearity generated in the amplifier and the output transformer (if one is present). All in all we get a complex system with coupled non-linearities – and one that produces pronounced linear distortion to top it all.

For a loudspeaker mounted in a cabinet that is open to the rear, we may neglect the non-linearity of the air. At low frequencies, we may – to start with – dispense with a consideration of the inductance so that as a first approximation, a non-linear mechanical subsystem and a non-linear magnetic subsystem remain. The mechanical non-linearity is found in the stiffness of the membrane-suspension, i.e. the inner centering (spider) and the outer fastening (surround). As the membrane is deflected slowly, force is directed against a progressive spring with its stiffness increasing as the displacement increases. The stiffness is a system-variable while the displacement is a signal-variable. If a system-variable is dependent on the signal, we always have a non-linear system. The non-linearity in the magnetic system clearly is the transducer coefficient (the force-factor): as system-variable *Bl*, it takes care of the proportionality between current and Lorentz force: $F = Bl \cdot I$. However, this proportionality requires that the system variable *Bl* is independent of the signal – specifically independent of the displacement. That is not the case here: with increasing displacement, the coil moves out of the magnetic field and therefore *Bl* decreases. A further effect may play a role in this scenario: two magnetic fields superimpose as current flows. One is constituted by the permanent field generated by the permanent magnet, the other is the AC field surrounding the voice-coil wire. Because the ferro-magnetic parts located in the magnetic circuit all show a non-linear characteristic (the specific magnetic conductance μ is field-dependent), "modulations of the magnetic field" may result. Some manufacturers seek to decrease the effect via short-circuit rings while others do not do anything about it, regarding it as typical. It is here where the peculiarities of the **guitar loudspeaker** begin: while for HiFi-speakers there is consensus that non-linearity must be as small as possible, opinions diverge considerably when it comes to guitar speakers. You may hear (or read) on the one hand that the guitar speaker is, after all, a loudspeaker too (correct), and thus what has been taught in the HiFi-domain for decades cannot be wrong (??). On the other hand, (positive) reviews including evaluations such as "dirty midrange" give rise to some hope that at least a few designers have recognized that sound-shaping function of the guitar loudspeaker.

We must not fail to mention here though, that not only among the manufacturers, but also among the players multiple opinions abound. You get the Jazz-dude who brutally chokes the hard-won brilliance of the guitar by bottoming out the tone control, the Country-picker with his piercing treble, the crunching-along Blues-man, the chainsaw-ing Metal-ist, the Jack-of-all-Trades cover-guy, and the folksy oom-pah-strummer. A consolidated drive towards standardized loudspeaker distortion may not be expected given such a heterogeneous population and mix of opinions.

Here's an example taken from loudspeaker history: **JBL**, the renowned American speaker manufacturer, looks back on a long tradition as supplier for cinemas, recording studios, and living rooms. Nothing but High-Fidelity – non-linear distortion is marginal as a matter of course: ... *"low distortion which has been always associated with JBL products"*. In the early 60's the demand for instrument speakers grows, and at JBL, a tried and trusted workhorse, the D-130, is modified to become the **D-130F**. The changes mostly relate to the air gap that was – according to statements by the designer (Harvey Gerst) – slightly enlarged in order to obviate damage. And then there's the designation: F is for Fender, the largest customer. Years later, the **K-130** follows with double the power capacity compared to its predecessor but still *"clean at any volume level"* (a quality that probably would not have always applied to the associated musicians). Both the D-130F and the K-130 were fitted with Alnico magnets, but the next generation – the **E-Series** – received ceramic magnets. This prompted JBL-mastermind John Eargle to state that Alnico was known for its *"inherently low distortion performance"*. However, according to him, the new E-Series is even better: *"The improvement has been in reducing second harmonic distortion",* obtained with the "*symmetrical field geometry"*. Given this upgrade, the loudspeaker is eminently suitable *"for vocals – and guitar"*. Presumably, this further added to the already long list of JBL-users shown in the adverts. Due to space-restrictions, this list cannot be commemorated in its full extent here, but the following may serve as an excerpt: Count Basie, Harry Belafonte, Tony Curtis, Sammy Davis jr., Doris Day, the unforgotten Carmen Dragon, Duke Ellington, Ella Fitzgerald, Hugh Hefner (!), Dean Martin, Frank Sinatra, not to forget Richard Nixon and "The Duke" John Wayne♣. Global super-stars, all of them – and all of them JBL-users. Such a feat of course calls the competition into the arena. And thus it was that **Electro-Voice** retaliates with a big swing, proclaiming: *"Symmetrical magnet gap structures have been promoted as desirable in a guitar speaker. We have found this to be a fallacy"*. Because: *"A coil moving in an asymmetrical magnetic gap will generate a mixture of odd and even harmonics, resulting in a more complex, richer sound."* To each his own … there's no accounting for taste.

So, let's not begrudge The Duke the undistorted JBL-sound of his electric guitar (hm … still thinking about that one …), and Joe Bonamassa his EV-sound chirping from the 4x12's – beauty is in the eye of the beholder. What can be said about **magnetic non-linearity** from a scientific angle? If you leave the pole-core (the cylinder in the interior of the voice coil) formed as a cylinder over its full length, as shown in **Fig. 11.63**, then an asymmetric scatter-field will result: the shape of the field above the coil is different from the shape below it. Reducing the core-diameter in the lower section, though – as it is shown in exaggerated fashion in the figure – will render the two stray-fields more symmetric. The result is that a symmetric Lorentz-force acts on the voice coil for both positive and negative displacement. As already mentioned, this force depends on current and displacement. While the current-dependency is desired, the displacement-dependency is not, because it generates non-linear distortion. For a symmetric field, the distortion is of even-order (even function) – given asymmetry distortion, odd-order also weighs in.

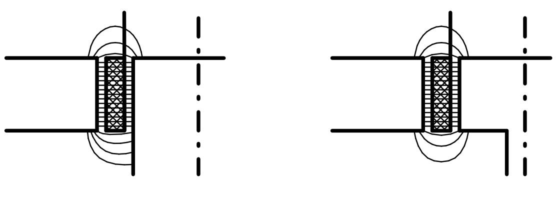

Fig. 11.63: Different designs of the pole-core. On the left it is purely cylindrical, on the right offset. The scatter-field generated outside of the air gap depends on the geometry of the pole-core.

♣ From JBL's 1968 loudspeaker brochure.

The field-limiting at both ends leads to a degressive, clipped current/force-characteristic that leads some people to conclude that the course of the oscillation would now also be limited – similar to an overdriven amplifier. This assumption, however, overlooks that the force is mapped onto the displacement (via Hooke's law) only in the range below the resonance. At resonance the membrane acts (in the simple model) as a damper, and above the resonance it acts as a mass [3]. If, in the range above the resonance, the motive force becomes weaker at strong displacement, this reduces the *acceleration* primarily. Of course there will be an effect on the displacement, as well, but: displacement and acceleration are in opposite phase. Or, to be somewhat more precise: the acceleration is the second derivative of the displacement. Analytically, this leads to a non-linear differential equation that can be solved approximately – with a big effort, however (the system is not just weakly but extremely non-linear).

The qualitative effects of the inhomogeneity may be studied rather well using the **simple membrane model** [3]. If we reduce the membrane to spring, mass and damper, and the electric side to the ohmic voice-coil resistance, we have a 2nd-order system that may be described via the **frequency of the pole** (resonance frequency) and the **Q-factor of the pole**. The pole-frequency unambiguously results from the stiffness and the mass, while two limiting cases are of importance for the Q-factor of the pole: open circuit and shot circuit on the electrical side. Given the open circuit, no current can flow and consequently the magnetic field will exert no force onto the membrane. The Q-factor of the pole depends solely on the mechanical parameters: $Q = \sqrt{sm}/W$. The purely mechanic dampening of the membrane is relatively small such that the Q-factor of the pole is considerably larger than 1 (5 is not uncommon). As the voice coil is shorted (or as an amplifier with a very small internal impedance is connected), the voice-coil resistance that is transformed from the electrical to the mechanical side acts as an additional dampening[♣], and the Q-factor of the pole drops below 1. Since the electromechanical coupling becomes smaller at large displacement (due to the inhomogenities in the field), the membrane-dampening decreases with increasing drive levels – the displacement tends to become too large, and not too small as it would with degressive limiting [3, Chapter 6.2.3].

In the asymmetrical magnetic field, the reset-forces acting at the extremes of the membrane-displacement are unequal (in terms of magnitude), and therefore the average force is not zero. A steady force of the frequency 0 Hz results that pushes the membrane out of its neutral position in the direction of the weaker of the two fringe-fields. Because in reality the two fringe-fields are never exactly identical, this effect always occurs: a small asymmetry suffices to make the membrane wander slightly from the idle position. This enhances the lack of symmetry, and the membrane continues to wander off – it is only stabilized by the onset of the counter-force exerted by the membrane-suspension. Therefore, 2nd-order distortion is to be expected in the range above the resonance – even if there is a symmetrical layout of the magnetic field. "Field-modulation" is a further source of 2nd-order distortion: part of the magnetic field generated by the flowing current superimposes onto the steady field of the permanent magnet, i.e. the flux density therefore fluctuates in sync to the excitation current. Because of this, the force obtains a share that is dependent on the square of the current – and that implies 2nd-order distortion. Another way of explaining the effect: the flowing current generates attracting forces between neighboring ferromagnetic parts (through which the field flows). These attracting forces are <u>in</u>dependent of the sign and therefore generate even-order distortion (just like a rectifier). Relief, if at all sought, could come in the form of a short-circuit ring. It forces the AC field out of the magnetic circuit, and the 2nd-order distortion decreases.

♣ The eddy-current brake in lorries and trains works based on a similar principle.

Analyzing the current fed from a stiff voltage source will give a first impression regarding the linearity (or non-linearity) of the loudspeaker. The mechanical membrane-impedance *F/v* is transformed to the electrical side with the square of the transducer coefficient. If we disregard the inductance, the electrical impedance consists of two components: the voice-coil resistance (e.g. 6 Ω), and the transformed mechanical impedance. Any non-linearity in the electrical impedance (given a stiff voltage source these would be **current distortions**) will consist of two components: a non-linear transducer coefficient (*Bl*), and/or a non-linear membrane impedance. In a loudspeaker, both these components are present: both the stiffness of the membrane-suspension and the transducer coefficient are displacement-dependent. The current-curves for the operation close to resonance are shown in **Fig. 11.64**, with the distortion being very significant. It should be noted that the voltage amounts to merely 10 V_{eff}, i.e. nominally only 12.5 W for this 8-Ω-speaker (deployed in a 60-W-amplifier). Moreover, since the impedance will be at its maximum at resonance, the power taken by the speaker will be even (much) less – we are far away from any undue overload situations. The curves reveal a strong 2nd-order distortion. The amplitude of the 2nd harmonic rises to up to 67% of the 1st harmonic; this would correspond to a harmonic distortion of k_2 = 56% (the approximation U_2 / U_1 should not be used anymore at such high distortion levels). It is beyond the aim of this chapter to localize or separate the individual roots of these distortions – the effort would grow too big. Rather, we will present comparative distortion measurements; these will consistently show that all investigated loudspeakers are strongly non-linear systems even at very moderate drive levels. Not that guitar players would generally be adverse to such characteristics …

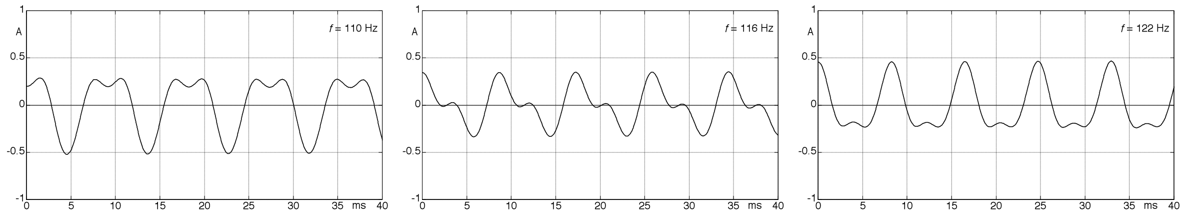

Fig. 11.64: Time-functions of the loudspeaker-current fed from a stiff voltage source, U = 10V; VOX AD60-VT.

The frequency responses of the distortion (**Fig. 11.65**) show that the maximum distortion of the current happens at the main resonance (116 Hz); it is here that the displacement is at its maximum. There are two reasons that 2nd-order distortion can rise to such heights: at resonance, the 2nd-order harmonic of the current is highest (due to the mentioned non-linearity), and at the same time the overall current becomes smallest (due to the rise in the impedance. The difference of the two levels (the distortion dampening) therefore has a pronounced maximum here. However, the current-distortion describes predominantly the electrical behavior – non-linearity in the sound radiation needs to be analyzed separately.

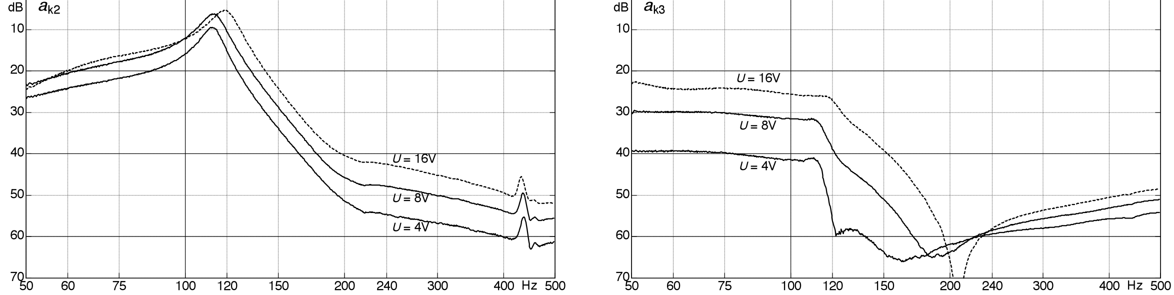

Fig. 11.65: Frequency-dependence of harmonic distortion in the current when fed from a stiff voltage source (as in Fig.. 11.64). The drive-level-dependent shift of the maximum is a result of the strong non-linearity.

In order to measure the (non-linear) distortion in the sound pressure, the speaker was operated from a stiff voltage source and mounted in the VOX AD-60VT enclosure. Measurements were taken in the AEC, with the microphone on axis at a distance of 3 m (**Fig. 11.66**). The results are not untypical for a dynamic woofer: at low frequencies very strong distortion is generated. Then, from 70 Hz up, the 3rd-order distortion drops off faster than the 2nd-order distortion, and above 150 Hz, the THD remains below about 1%. Compared to the analysis of the current (Fig. 11.65), the distortion has mostly increased.

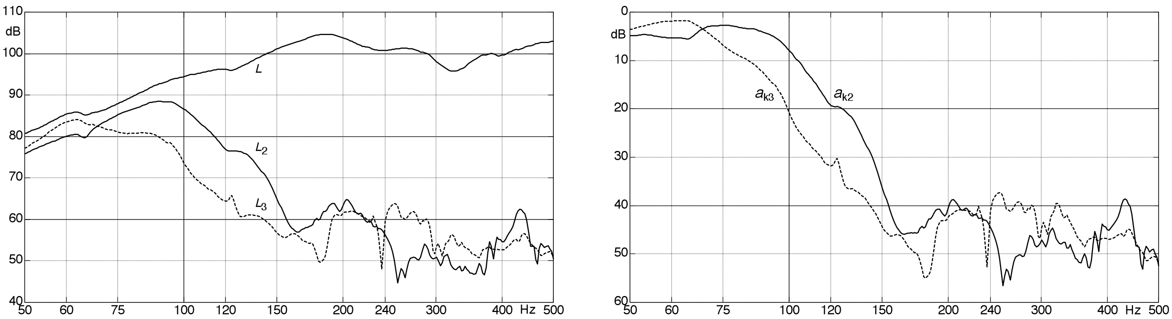

Fig. 11.66: Frequency responses of SPL (left) und distortion level. VOX AD60-VT, U = 10V.

However, a THD of 1% is not the professed aim for a guitar loudspeaker – some speakers easily reach ten times that distortion: **Fig. 11.67** illustrates measurements with another Celestion loudspeaker, operated with the same voltage and in the AD-60VT-enclosure: the Celestion “Blue”. This speaker is not broken – far from it; it’s just that the input voltage of 10 V already pushes it close to the borders of its power capacity (15 W). On the other hand, this should not be taken as evidence that a THD of 10% would be typical for getting near to the power limits: the Vintage-30 speaker (depicted below the “Blue”) is specified at 60 W and, at 10 V, distorts similarly to the “Blue”. Since the concept behind the Vintage-30 is that it should be a descendant of the “Blue”, it is only sequacious that it should distort like the latter.

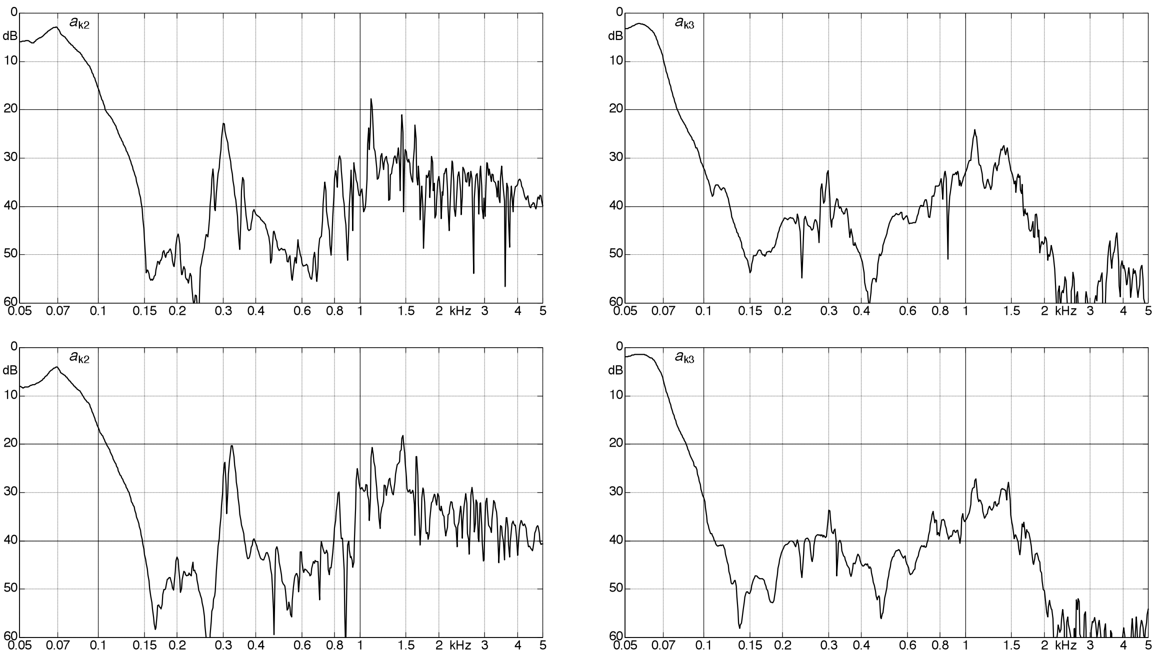

Fig. 11.67: 2nd- and 3rd-order distortion. Enclosure: VOX AD-60VT, U = 10V, d = 3m.
Upper row: Celestion "Blue" (P_{max} = 15W); lower row: Celestion Vintage-30 (P_{max} = 60W).

Documentations on loudspeaker non-linearity are often limited to measurements of the harmonic distortion; this may have to do with the fact that such measurements are a standard tool in systems analysis. Since Brüel&Kjaer has released its legendary instrumentation-combination of the 2010/1902-devices, difference-frequency measurements are also common for band-limited systems – but there is a further distortion-mechanism that is found especially in loudspeakers: **sub-harmonics**. This term means to describe the generation of distortion tones that have a frequency lower than the excitation frequency, e.g. $f/2$ or $f/4$. **Fig. 11.68** depicts, accordingly, two spectra derived from the sound pressure. A sinusoidal voltage (10 V) was imprinted at the loudspeaker connectors, with f = 1.6 and 1.5 kHz. Highlighted in grey are those spectral lines (broadened by leakage) that could be expected as regular "harmonic distortion"; in addition, however, we see a sub-harmonic developing, and frequency-multiples of it. The double-peaks in the right-hand diagram point to fast time-variant processes: the spectrum resulted from a sweep, and the "sub-harmonic distortions" change their level very fast.

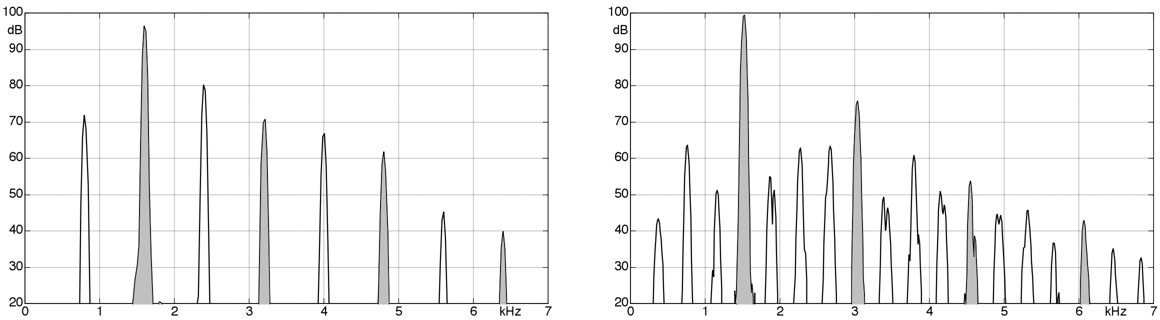

Fig. 11.68: Sub-harmonics at half (left), and a quarter (right) of the excitation frequency.

Such sub-harmonics appear – if at all – only in small frequency ranges. **Fig. 11.69** shows two spectrograms representing the level as grey value across the f/t-plane. The bottom rising curve belongs to the level of the first harmonic, the levels of the higher harmonics follow above. The grey dots or groups of dots appearing in the right half of the diagram point to sub-harmonics (or their frequency-multiples). The speaker analyzed on the left (Jensen P12-N) shows sub-harmonic distortion only at an excitation frequency of about 1760 Hz, while the speaker on the right (Celestion G12-Century) features it in several ranges from 920 Hz up. Both speakers have an impedance of 8 Ω and both were measured at 10 V. The C12-N is specified with a power capacity of 50 W, and the G12 at 80 W – neither speaker is therefore operated close to any power limit.

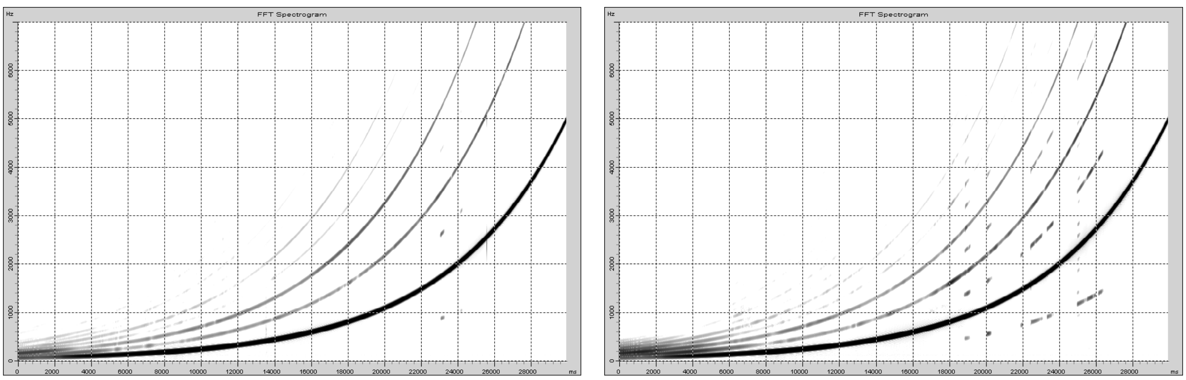

Fig. 11.69: Sweep-spectrograms f = 50 – 5000 Hz, U = 10V. The speaker analyzed on the left generates sub-harmonics only at about 1760 Hz while the speaker on the right does so in several frequency ranges. Abscissa-scaling: sweep-time = 0 – 30 s; ordinate-scaling: frequency = 0 – 7 kHz.

The generation of sub-harmonic distortion cannot be explained merely via a curved transmission characteristic. As **Fig. 11.70** shows, a tilting vibration of half the frequency is superimposed. Mathematics laconically (and correctly) explains such phenomena with “solution of a non-linear/time-variant differential equation”, physics offer “parametrically excited eigen-oscillations of a system with time-variant system-variables”. Time-variant quantities are indeed easily imaginable: the membrane deforms, and the location-dependent stiffness of the membrane is certain to be dependent on load – and therefore on time. Moreover, the oscillation of the membrane by no means needs to be one-dimensional: tilting- and tumbling-movements are possible, and the overall system is of a complicated, non-linear nature. We may expect sections of the membrane oscillating with the same phase – but of course not the whole area; there will be phase shifts, and since the system parameters are time-variant, most probably there will be time-variant phase shifts, as well. Simple models fail here, e.g. since already the superposition principle may not be applied anymore.

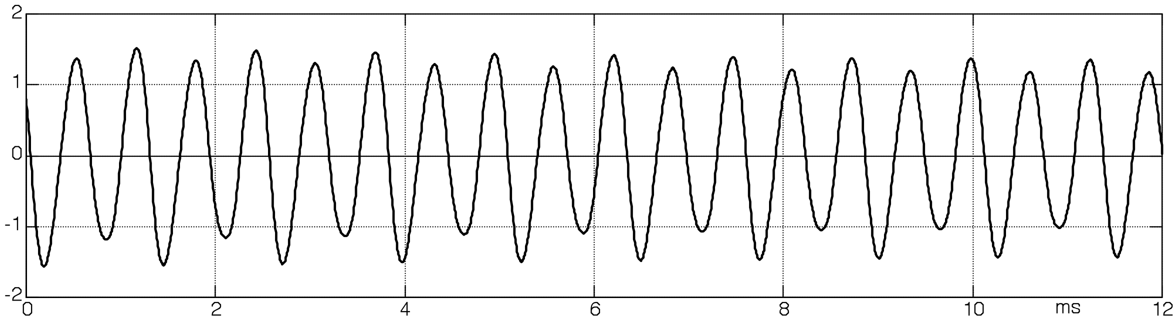

Fig. 11.70: Time function of sound pressure in a sub-harmonically distorted sinusoidal tone: f = 1,6 kHz.

Fig. 11.71 is targeted at showing at showing another example of the complexity of sub-harmonic distortion: above 1.5 kHz, this loudspeaker generates sub-harmonics not only at half the excitation frequency, but – among others – also at $f/4$ and $f/5$ (and at the multiples).

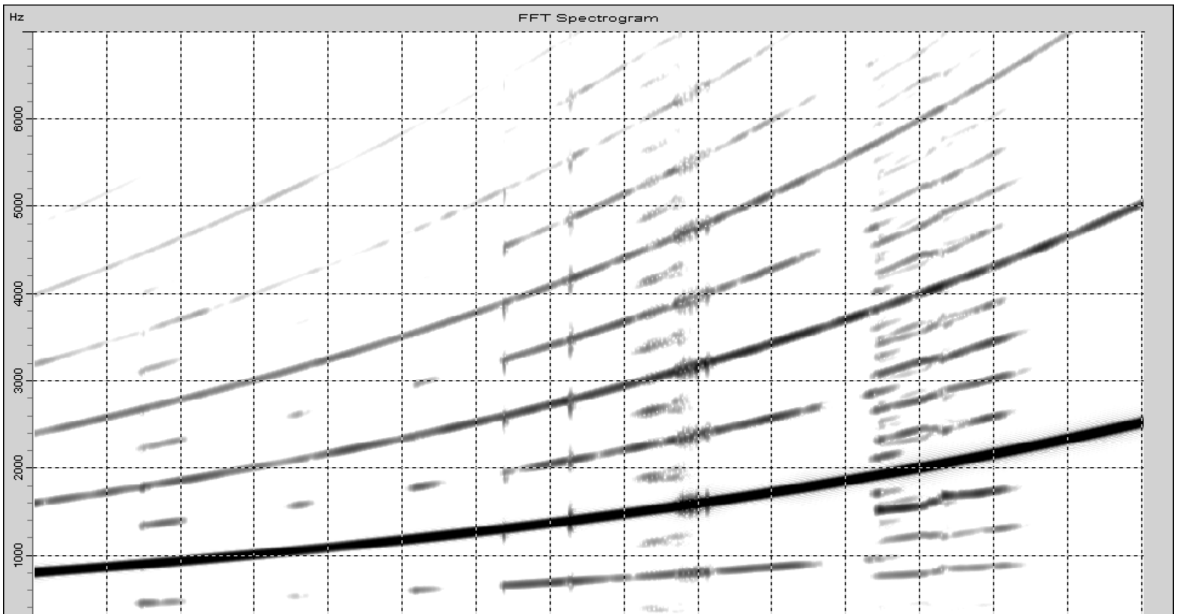

Fig. 11.71: Jensen P12-R, sweep-spectrogram, f = 800 – 2500 Hz, U = 10V.

The **levels of the sub-harmonics** follow their own laws, not even showing the power laws to be expected for standard models. IN **Fig. 11.72**, the level of sub-harmonic (*f/2*) is shown dependent on the level of the primary tone (*f*). For primary tones below a certain threshold (here at just under 108 dB), there is no sub-harmonic at all. Going across that threshold, the sub-harmonic builds up. As we reduce the level of the primary tone below the threshold value, the level of the sub-harmonic remains constant first – only as the primary tone falls below about 104 dB, the sub-harmonic disappears again.

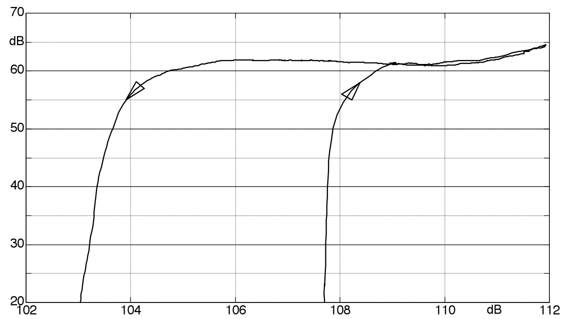

Fig. 11.72: Level-hysteresis (left), distortion spectrum (right). Celestion Vintage-30, *f* = 844 Hz.

Fig. 11.73 more precisely depicts the evolution of the level for three frequencies. The generator level rises by 25 dB during 30 s: the corresponding measured sound pressure levels are shown. At 1081 Hz, no sub-harmonic is created and the levels grow monotonously. At about 1.3 kHz, however, a sub-harmonic appears around -11 dB (50W / 12.5 = 4 W), and this has effects on all measured sub-harmonics. From -9 dB, there are audible beats.

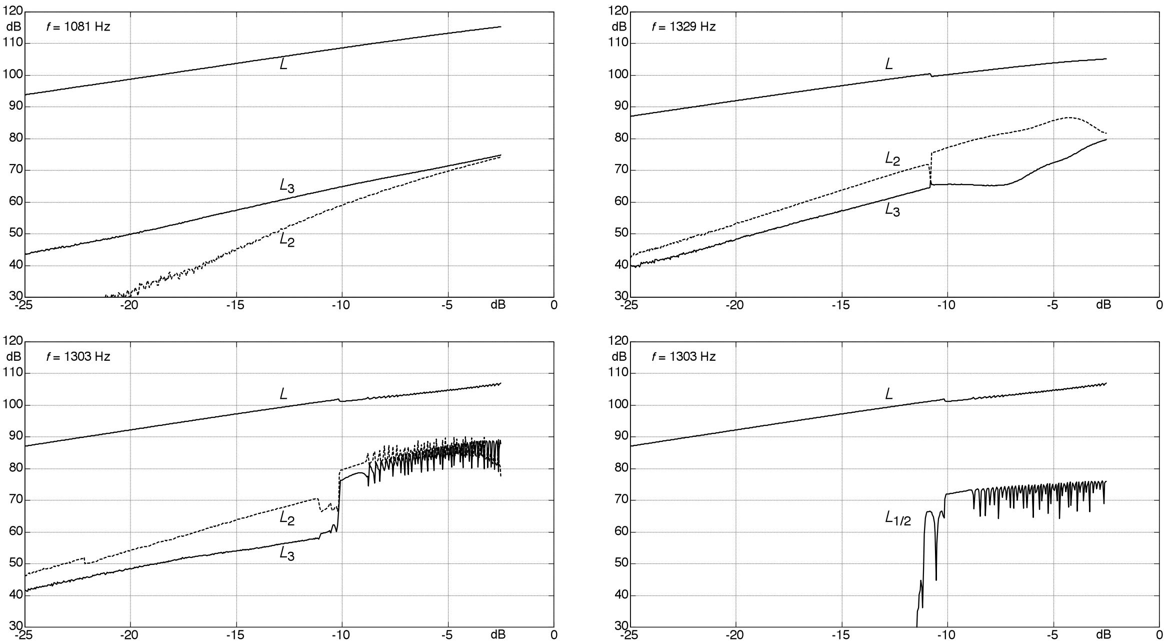

Fig. 11.73: Sum SPL *L* and distortion level, Eminence L-105; 0dB = maximum power.
Lower left: for f = 1303 Hz, the level development of the sub-harmonic (*f*/2) is shown.

The following **Fig. 11.74** depicts the non-linear behavior of several loudspeakers in an overview; supplementary measurement data are added in the last of the three diagrams.

Fig. 11.74a: Distortion suppression (harmonic distortion attenuation) a_{k2}, a_{k3} of various loudspeakers. The distortion suppression of the sub-harmonic is, respectively, included at the upper right in the right-hand.

These figures are reserved for the printed version of this book.

Fig. 11.74b: Distortion suppression (harmonic distortion attenuation) a_{k2}, a_{k3} of various loudspeakers. The distortion suppression of the sub-harmonic is, respectively, included at the upper right in the right-hand.

These figures are reserved for the printed version of this book.

Fig. 11.74c: Distortion suppression (harmonic distortion attenuation) a_{k2}, a_{k3} of various loudspeakers. The distortion suppression of the sub-harmonic is, respectively, included at the upper right in the right-hand. Given the nominal impedance of 8 Ω, the voltage (10 V_{eff} from a stiff voltage source) results in a power of 12.5 W. All measurements were taken in the AEC, with the 12"-speakers mounted in the VOX AD-60VT enclosure, the 15"-speakers mounted in an air-tight enclosure measuring 36x74x40 cm^3, and the 10"-speakers in an air-tight enclosure measuring 39x39x25 cm^3.

These figures are reserved for the printed version of this book.