Loudspeakers for AEC: Measurement and Linearization

Adaptive filter can average over uncorrelated random background noise.

Adaptive filter can not average over correlated LTI distortions.

The performance of the FSAF[1] AEC is hard limited by loudspeaker's non-Linear-Time-Invariant-ness (non-LTI-ness). The assumption of "all loudspeakers are essentially the same" is been useful only for poor, essentially half-duplex AEC. For an AEC striving to perform in the maximally transparent full-duplex mode, this assumption is wrong and must be explicitly discarded. LTI distortions vary wildly between loudspeakers and can not be easily identified during a "regular" teleconferencing. Therefore, a well-performing AEC must be tied to a specific hardware. An adaptive filter can be more or less generic but a good AEC can not be generic.

1 INTRODUCTIONS

1.1 BASICS

Many words start with weak fricatives and plosives. AEC must block any echo of loudspeaker [non-]LTI distortions but pass the lowest possible double-talk to preserve intelligibility, trust and mutuality. However, it's practically impossible to distinguish echoes of arbitrary LTI distortions from the double-talk on the fly, in real time, as it happens. Therefore, AEC must be capable of estimating | foreseeing | predicting the variance and spectral distribution of loudspeaker echo LTI distortions for any reasonable signal transmitted, to insert as little signal alteration | attenuation | suppression as possible. The same estimate shall be used for step size control:

$\mu_t = {v_t}^2 / ({v_t}^2 + {\Sigma_t}^2)$; where

${v_t}^2 = x_t^H D_t x_t$; a priory estimation of residual error, before the noises.

$D_t = E\{(h_t - h)(h_t - h)^H\}$; dispersion matrix

${\Sigma_t}^2 = {\sigma_{n,t}}^2 + {\sigma_{u,t}}^2 \quad \ldots + {\sigma_{LTI,t}}^2$; total variance of all noises and distortions sensed on the output, including the expected [echo] LTI distortions' variance. More than often LTI distortions dominate over background noise.

Popular music recordings have spectra falling at ~-6dB/octave. Decade-wide woofer, mid woofer and tweeter must have roughly equal power handling capabilities. Classical music and voice spectra decay at the -12dB/octave after, say, 5kHz, which somewhat relaxes power handling requirements to tweeters on such programs only.

The most important type of loudspeakers for AEC is mid-woofers (a.k.a. midrange) working in the 300Hz … 3kHz range - which coincides with old PSTN telephony bandwidth.

I would like to pre-emptively apologize for 1) the lengthy explanation in the "We shall tell it at length, thoroughly, in detail – for when did a narrative seem too long or too short by reason of actual time of space it took up? We do not fear being called meticulous, inclining as we do to the view that only the exhaustive can be truly interesting" style instead of "For every complex problem there is an answer that is clear, simple, and wrong", and 2) statements that some less mathematically literate readers may find offensive. If any of this may affect you, I kindly ask you to give this paper a miss.

---

[1] Fast Subband Adaptive Filtering (FSAF) - File Exchange - MATLAB Central (mathworks.com) , part 4 is a mandatory prerequisite reading.

1.2   [OVER]SIMPLIFIED THEORY OF [MID-WOOFER] LOUDSPEAKERS

The loudness is sound pressure, often expressed in dBSPL, re 20µPa, which is deemed to be a median threshold of sensitivity to an 1kHz tone in a headset.

Sound pressure is a function of air particles acceleration, 2nd derivative of instant air density wave vibrations. For the same loudness, (air particles and) the cone excursion amplitude changes as *~1/f²*. If a subwoofer max excursion is, say, 10mm on 20Hz, the corresponding tweeter excursion is (ignoring *Sd* differences) only 10nm on 20kHz.

1.2.1   Main factors

There are only a few major factors affecting the loudspeaker performance most:

- Structure, size and shape of the cone, often referred as a standard class (5.25", 6.5", etc) or *Sd* - effective radiating surface area, with *d=2*(Sd/π)^{1/2}* and cone/dome depth *Cd*, mm
- *Mms* - mass of all moving parts in the cone - voice coil assembly, and the "equivalent" mass of suspension.
- *Fs* – free air resonant frequency of the cone, defined by *Mms* and the suspension compliance.
- *Le* – voice coil inductance at 1kHz
- *Re* – voice coil DC resistance
- *Bl* – a multiplication of *B* – magnetic flux density in the gap, and *l* – length of voice coil. The force applied to the cone assembly by the current *I: F=(B*l)*I*. As *B* and *l* are always used together, the symbol *"Bl" (="B*l")* is often perceived as one.

1.2.2   …and their influence on FR, diffraction and sensitivity

- *f < Fs*, a high-pass second order filter on *Fs* applies, +12dB/octave from *Fs* down.
- *f > Fs*, loudness ~ *Sd*Bl/(Mms*(Re+j*ω*Le(ω)))*, flat, piston alike, with sensitivity represented as dBSPL on 1W at 1m or dBSPL on 2.8V at 1m into *2π* half-space, anechoic.
- *f > F_{dir} = c/(π*d),* where *c* = speed of sound (~340m/sec, depending on air pressure, humidity, etc), a +6dB/octave: on-[loudspeaker-]axis and -12dB/octave: 90^0 off-axis. The cone radiation pattern changes from omnidirectional to more and more unidirectional, in the same manner as for uniform beamforming. If you do not keep the same constant directivity for *$f_{crossover}$ < f < 20kHz*, BBC dip will "happen".
- *f > F_{cone}* which is defined by a single slope triangle filter in a time domain from 0 to the cone depth. For *Cd*=70mm*: L=cd/c*; Sampled on *fs=48kHz: cd*fs/c* = (70mm/7mm) = 10. For off-axis angle *θ* the filter elongates as *cos(θ)^{-1}.* Dust cap shortens the filter in an obvious way. Most often, *Cd* is too low.
- *f > F_{Le} = Le(f)/(2π*Re)* – a partial order low pass filter applies, usually -3…-3.5 dB/octave, due to the "shorting ring" which reduces inductance, *usually Le ~ 1/sqrt(f).* See [2], [3] for further details.

---

[2] Transient Loudspeaker Model with Shorting Ring:Finite Element Method Magnetics (femm.info)
[3] Microsoft Word - FaradayRingsVoiceCoilImpedance.doc (geocities.ws)

Loudspeaker FR for diameter=100mm, depth=35mm Le=0.78mH
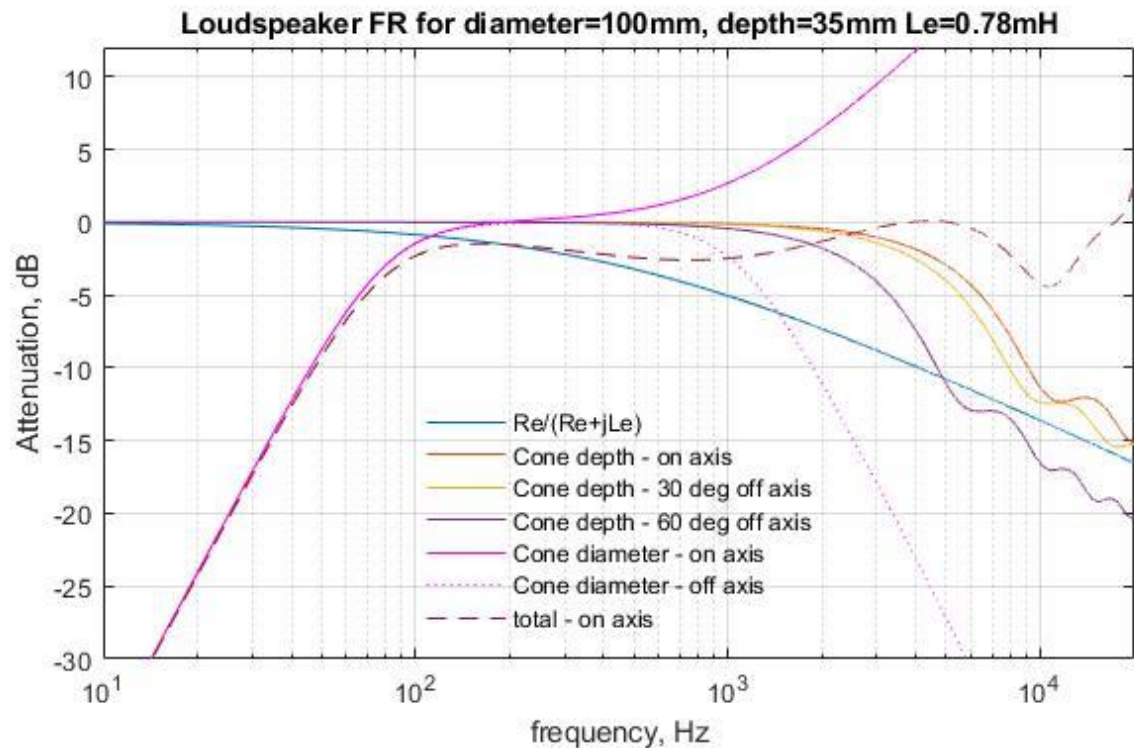
The rise for f>10kHz most often absent due to the effects which were not mentioned.

- $f >= F_{breakup}$ where the cone loses piston mode and starts to resonate and the model above loses applicability. Air resistance drops down because different areas of the cone move in opposite directions. There will be spikes at FR which are not repeatable. Much of the development effort was spend on attempts to find optimal cone shape and materials, from pure beryllium or aluminum to rocket science composite sandwiches of carbon fibers with viscous filling, to find the lowest density material with maximum strength cone | ark | dome shaped. This challenge was by no means unique, and the traditional theoretical mechanics of solid states and envelopes could have helped to avoid many pitfalls.
- The reported FRs, even from the vendors, frequently contain somewhat periodic ripples of various frequencies. More often than not, they come from an excessively reverberant test chamber. The frequency period of FR ripples Δf corresponds to a dimension of the test chamber: $\Delta x=n*c/\Delta f$.
- Be aware of room gain and corners. A good introduction: What is Subwoofer Room Gain? | SVS Sound Experts Blog (svsound.com) See dedicated literature for more details.

1.2.3   Linear behavior

Ideally linear acoustics, a well-developed scientific discipline, describes loudspeaker's linear behavior quite well. A competent Ph.D. or Sr. Engineer can [use standard constrained optimization methods to] find a desirable compromise between parameters despite everything being interconnected with everything else in loudspeakers. Room acoustics is easier to measure than to calculate with practically required precision. Off-axis responses into anechoic space can be predicted theoretically sufficiently well, so we study only the measured on-axis responses.

1.2.4   Harmonic and intermodular distortions

Let's look at IMD of Peerless TG9FSD1004, with voice on 2500Hz and base swept from 80Hz to 200Hz:

signal spectrogram [/recs/spk/220805/peerless/imd80/70:akg.p170]
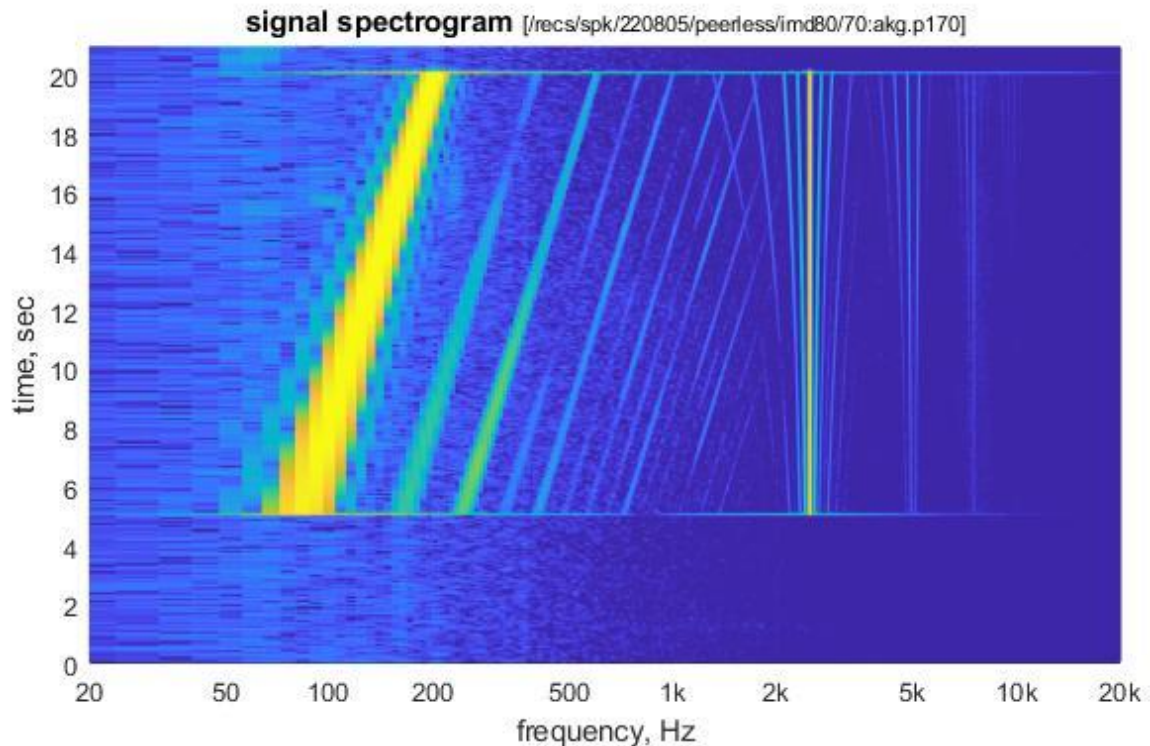
Let's observe that while base's harmonics fall with frequency increase, voice harmonics and intermodulates grow. The same observation is true for other drivers I tried, which contradicts mainstream "audioscience" theory.

1.2.5   Non-Linear behavior

An established theory by Dr. W. Klippel[4] describes non-linear distortions in the loudspeaker drivers as dependant on cone - voice coil assembly displacement or velocity. Therefore, according to the Fourier transform theory, the displacement-related distortions must decline as $1/f^2$, and the velocity-related distortions must decline as $1/f$. N^{th} order harmonic distortions must grow (and decay) as the input [sine wave] amplitude a^N, or 6.02dB*(N-1) faster than input. Or so they say.

1.3   IMPORTANCE OF LOUDSPEAKER MEASUREMENT

There is a wild variety of pseudo-scientific theories on the loudspeaker quality and inaudibility of various distortions, usually coupled with complete disregard to rooms' acoustic treatment, all of which belong to the dustbin. There is a history of disagreements between "objectivists" and "subjectivists" as related to the loudspeaker quality. Subjectivists are correct: the chirp and MLS measurements say next to nothing about audio quality; they imply a theory which is out of place and time.

While music is a magic (as far as it gets in this world), loudspeakers are not magical. Loudspeakers are fully & objectively measurable within standard System Identification approach - which measurements do not imply a theory. A scientific theory must be based on objective repeatable measurements and statistics, and must not contradict any relevant measurements. Meaningless, out of place and time, measurements do not prove existence of magic. There is no such thing as "*The **Art** of Designing Loudspeakers*".

---

4 Transducer Nonlinearities (Curve Shape) (klippel.de)

2   METHOD

Measurements were done in a listening room with ½ of wall surface covered by acoustic foam, and the rest by diffracting 3/4-full open bookshelves. RT60 is below 200ms, FR ripples are 2-3 times less than in the loudspeaker vendor datasheets[5]. Recordings were made using a close pickup [super-]cardioid condenser microphones AKG P.170 and Behringer B-5 (modified by increasing circum-linear path between diaphragm sides) with dual calibration:

1. Apex.220's 6mm noisy electret high-frequency FR was corrected using known FR of Focal Spectral 918.1 (earlier measured by a certified B&K mic) and Genelec 8020.
2. AKG P.170/B-5 FR was corrected for the chosen distance to match Apex.220's FR. Both directivity and proximity effects allow reducing distant low-frequency noise and room influence. Figure-8, second-order differential mic, or endfire microphone array would be even better.
3. Setup location and direction maximally attenuate standing waves' influence. Floor & ceiling reflections delay $\sim$9', microphone to driver < 1'; this setup is low-frequency calibratable because loudspeakers are omni in LF, so an anechoically measured driver can serve as the reference.
4. RIR modified Gabor analysis allows to separate loudspeaker and the room.
5. Working knowledge of signal processing, human voice tract, system identification and dereverberation algorithms is required to prevent messing up. Not a technician job.

Loudspeakers under test were mounted onto a closed nautilus shaped box, full of acoustic foam, verified to handle [known] 4" & 5.25" loudspeakers well enough. The 6.5kg box was suspended from a pole with horizontal $f_r\sim$3Hz and $Q\sim$1.0. The maximum cone weight was <10g. The nonlinear impairments due to Newton's 2nd Law $(ma_1=-Ma_2; p_2=p_1*(m/M)*(S_{box}/Sd);)$ consequent radiation by box walls, and suspension's non-linearity were estimated to be below -80dB re loudspeaker on f>100Hz. You may prefer to suspend your test box from the ceilings (which has even lower f_r = good solution[6]) but be cautious of spikes even if they are "super" or/and golden[7].

- SpkId.m & FSAF package Fast Subband Adaptive Filtering (FSAF) - File Exchange - MATLAB Central (mathworks.com),
- Audio Toolbox Audio Toolbox - MATLAB & Simulink (mathworks.com),
- Signal Processing Toolbox Signal Processing Toolbox - MATLAB (mathworks.com) ,
- MatLab R2019b MATLAB Home - MATLAB & Simulink (mathworks.com),
- Dell i7-11700K XPS 8940 Dell XPS Desktop Computers - Desktops & All-In-One PCs | Dell Canada,
- Focusrite 2i2G3 and 18i20G1 Scarlett 2i2 | Focusrite, 24bit 48kHz mode,
- TPA3255 amplifier AIYIMA A07| 2.0 Channel | 300Wx2 | HiFi Amplifier ,
- 36V/4.5A power supply. PS was quite problematic; you will see lots of N*60Hz lines on spectrograms. "Lab" regulated stabilised power supplies fired even worse.

Recorded channels:

1. AKG P.170 / Behringer B-5
2. Apex.220

---

[5] A properly calibrated and certified anechoic chamber is always desirable but not required.

[6] 1m was originally defined as the length of a pendulum making a half-swing in 1 second (that's why g $\sim$= π^2).

[7] Vibration isolation is an almost universal problem: refrigerators, washers, driers, bicycles, cars, HVACs, engines, submarines, planes, rockets, etc use rubber | pneumatic | wire mesh | etc devices, passive or active, claiming being based on some scientific theories. Audioscience proudly rejects all those pathetic claims and declares:" The use of spikes with Hi-fi components improves sound quality… The use of spikes gives the best sound!" http://www.soundcare.no/ (now http://www.seas.no/) with prices up to and beyond $2000
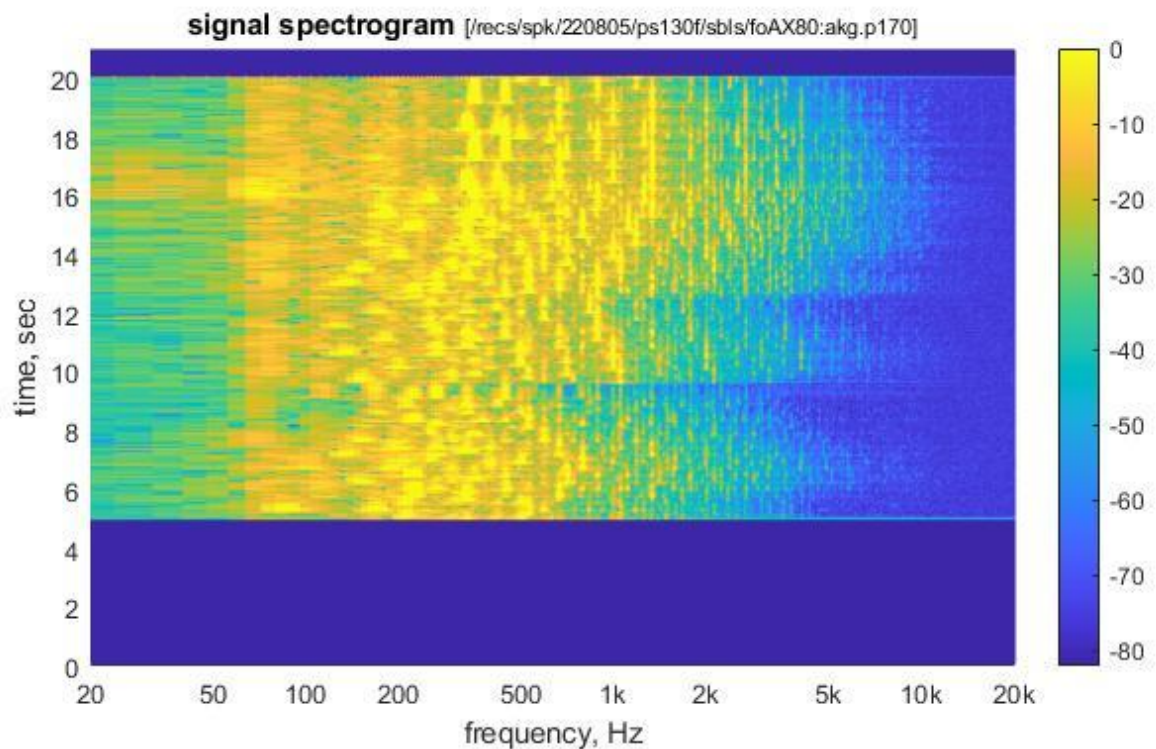
3. Amplifier output voltage
4. Amplifier output current, on 0.22 Ohm load

Excitations, 15sec each, aligned to -25dBFS RMS, on 60, 70 and 80 dBSPL, with 5 seconds of pre-silence and 1 sec of post-silence:
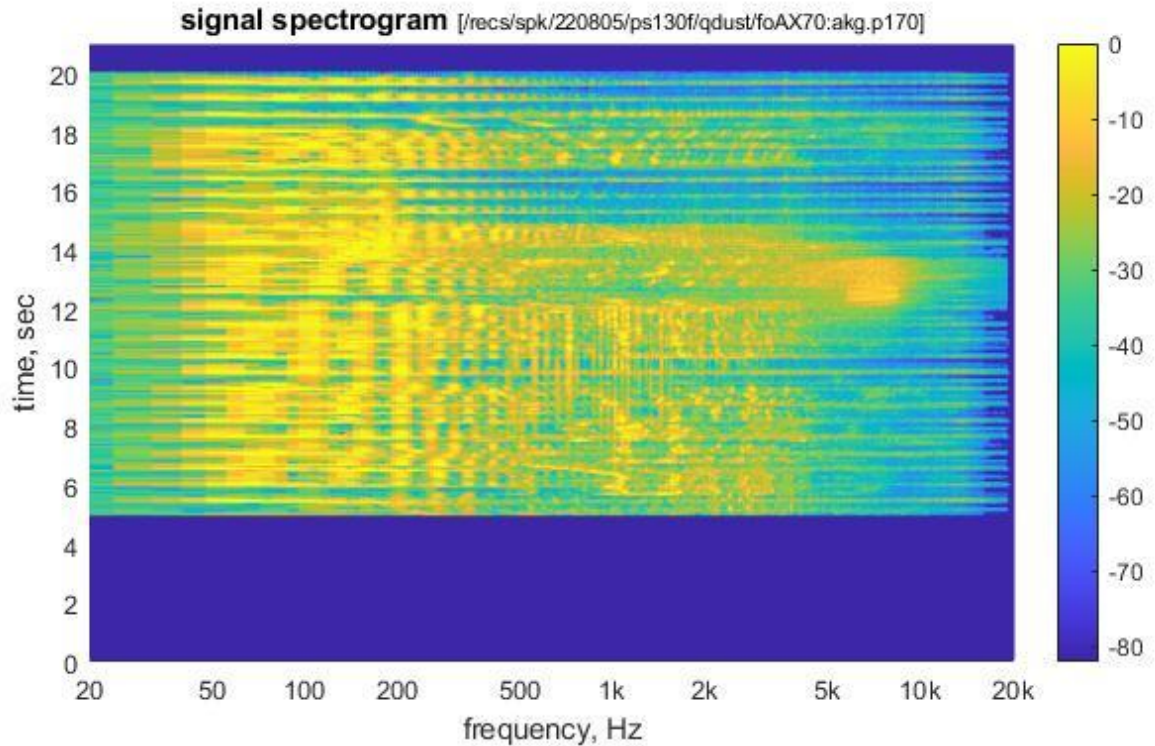
1. ITU-T© P.501 48kHz recordings to test convergence and double talk (used exactly as intended)
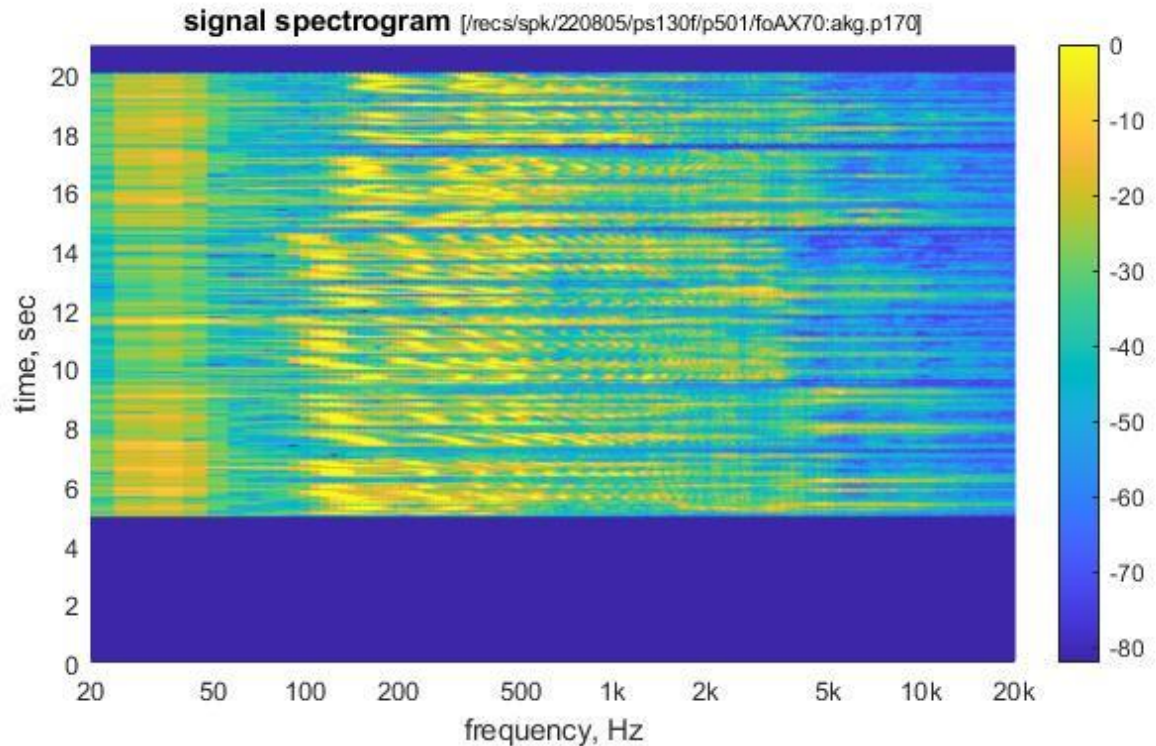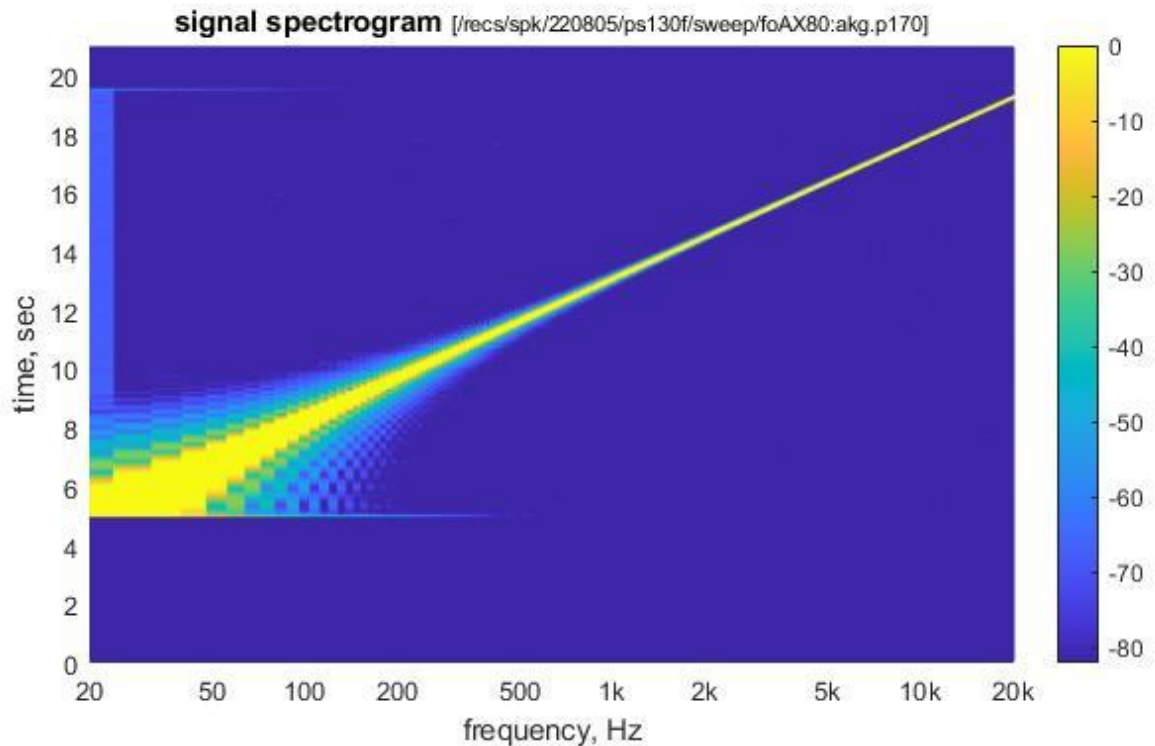

2. A Piano music clip (allegedly "fair use", copyrights of respecting owners)

3. A Rock music clip (allegedly "fair use", copyrights of respecting owners)

signal spectrogram [/recs/spk/220805/ps130f/qdust/foAX70:akg.p170]

4. Exponential chirp

signal spectrogram [/recs/spk/220805/ps130f/sweep/foAX80:akg.p170]

5. White gaussian noise with various colorations

I contacted the labels of the allegedly "fairly used" copyrighted materials to make sure they don't mind this use case. No one answered, so far. Most likely they don't care about small fish … but who knows? No recordings, however distorted, could be provided till written permissions are received.

3 RESULTS

There were ~20 drivers tested as candidates for desktop loudspeakers (a half was dropped as clearly being a junk):

- They do not have to deliver 115dBSPL as floor standers; they do not have to deliver 105dBSPL as studio monitors; 70 dBSPL RMS, peaking at 95 dBSPL, shall suffice. More is beneficial but not necessary.
- They should be capable of reproducing some base – otherwise people will not consider them for an AEC does not benefit this end user directly.
- They shall be "optimal" for reproducing speech at 60…70 dBSPL RMS.
- They are between 3" and 5.25", with sensitivities between 82 and 92dBSPL/W at 1m. Smaller low sensitivity drivers do not have a chance to have low LTI distortions at the required loudness levels, large drivers require large boxes which won't fit on the desktop.
- They shall fit into a sealed box under 5.5"W * 6"D * 18"H, the smaller the better.
- Their wholesale price shall be in a reasonable proportion to the $/1k price of a single core floating-point 400+ MHz DSP with sufficient internal memory, like Shark ADSP-21489 or similar Shark+, or similar C67xx TI DSPs, currently in the $15…30 range. There is no sense in wasting indeterminate amount of time trying to accomplish the task with < $2 drivers, nor with > $200+ drivers.
- There is also a limit, like 15…20%, most people would be ready to spend on the desktop speakers (with or without AEC), relative to the price of their laptop.

Current drive linearization potential was computed by:

- Calculating LTI residual at condenser microphone
- Calculating LTI residual at driver's current
- Calculating linear correlation function from current to microphone residual
- Subtracting correlated part from the microphone LTI residual
- Comparing this residual of the residual to the original LTI mic residual

This method does not work well if the current LTI residual does not have good SNR re sensing noise.

How they loudspeakers' LTI distortions sound? Subjectively:

- LTI distortions on a piano shift a fine concert instrument towards salon junk because proper piano overtones are not harmonical.
- Distortions on voice sound like the singer | speaker is somewhat nervous, hysterical, and untrue.
- A few days of listening to the LTI distortions make you spot them everywhere easily.
- Actually, you have been noticing the slightest nuances in the lack (or presence) of relaxation in the human voice subconsciously – but you used to ignore them in a recorded voice as inevitable.
- Then you learn to pay attention to them because you know how they sound.
- Then you learn to distinguish between driver's distortions and the real singer's (lack of) virtuosity.
- Then you learn to acknowledge that so many young people spending a small fortune on ultra-low distortions headphones do have a point.

3.1 FOCAL PS130F

Traditional 5.25" design for a closed [automotive] box, Fs=92.5Hz, sensitivity 87 dB/W/m, Mms=9.1g, Bl=4.56 N/A, effective diameter 104mm, Re=3.3 Ohm, Le(1kHz)=6.92uH (typo?).

arete [/recs/spk/220805/ps130f/wnse/brown/foAX70:akg.p170]
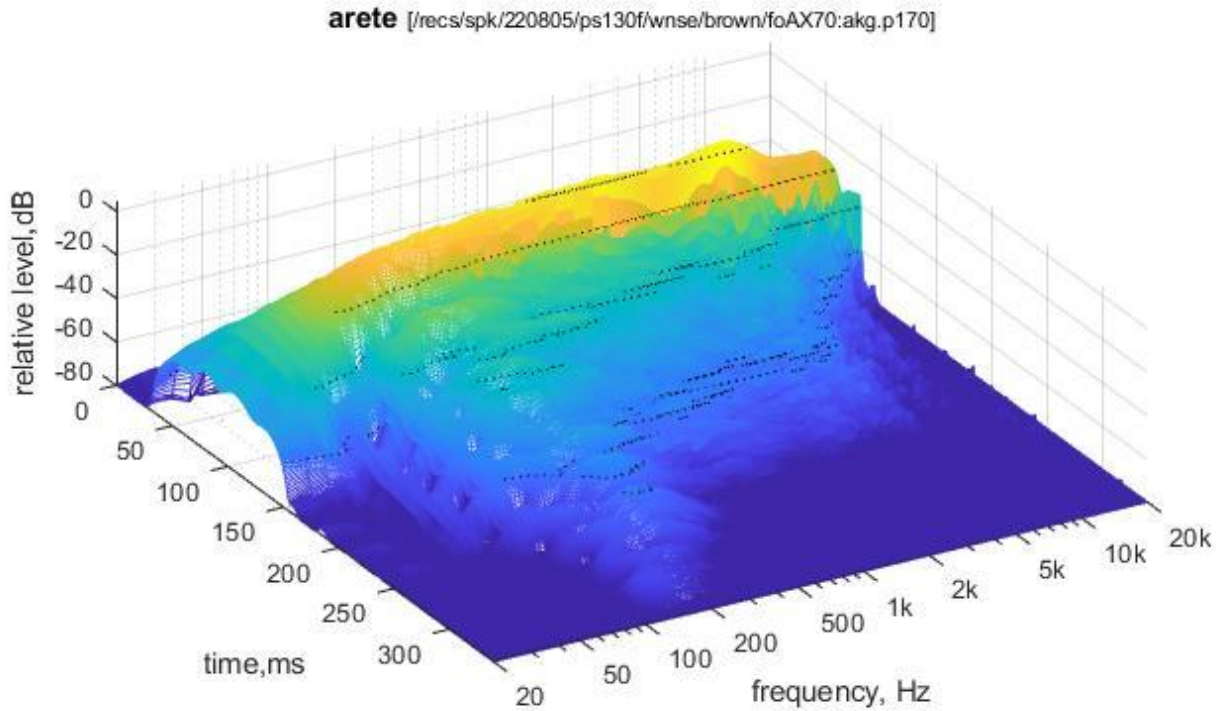
The RIR starts below -40dB for f>500Hz, everything above it – the driver itself.

3.1.1   Harmonic Distortions

sweep main and harmonics at mic:[/recs/spk/220805/ps130f/sweep/foAX80:akg.p170]
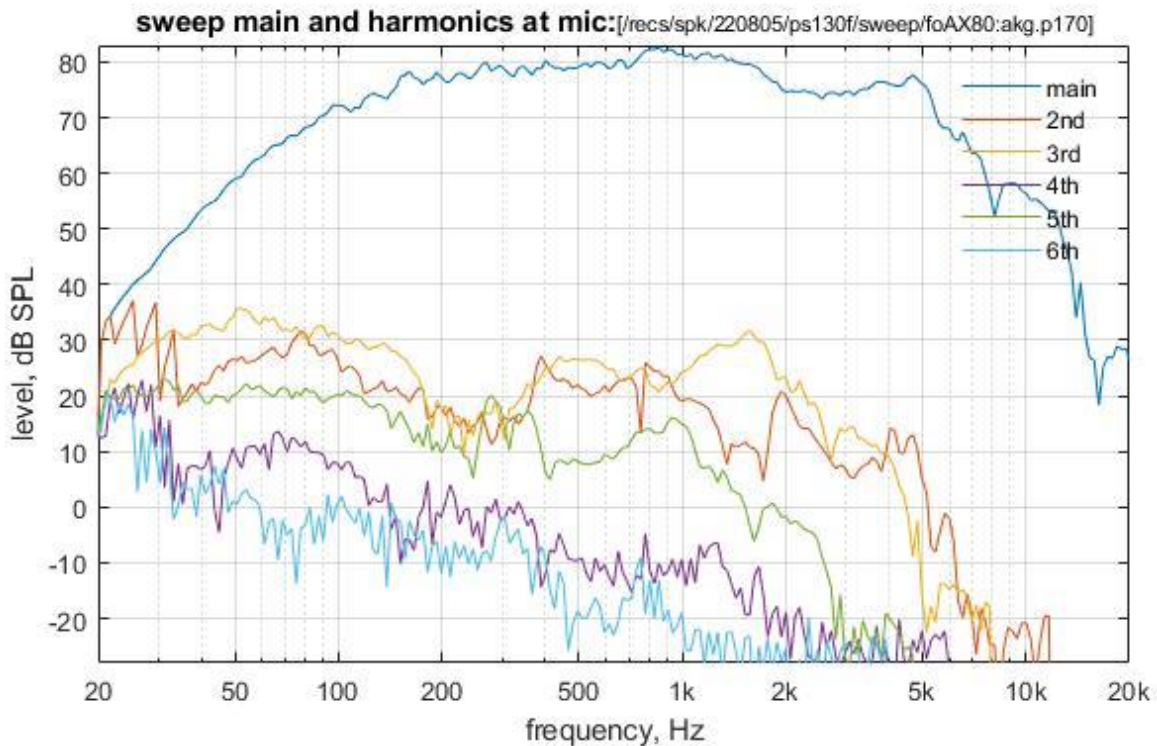
80dBSPL @ 1m; for f>150Hz, harmonic distortions are < 0.3% (-50dB). Room reflection distortions are +-2dB between 100Hz and 500Hz, and negligeable > 1kHz (see arete plot above for explanations). The line under deep valley on high frequencies is the inherent defect of the chirp | sine sweep method. Get over it.
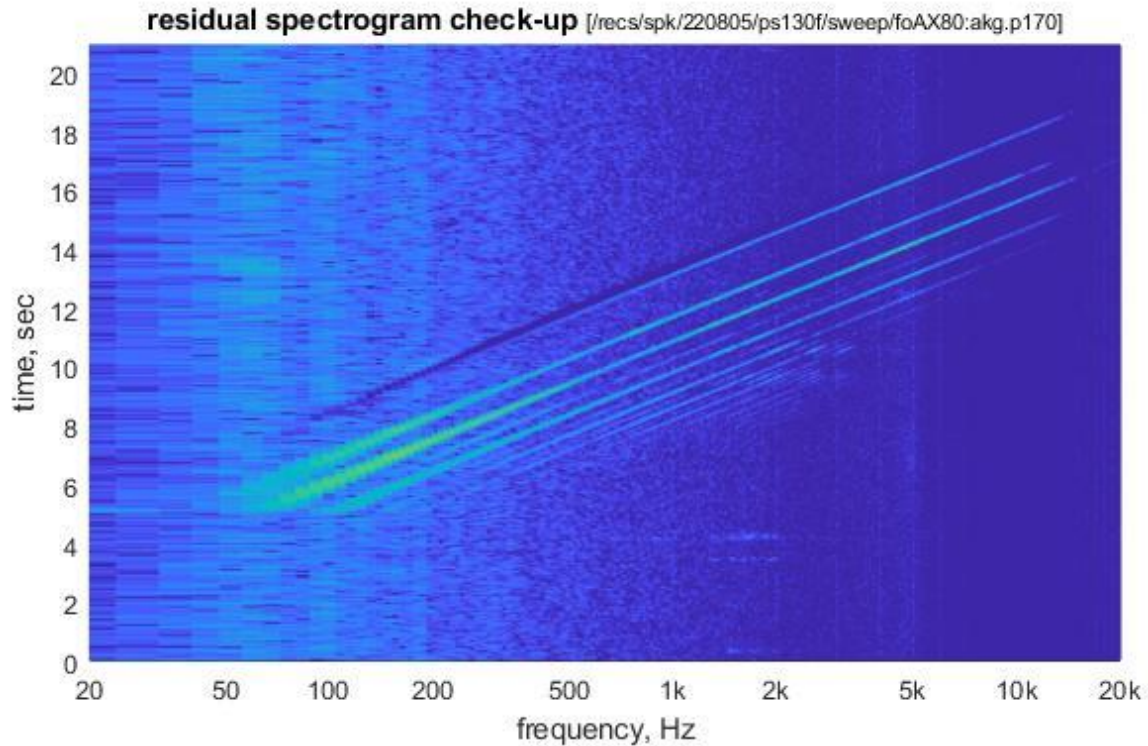
FSAF ReLS has fewer processing artifacts - but does not lower THD:
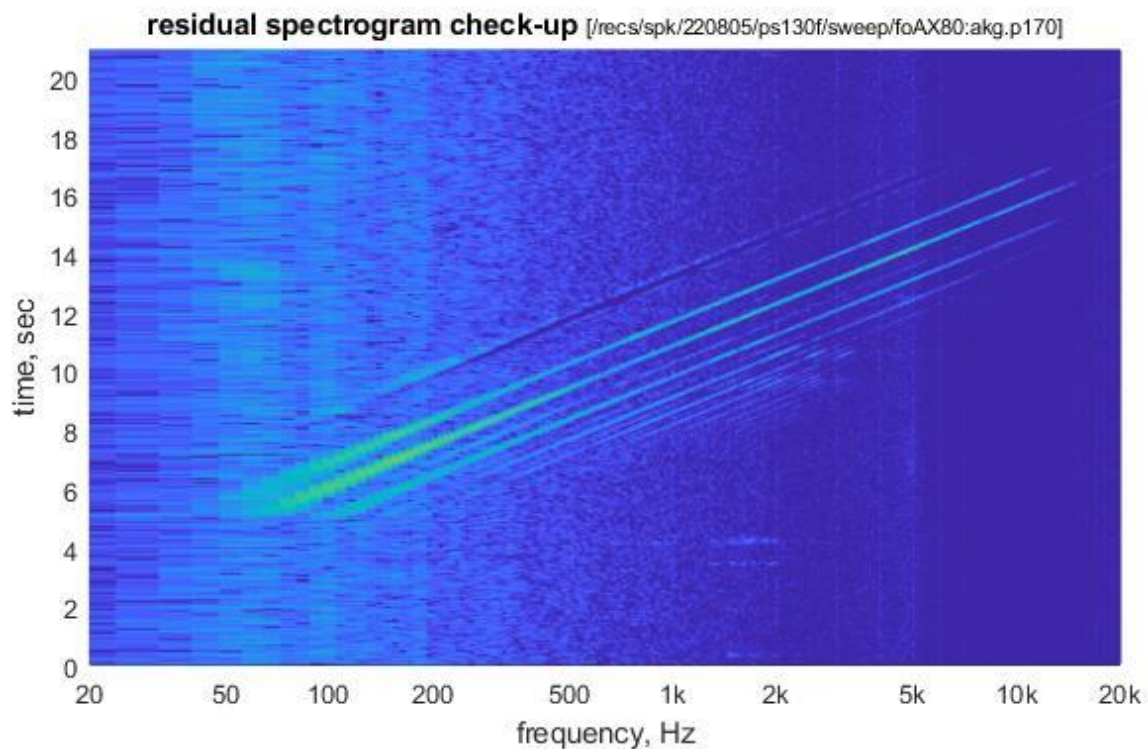
2nd harmonic re output (50:10:80) dBSPL behaves as expected, dropping ~20dB per 10dB decrease in output:
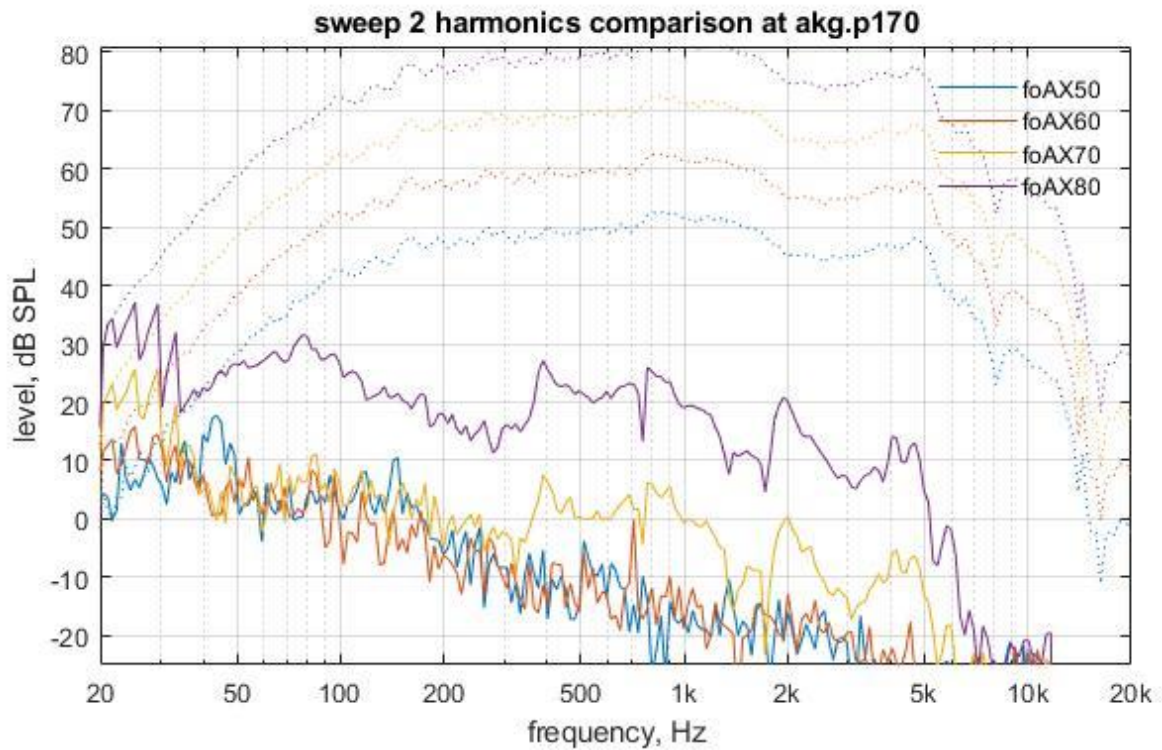
3rd harmonic re output (50:10:80) dBSPL does not behave as expected, staying about the same 50dB below output level plus ~2 dB per 10dB of output decrease:
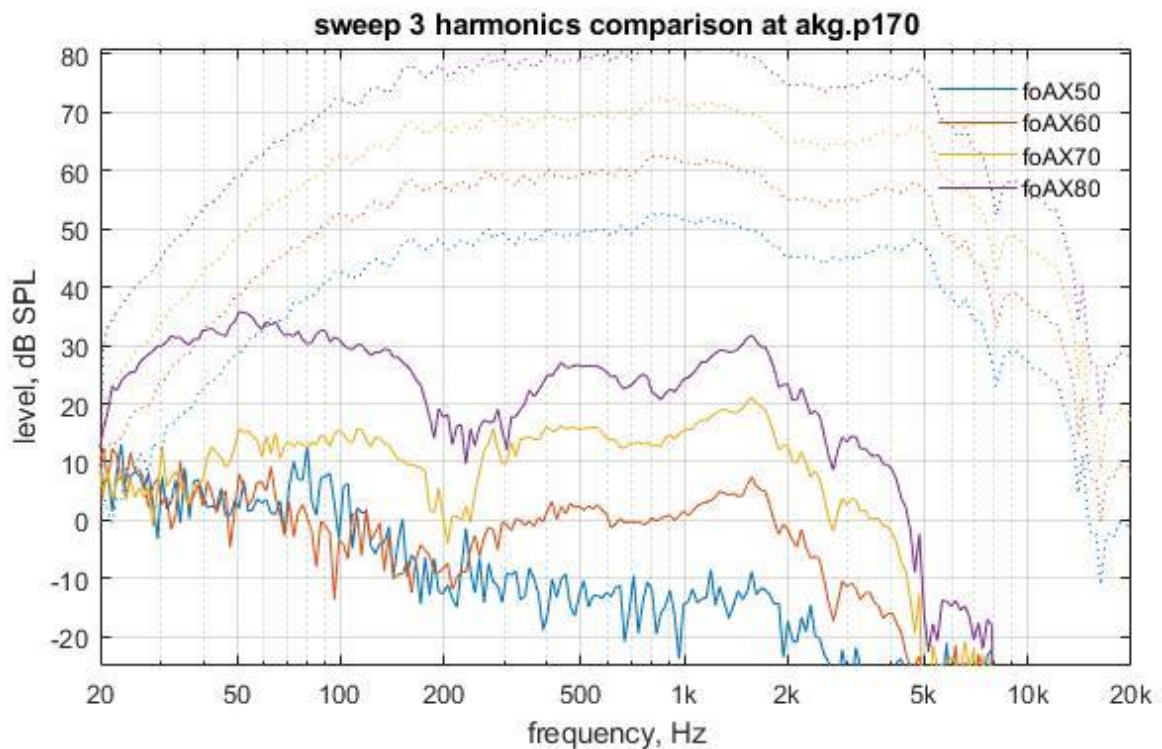
5th harmonic behaves in the same way, -65dB re output:

sweep 5 harmonics comparison at akg.p170
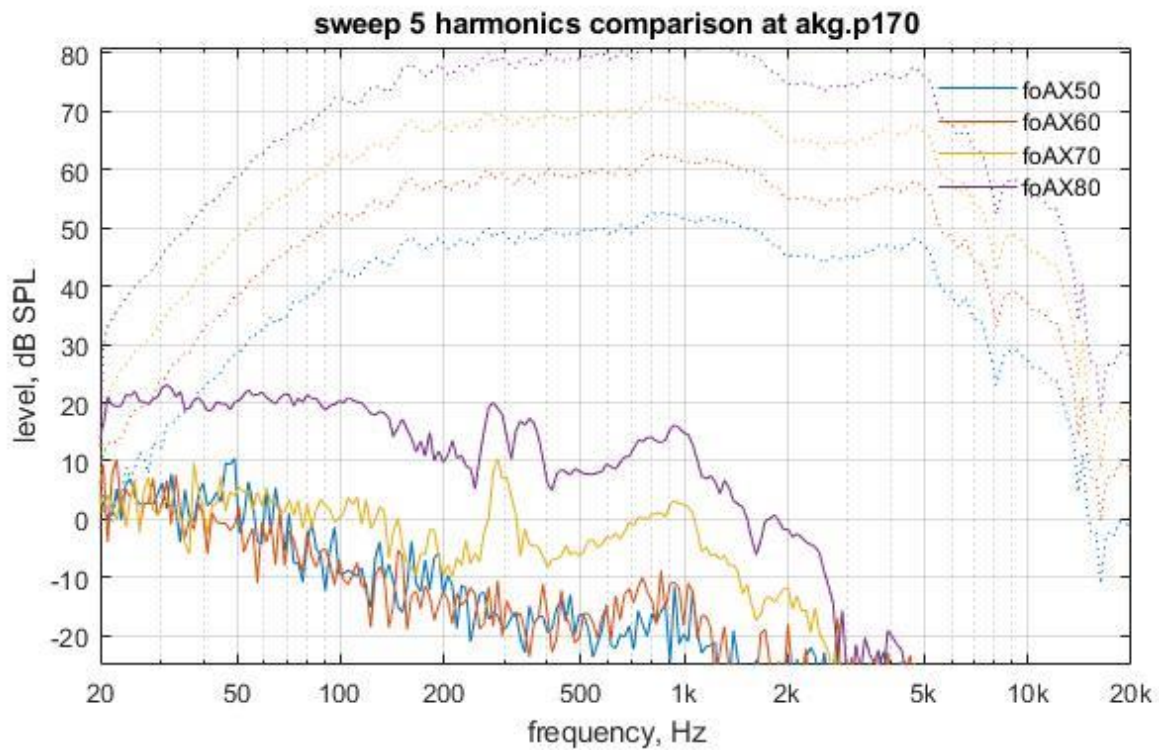
The shape of 3rd harmonic curve(f) more or less repeats the main (f*3). Harmonics do not decrease as f^{-1} nor as f^{-2}. Their shape is stupendous. How such a simple device can produce such complicated curves?

3.1.2   LTI distortions on brown noise

spk / mic / residual / nse / !LTI spectra [/recs/spk/220805/ps130f/wnse/brown/foAX80:akg.p170]
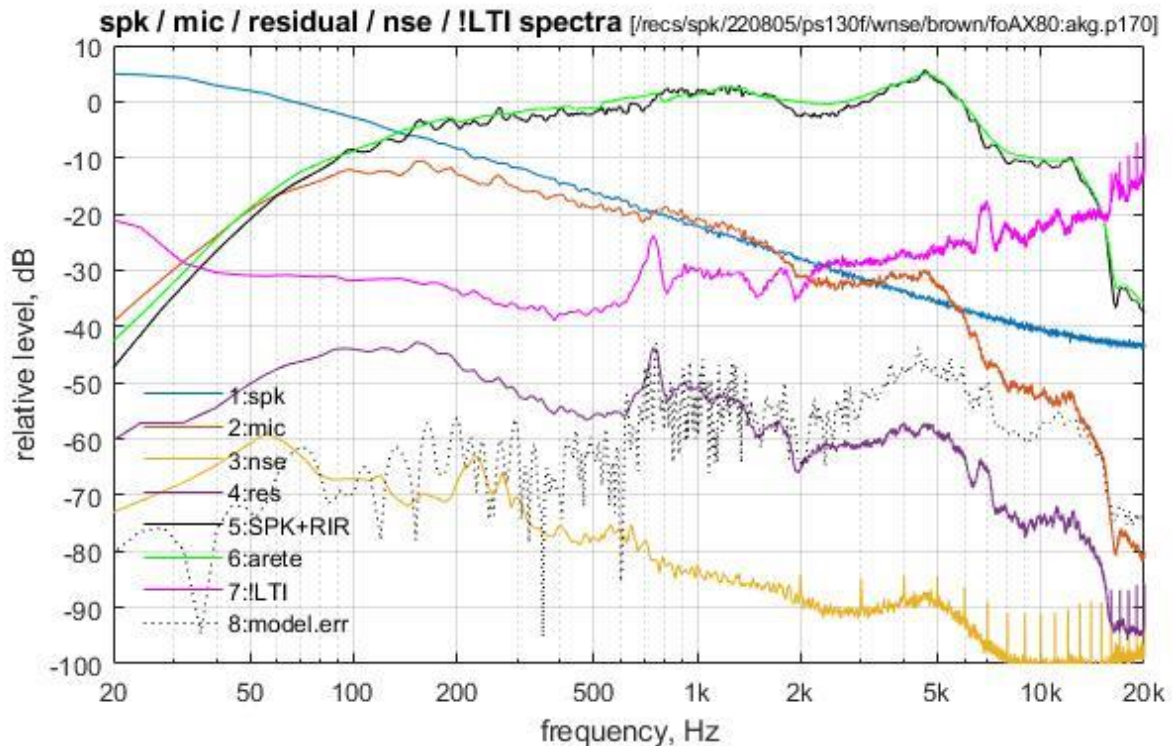
The LTI distortions are much higher than you could have imagined after studying harmonic distortions; the shape of LTI distortions in HF more or less repeats the driver FR. Ignore 240Hz noise bump below.
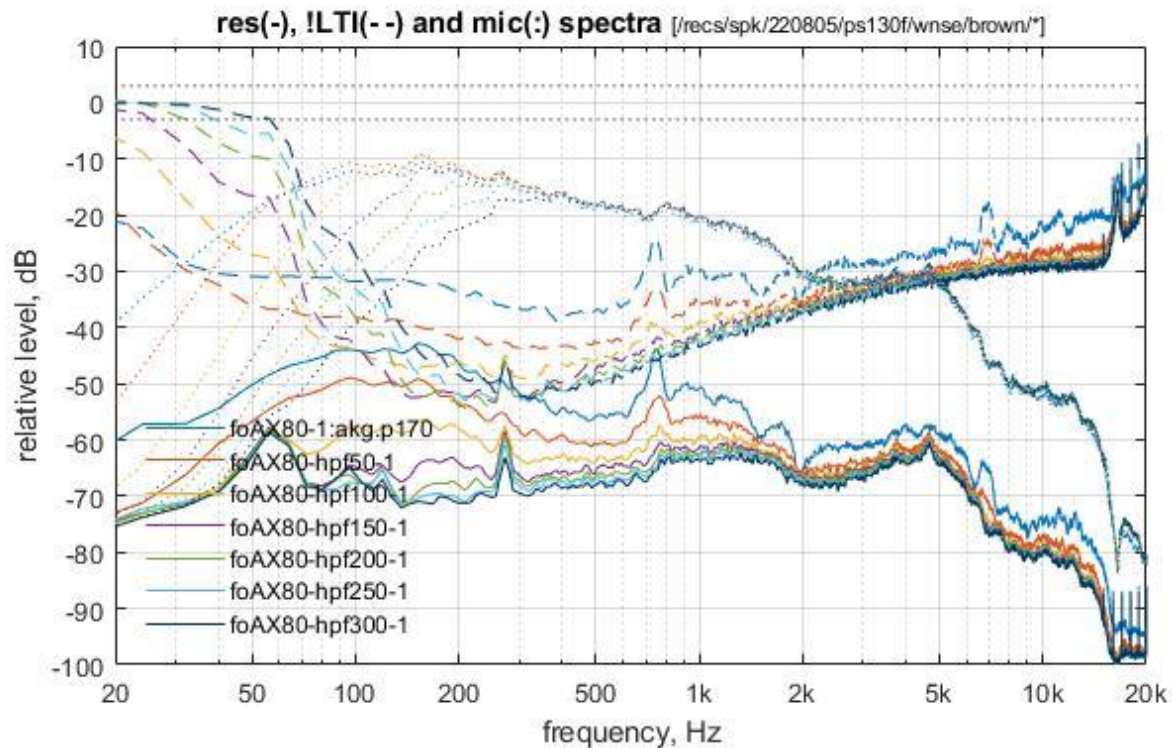
The mid-range distortions heavily depend on the level of low-frequency excitation, here progressively high-pass (0:50:300) filtered with spike at 750Hz disappearing.

As a mid-range (250Hz…2.5kHz) driver, PS130F easily survives up to 90dBSPL RMS (105 peak) and degrades only when pushed to 115dBSPL.
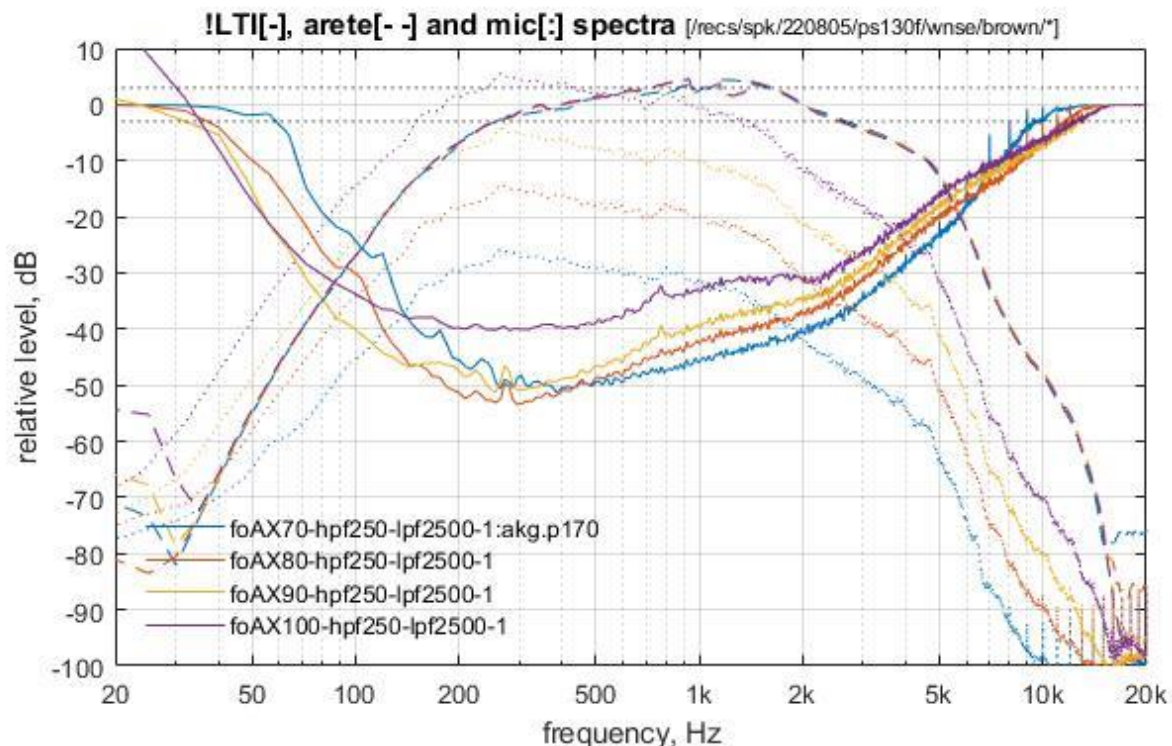
Before you attempt to use any driver in 2-way configuration, see how its LTI distortions grow with bandwidth widening, here on 70dBSPL, for the same HPF=50Hz, and increasing LPF=[250 350 500 700 1000 2500]:
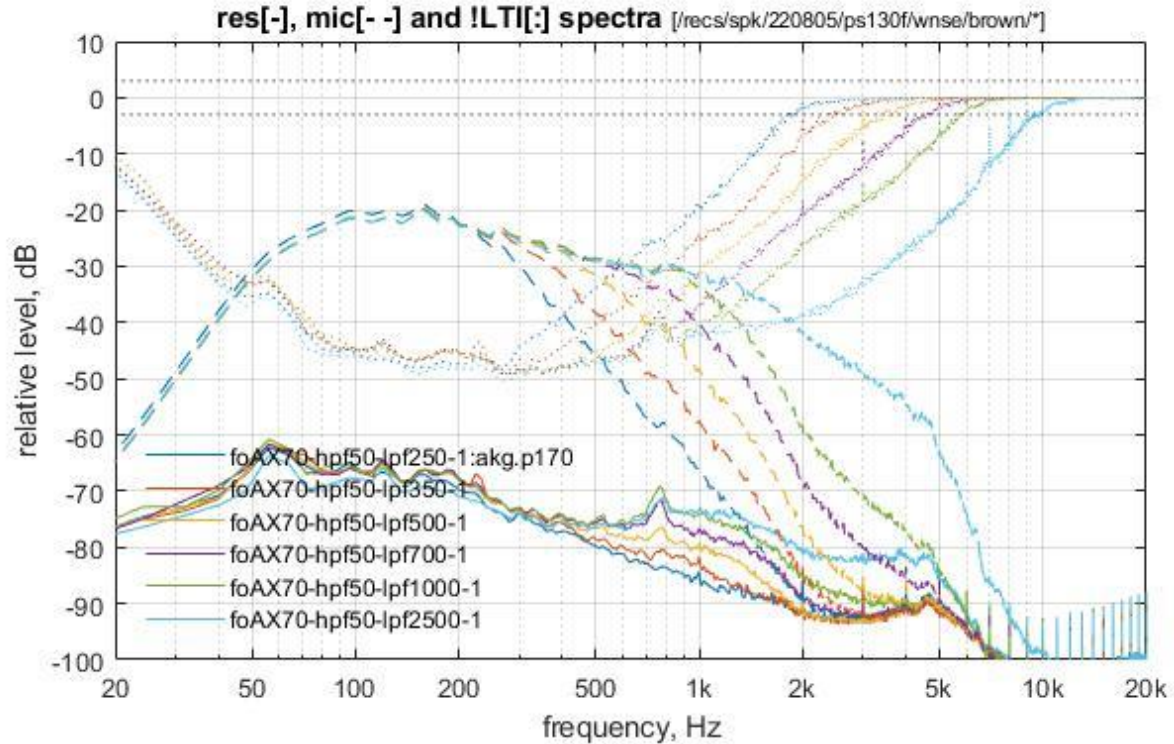
And the same on 80dBSPL:
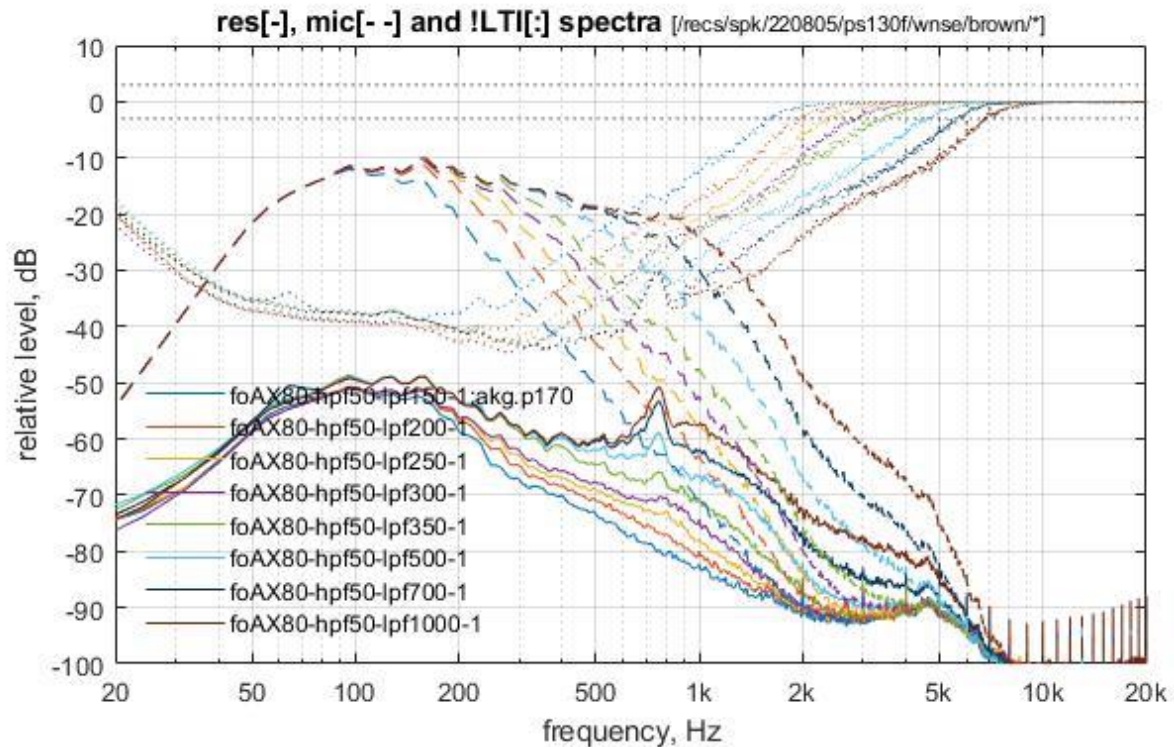
AEC needs to know how wide LTI distortions spread. Here we feed an octave wide noise and watch the LTI distortions of the lowest octave 63-125Hz excitation spread beyond 2kHz, etc:
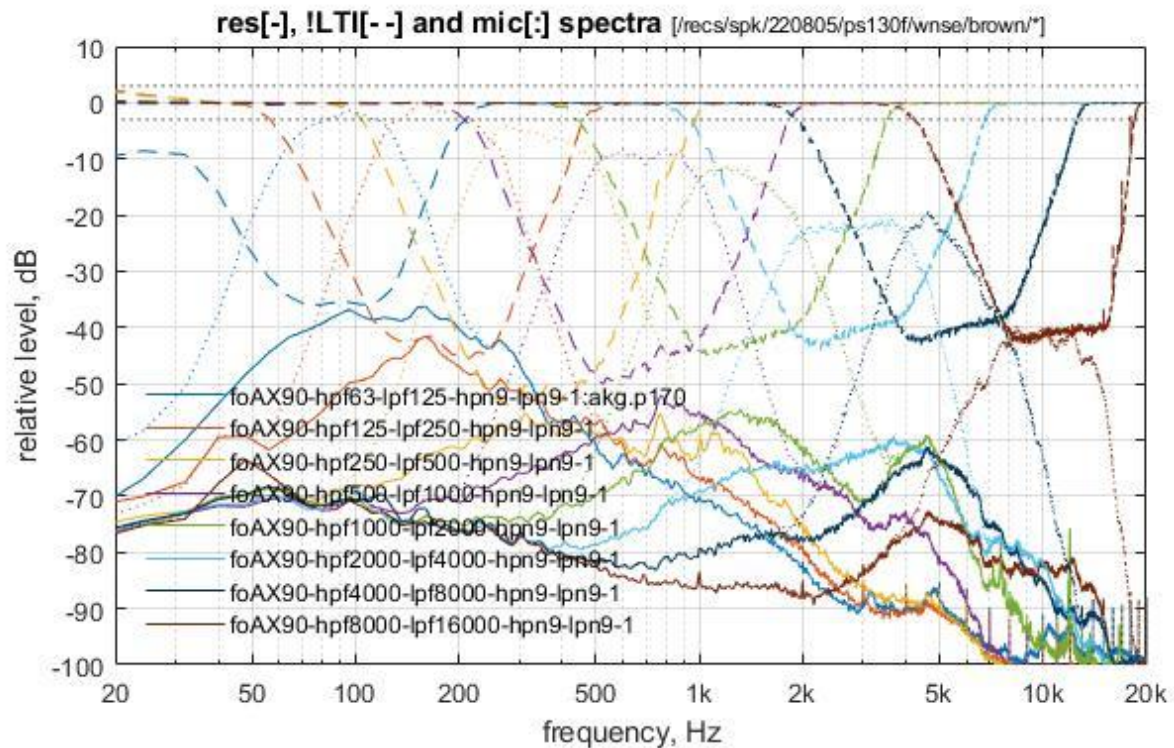
Linearization potential for mid-range usage on (70:10:90) dBSPL is up to 20dB:
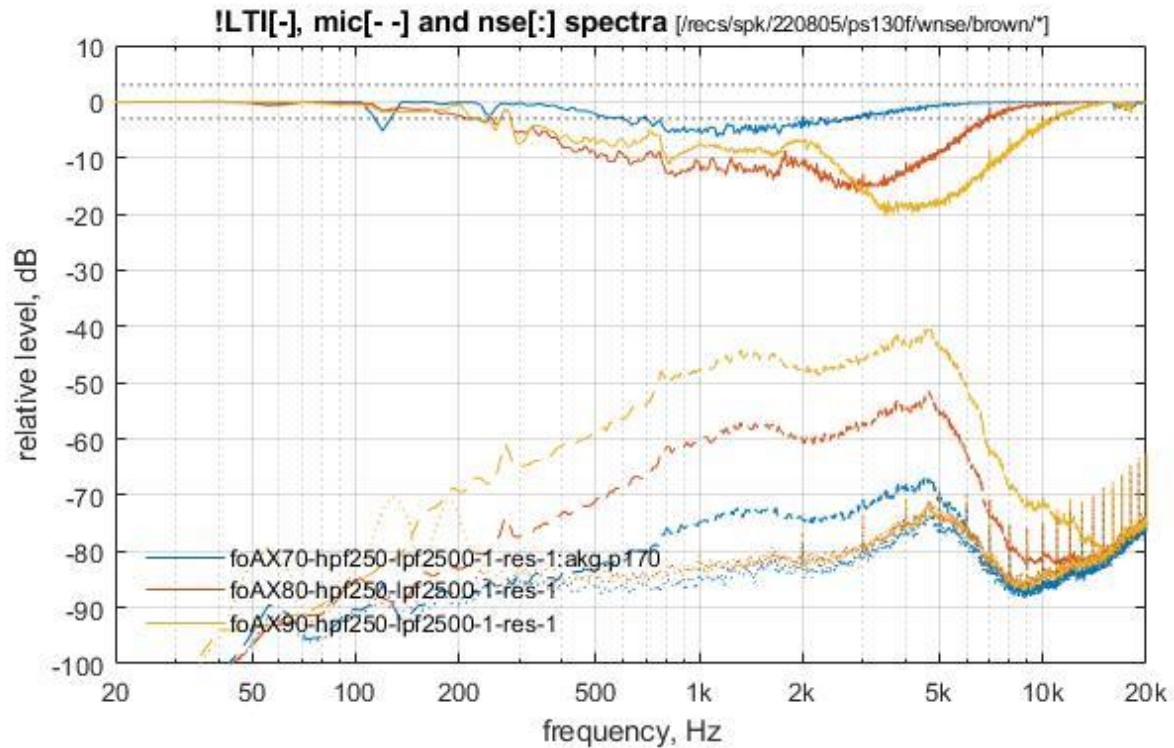
3.1.3   LTI distortions on ITU-T P.501 speech
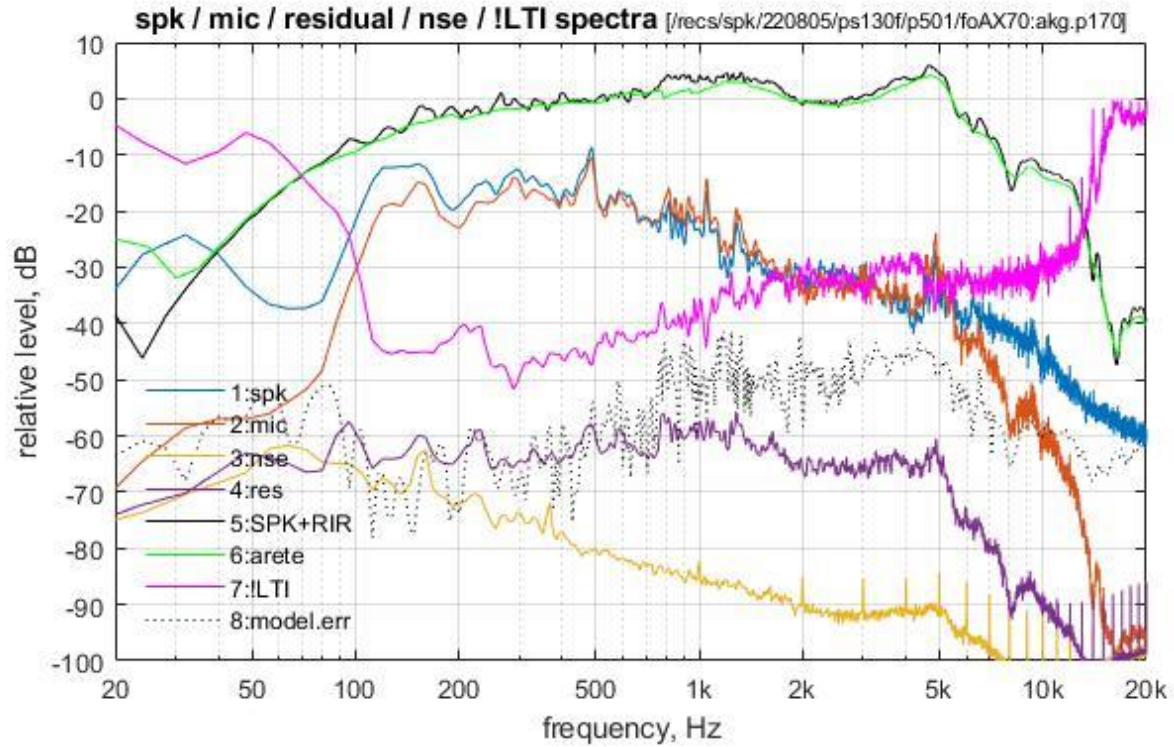
Spectrogram for 70 dBSPL shows LTI distortions on vowels. For an AEC with lower direct-to-reverberant ratio, LTI distortions will spread over following weak fricatives and plosives:
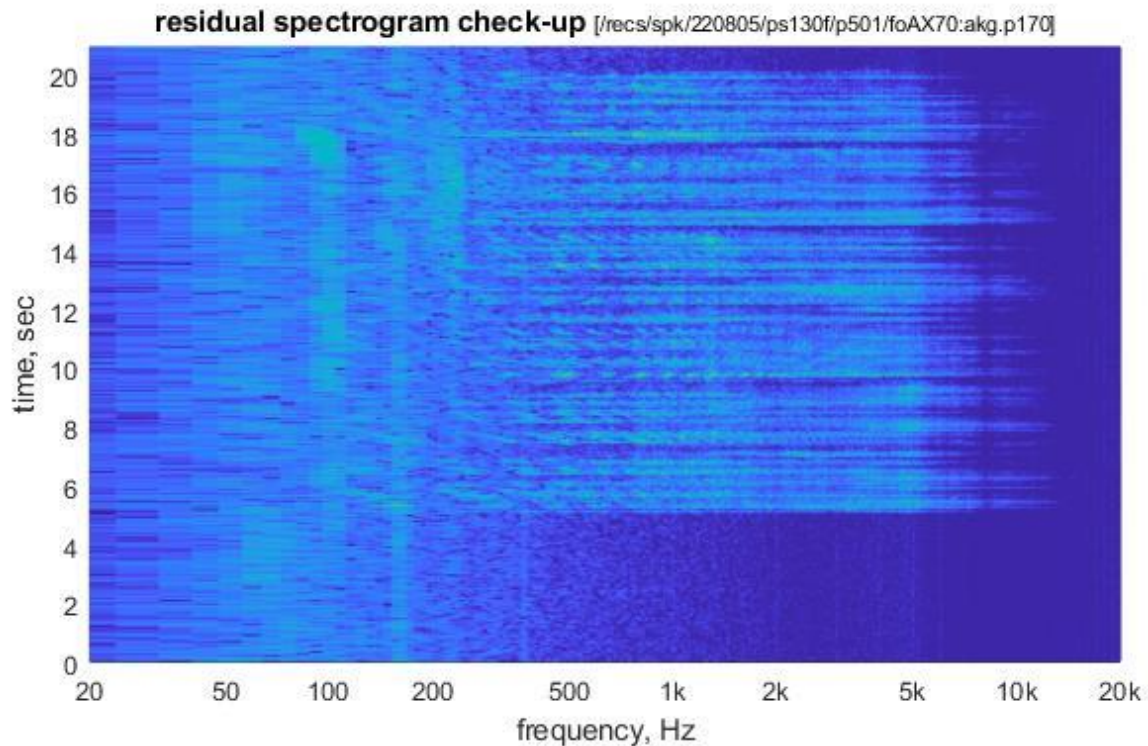
LTI distortions' SNR changes very little, if any with varying the signal level from 50 to 80dBSPL:
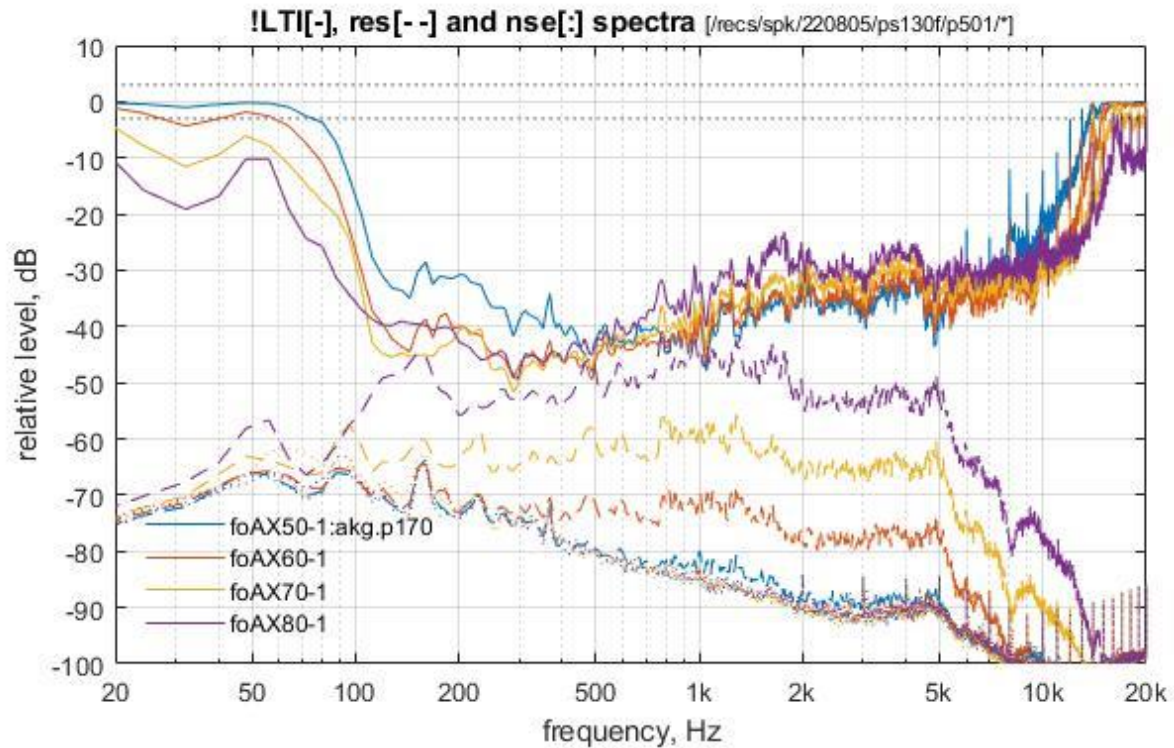
...which means they are mostly defined by the same factor which defines constant SNR 3ʳᵈ harmonic.

The LTI distortions for the case of using the driver only as a woofer:
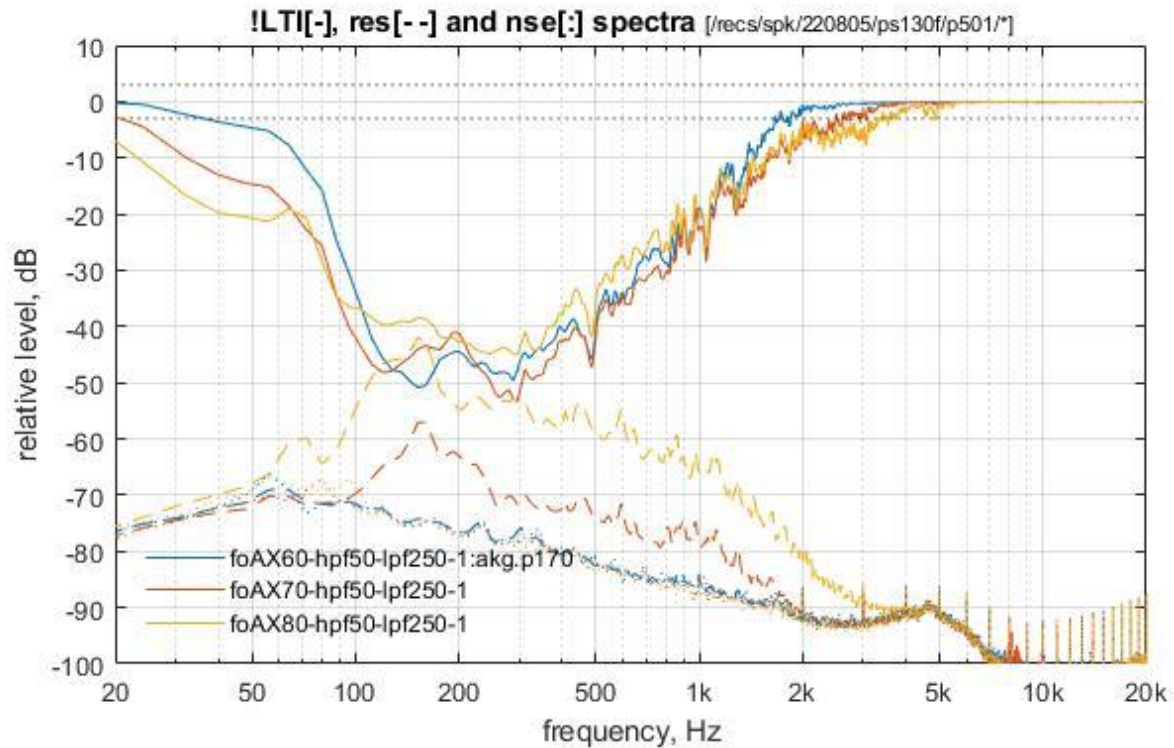
And only as mid-woofer:
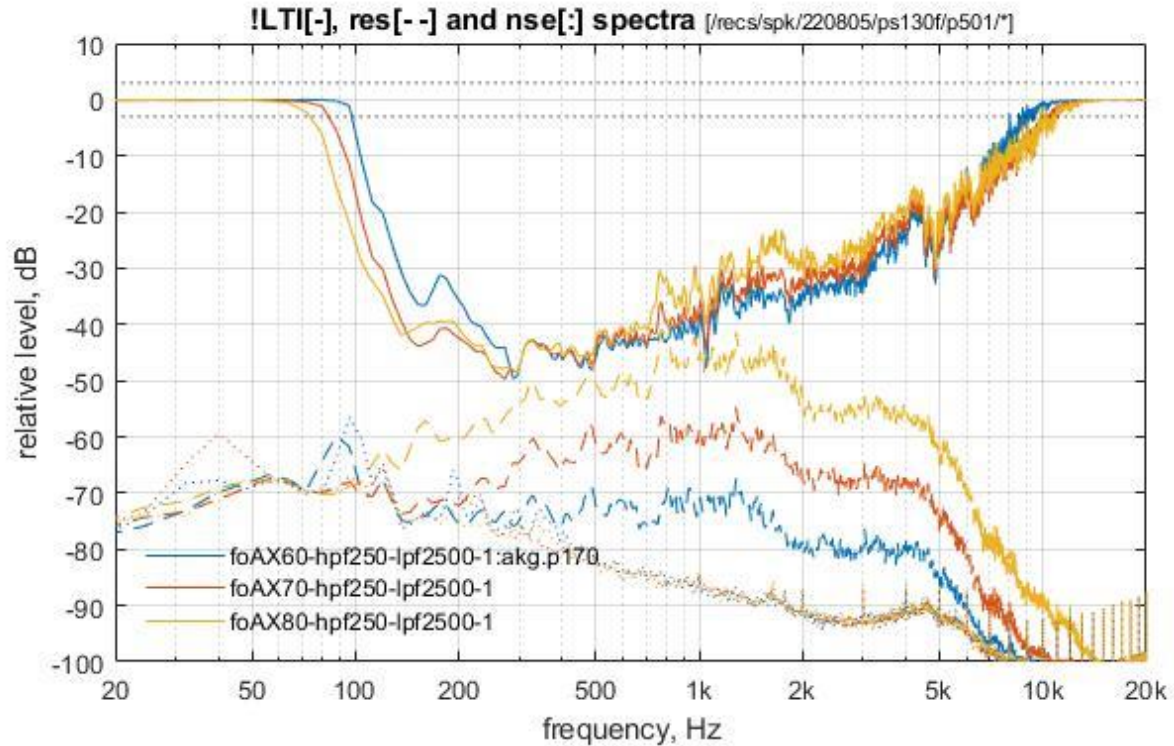
... does not gain much, if any, vs using of this driver in a 2-way design - on speech.

Residual spread for octave-wide speech excitation, with very long high-frequency tails:
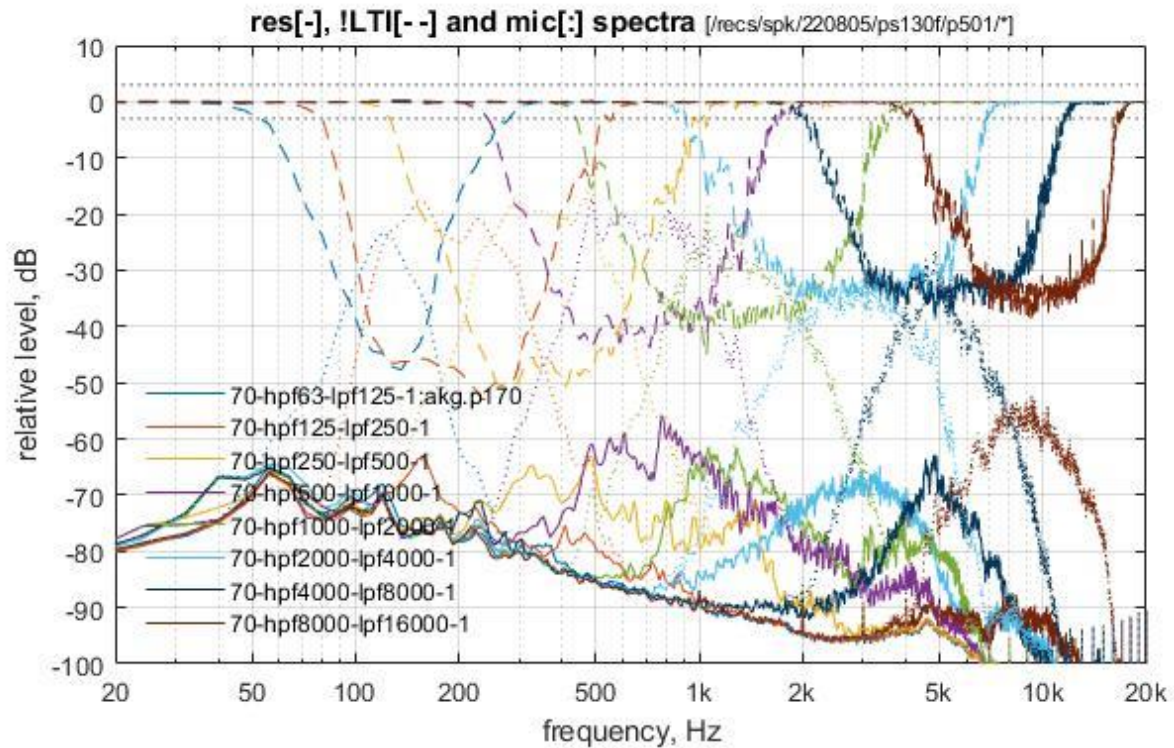
3.1.4   LTI distortions on a Rock music clip
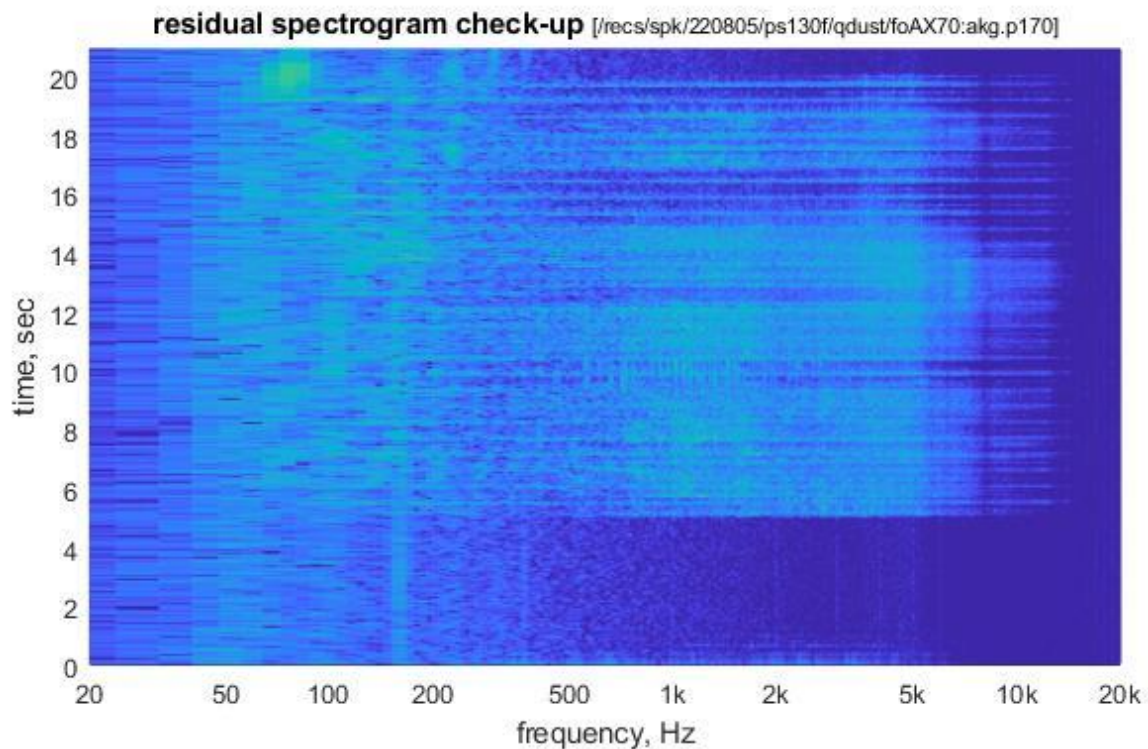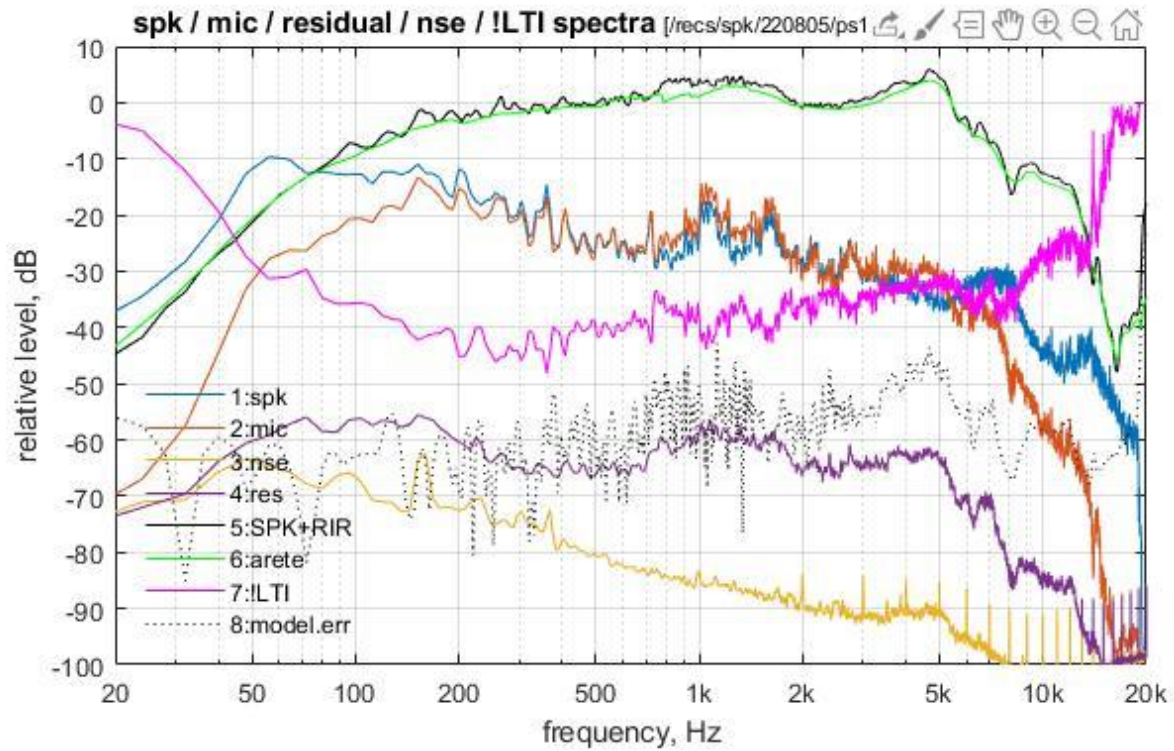
The peak to average ratio is lower therefore the LTI distortions are lower too.

On rock music, LTI distortions also show little dependence on volume:

!LTI[-], res[- -] and nse[:] spectra [/recs/spk/220805/ps130f/qdust/*]

Legend:
foAX50-1:akg.p170
foAX60-1
foAX70-1
foAX80-1
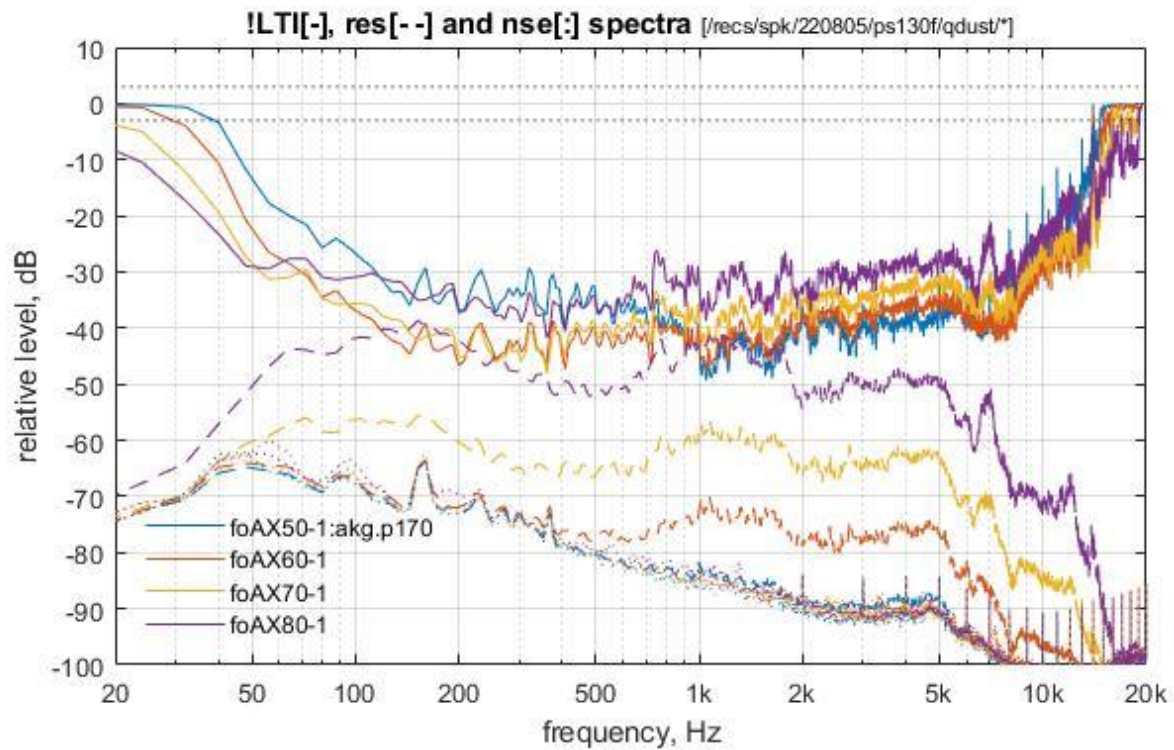
There are benefits to splitting 2-way into woofer and mid-woofer but they do not exceed 8dB in mid-frequencies:

!LTI[-], res[- -] and nse[:] spectra [/recs/spk/220805/ps130f/qdust/*]

Legend:
foAX60-hpf250-lpf2500-1:akg.p170
foAX70-hpf250-lpf2500-1
foAX80-hpf250-lpf2500-1
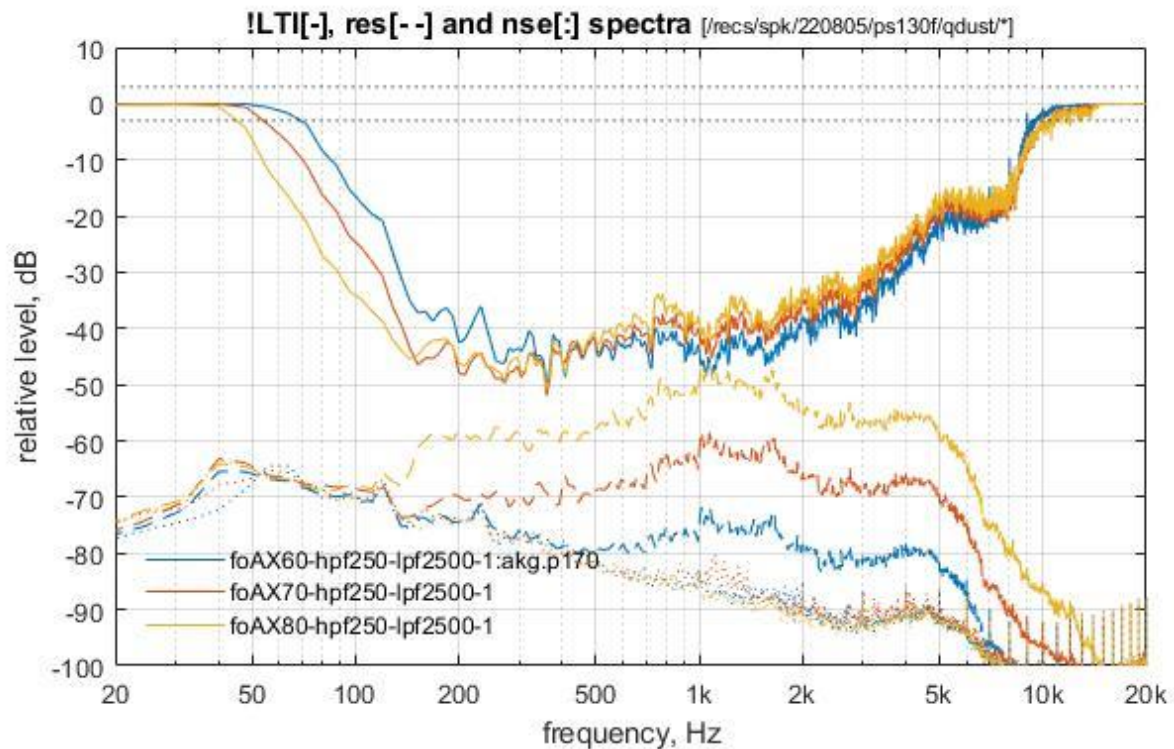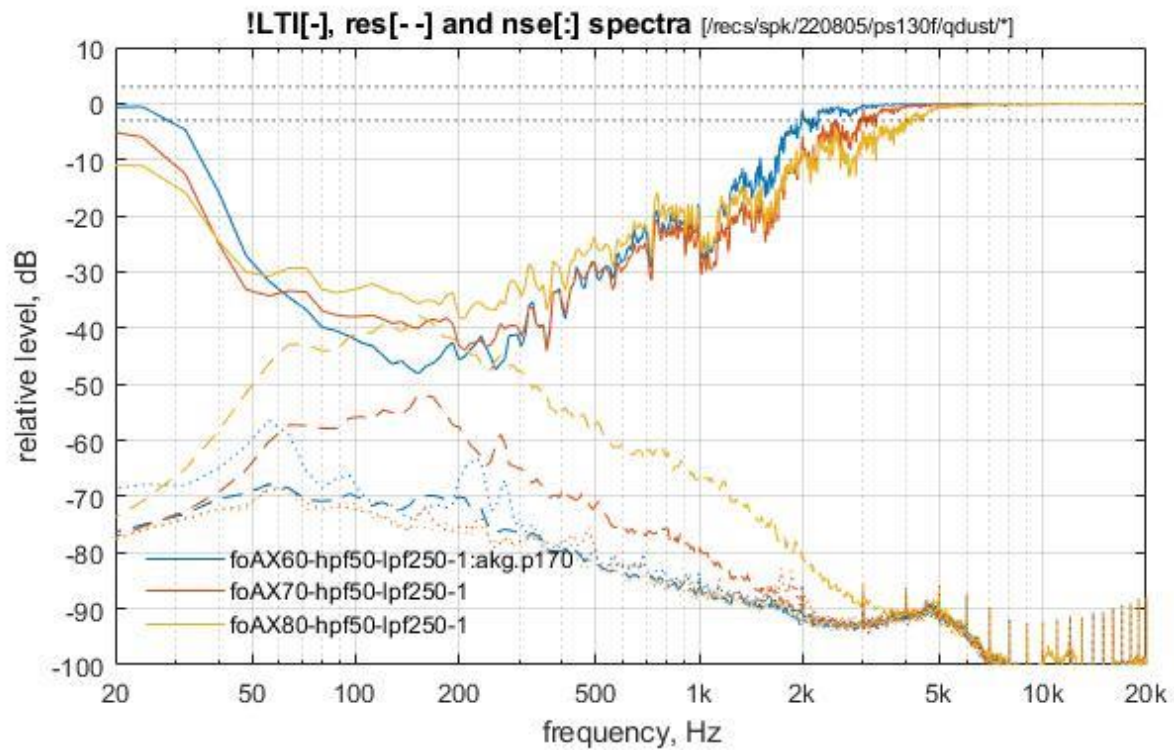
Linearization potential for PS130F as the woofer is about 10...15dB:

… same for mid-woofer:

!LTI[-], res[- -] and mic[:] spectra [/recs/spk/220805/ps130f/qdust/*]

foAX60-hpf250-lpf2500-1-res-1:akg.p170
foAX70-hpf250-lpf2500-1-res-1
foAX80-hpf250-lpf2500-1-res-1
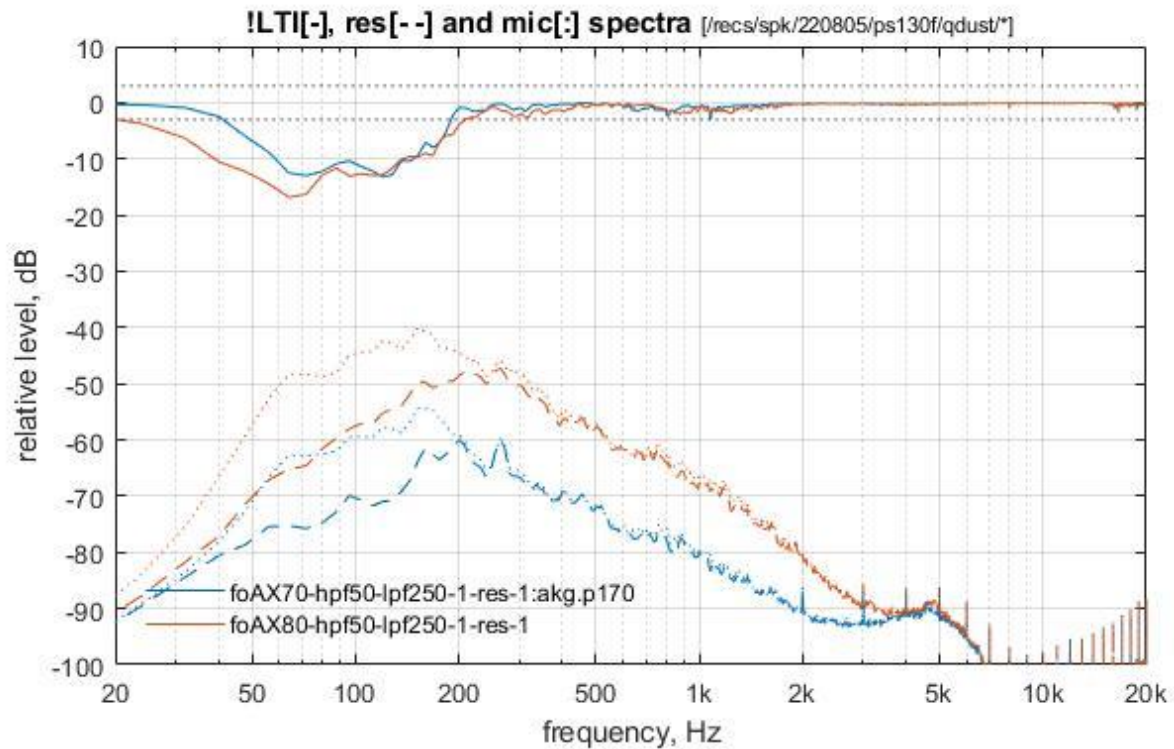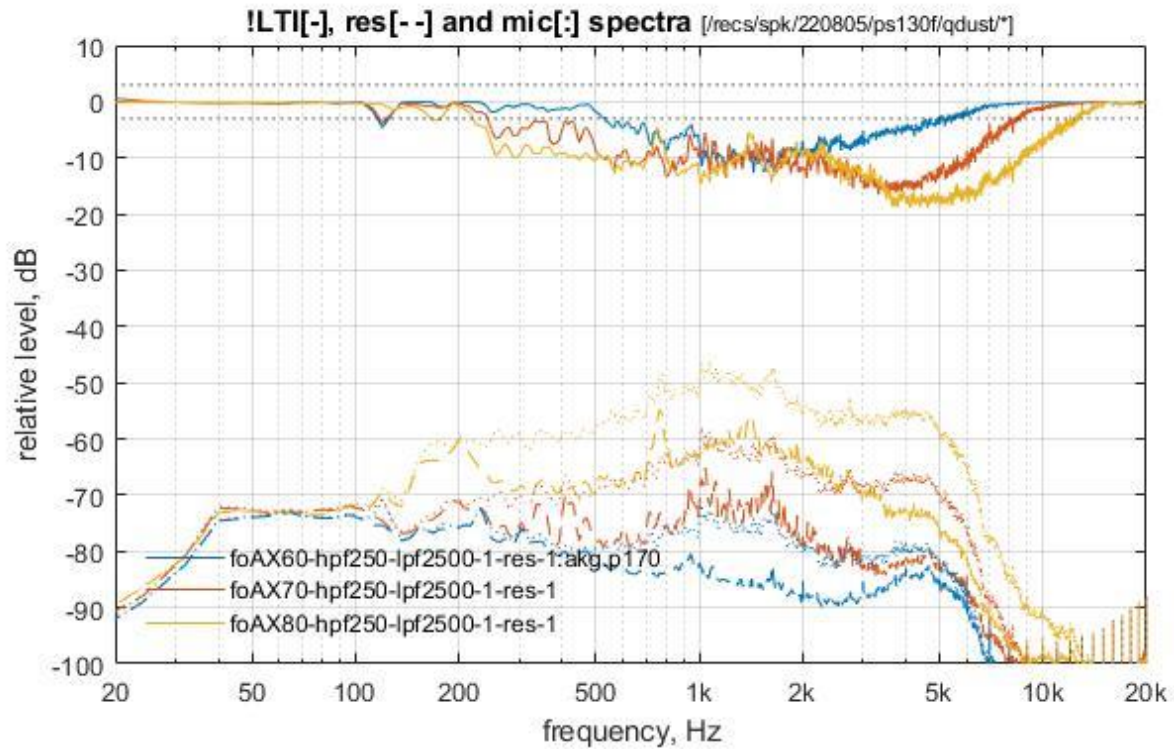
… and almost the same for full-band:

!LTI[-], res[- -] and mic[:] spectra [/recs/spk/220805/ps130f/qdust/*]

foAX60-1-res-1:akg.p170
foAX70-1-res-1
foAX80-1-res-1
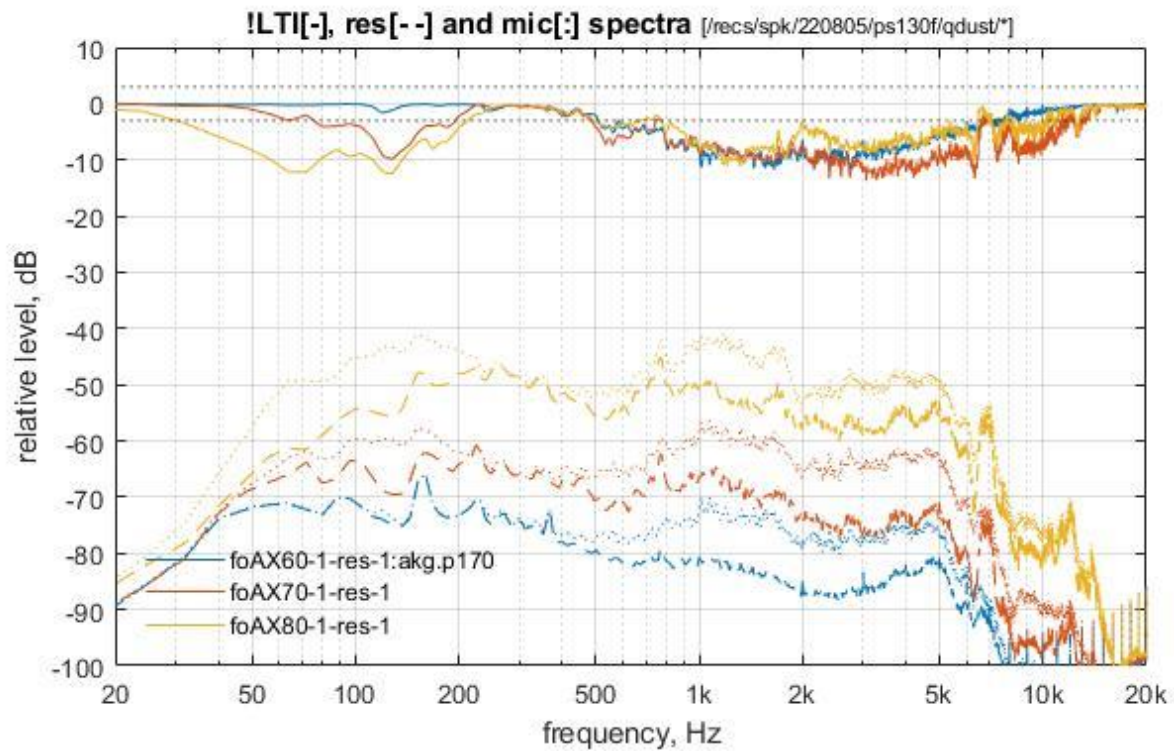
Interesting that current drive linearization is better if the signal bandwidth is narrower. Here is distortion spread graph for octave wide excitation:
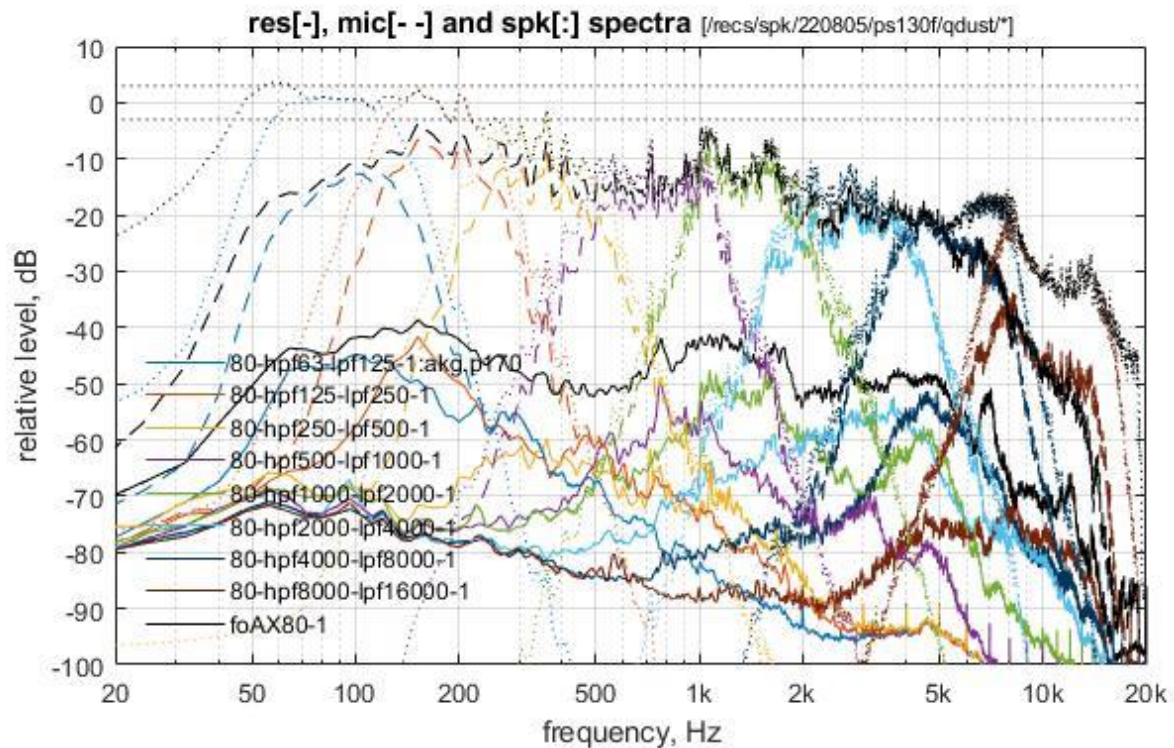
.. and here are the same LTI residuals after being subdued by current feedback (except for the stubborn 750Hz spike):
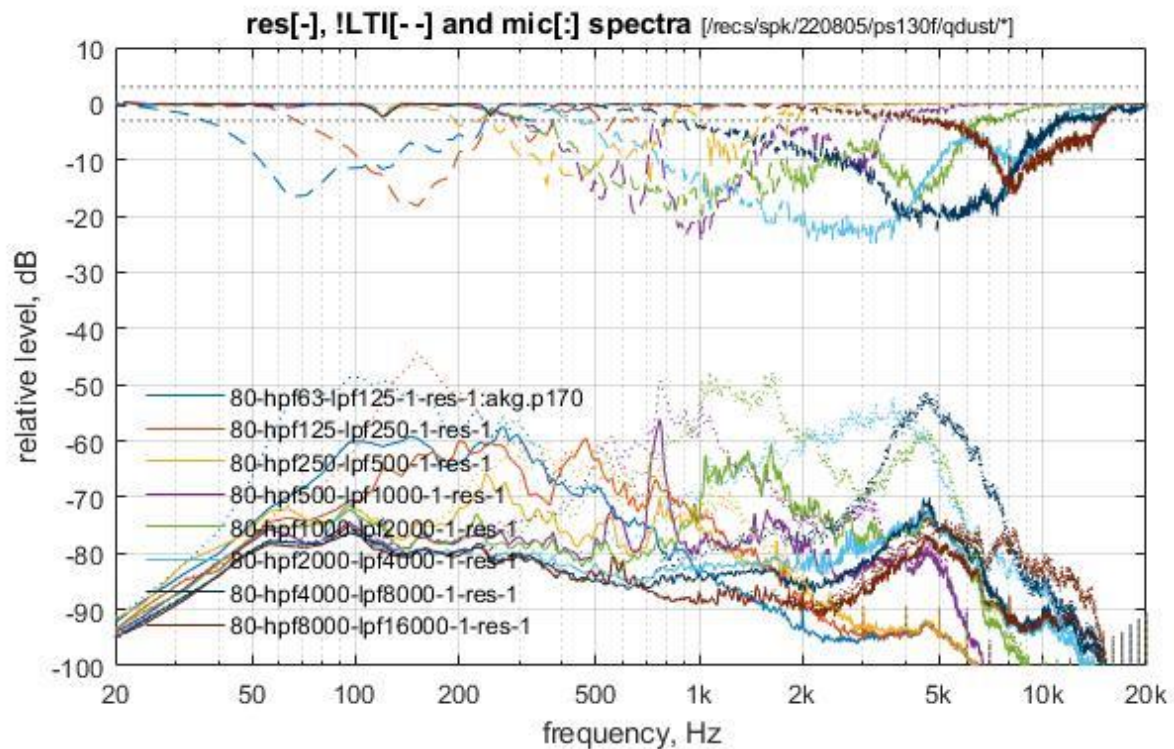
It appears that the more ways is your loudspeaker, the better.

3.1.5   LTI distortions on a Piano music clip

In the essence, LTI distortions follow the program's maximums, keep the same SNR re signal regardless of the program (speech / rock / piano / classical / etc), influenced by low frequencies, can be diminished by driver's current drive linearization which works better when the signal bandwidth is narrower.
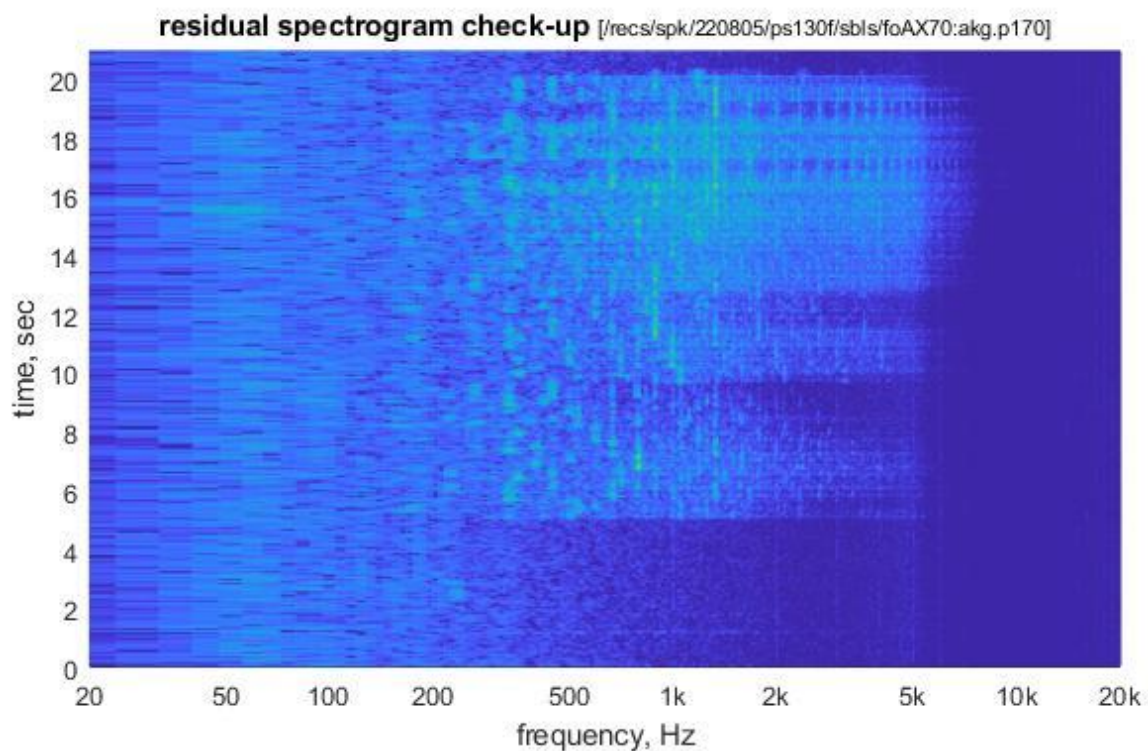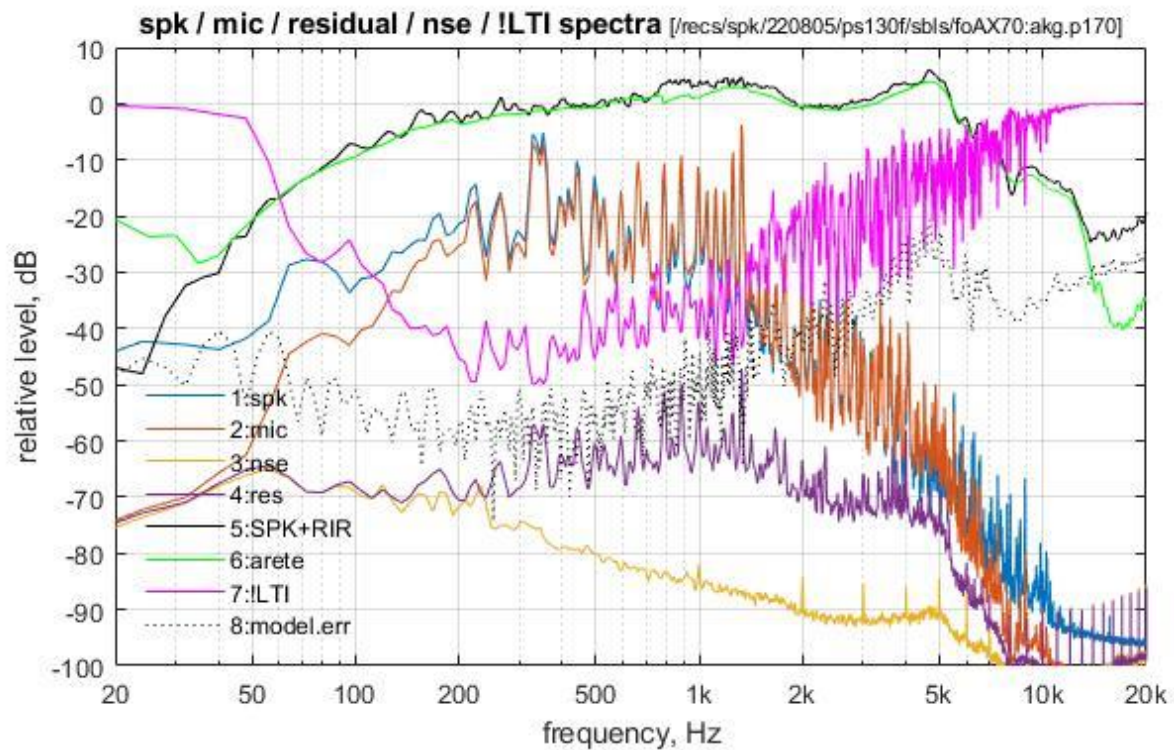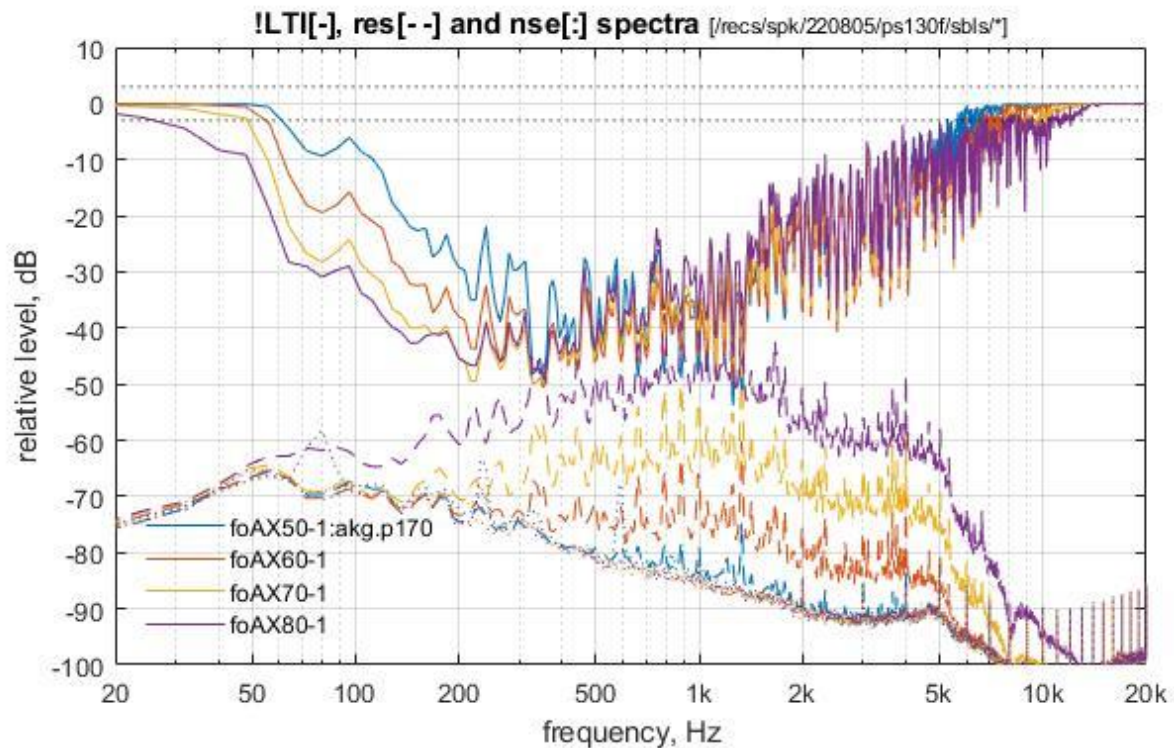
I could have added many more illustrations but I don't feel they are strictly necessary.

Focal PS130F is a good all-around 5.25" mid-woofer without any significant flaws. It appears that Focal engineers have done their share of mistakes but learned from them – which is very unusual for audio equipment vendors.

3.2   HI-VI RESEARCH DM-7500

"SWANS has quietly developed into one of the world's leading loudspeaker manufacturers". Swan (swanspeakers.com); hivi.com - is not a wild exaggeration. DM-7500 is a midrange dome tweeter with, closed box. Fs = 300Hz, Re =4.1 Ohm, diameter 75mm (same size as voice coil), sensitivity 92dB (SPL on 1W at 1m) which is a somewhat non-traditional design. Well worth considering for AEC despite some minor deficiencies.

3.2.1   Harmonic distortions
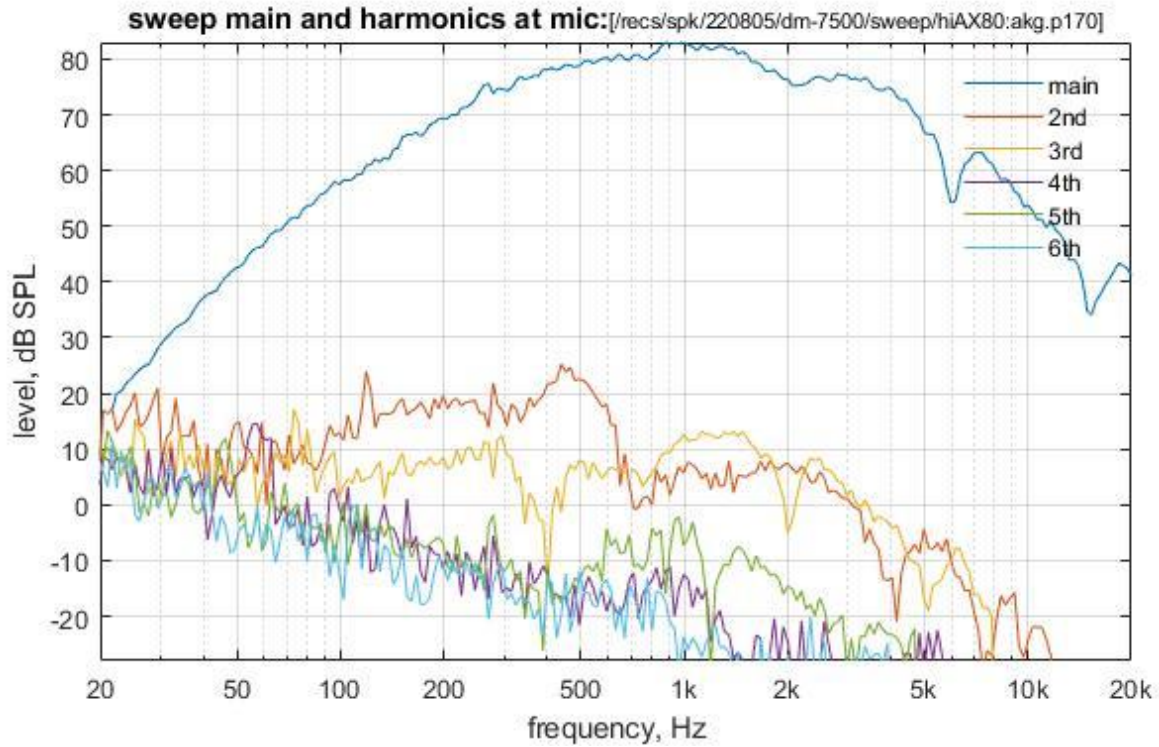
The harmonic distortions are 10-15dB lower than in PS130F. First resonance of dome breakup at 6kHz.
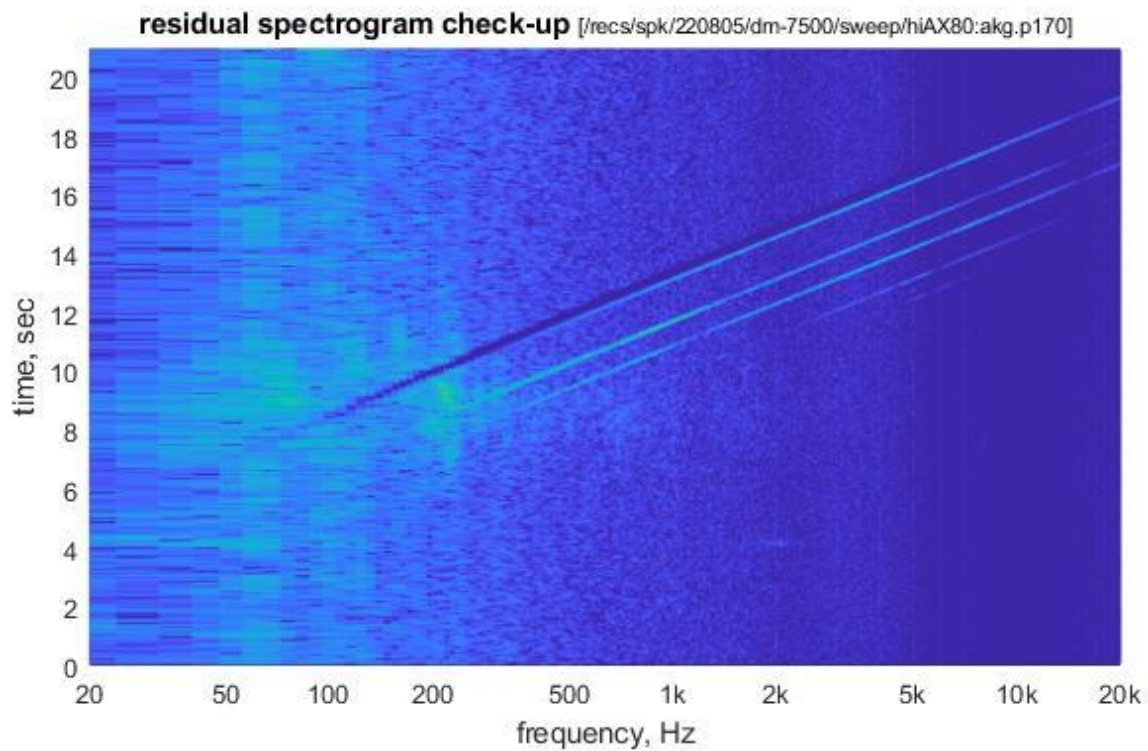
The same problem with 2nd and 3rd harmonics:

sweep 2 harmonics comparison at akg.p170

sweep 3 harmonics comparison at akg.p170
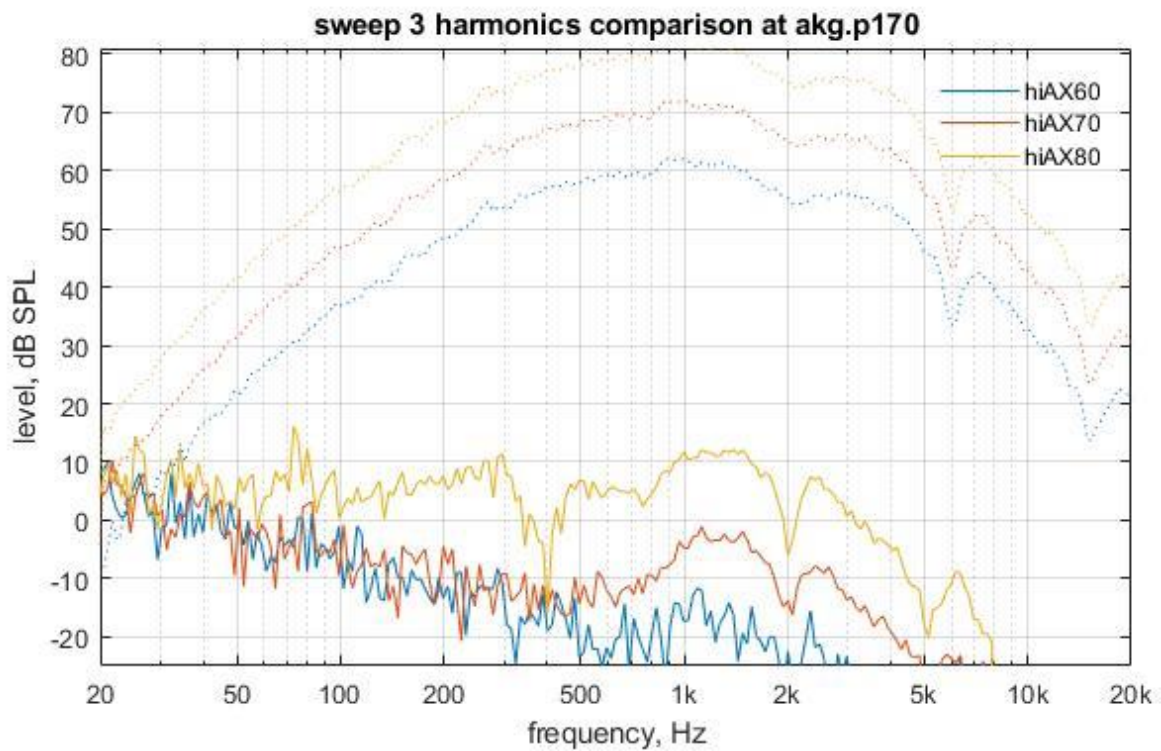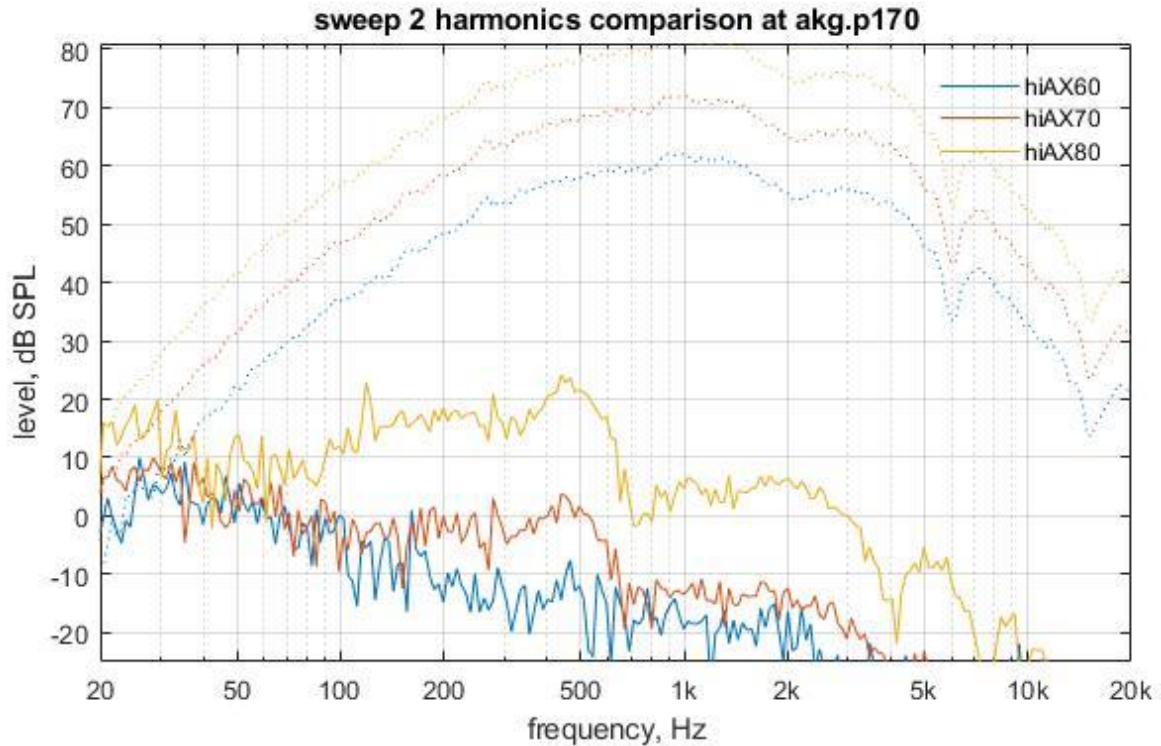
The 3rd harmonic is ~70dB below the output. It decays a bit faster than for PS130F but still very far from 30 dB per 10dB of input.

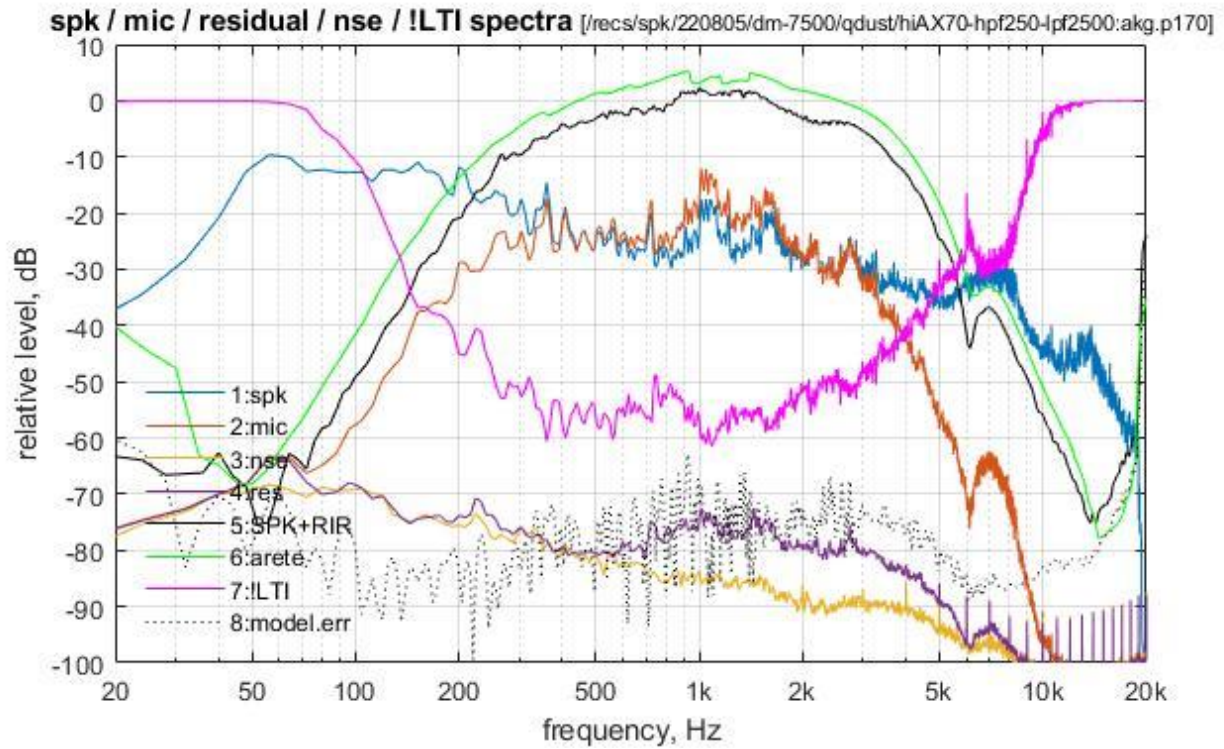### 3.2.2   LTI distortions on brown noise

Already limited to 250-2500 Hz mid-range:
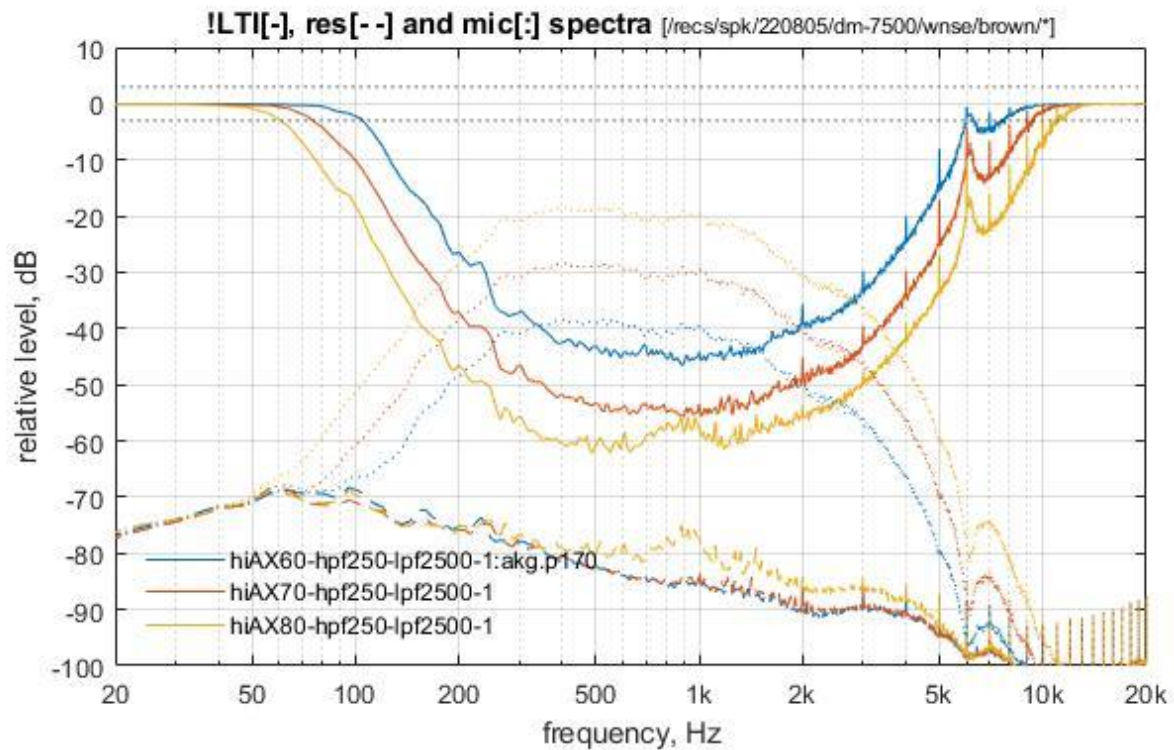
The distortions are mostly below the noise floor, except a small hump between 500 and 1000Hz on 80dBSPL. Not much sense to dig deeper on noise.

3.2.3   LTI distortions on ITU-T P.501 speech

residual spectrogram check-up [/recs/spk/220805/dm-7500/p501/hiAX70-hpf250-lpf2500:akg.p170]

!LTI[-], res[- -] and mic[:] spectra [/recs/spk/220805/dm-7500/p501/*]

hiAX60-hpf250-lpf2500-1:akg.p170
hiAX70-hpf250-lpf2500-1
hiAX80-hpf250-lpf2500-1

3.2.4 LTI distortions on a Rock music clip

spk / mic / residual / nse / !LTI spectra [/recs/spk/220805/dm-7500/qdust/hiAX70-hpf250-lpf2500:akg.p170]

Legend:
- 1:spk
- 2:mic
- 3:nse
- 4:res
- 5:SPK+RIR
- 6:arete
- 7:!LTI
- 8:model.err

residual spectrogram check-up [/recs/spk/220805/dm-7500/qdust/hiAX70-hpf250-lpf2500:akg.p170]

The peak-to average ratio is lower for Rock, thus the hump moved up on the RMS scale. The octave-wide excitation to check the distortion spread:
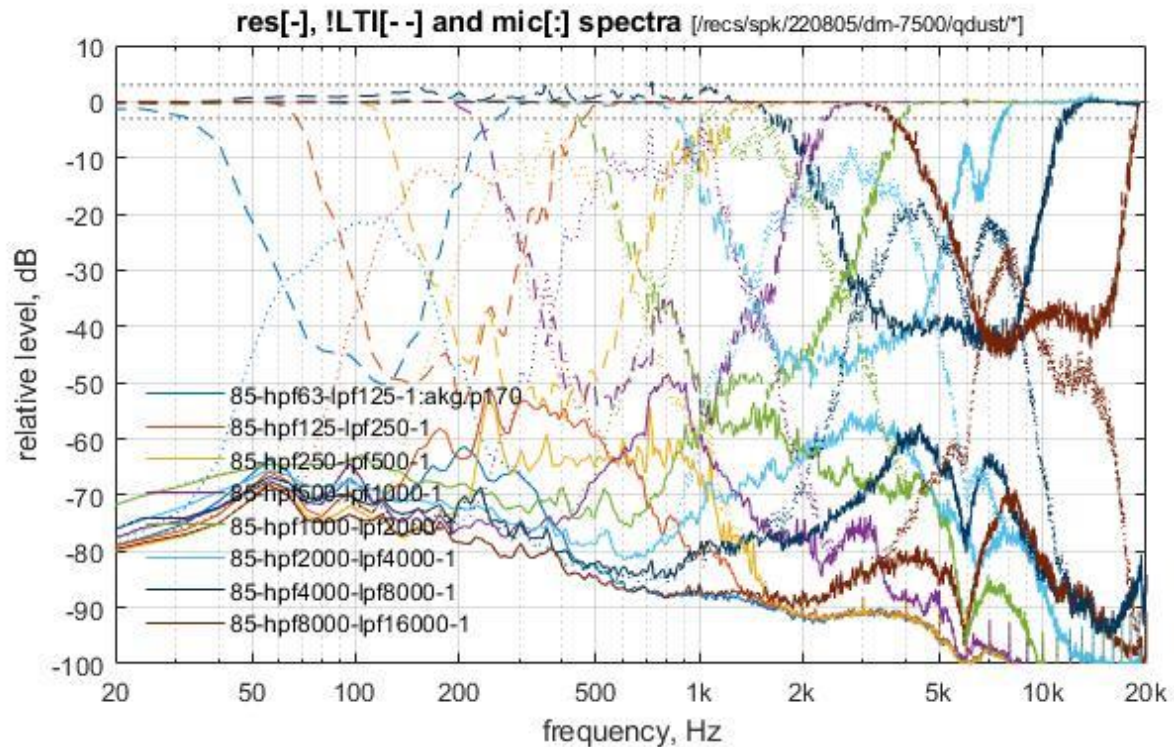
... which shows the 500...1000Hz hump. However, it disappears on current drive:

res[-], !LTI[- -] and mic[:] spectra [/recs/spk/220805/dm-7500/qdust/*]

Legend:
- 85-hpf63-lpf125-1-res-1:akg.p170
- 85-hpf125-lpf250-1-res-1
- 85-hpf250-lpf500-1-res-1
- 85-hpf500-lpf1000-1-res-1
- 85-hpf1000-lpf2000-1-res-1
- 85-hpf2000-lpf4000-1-res-1
- 85-hpf4000-lpf8000-1-res-1
- 85-hpf8000-lpf16000-1-res-1
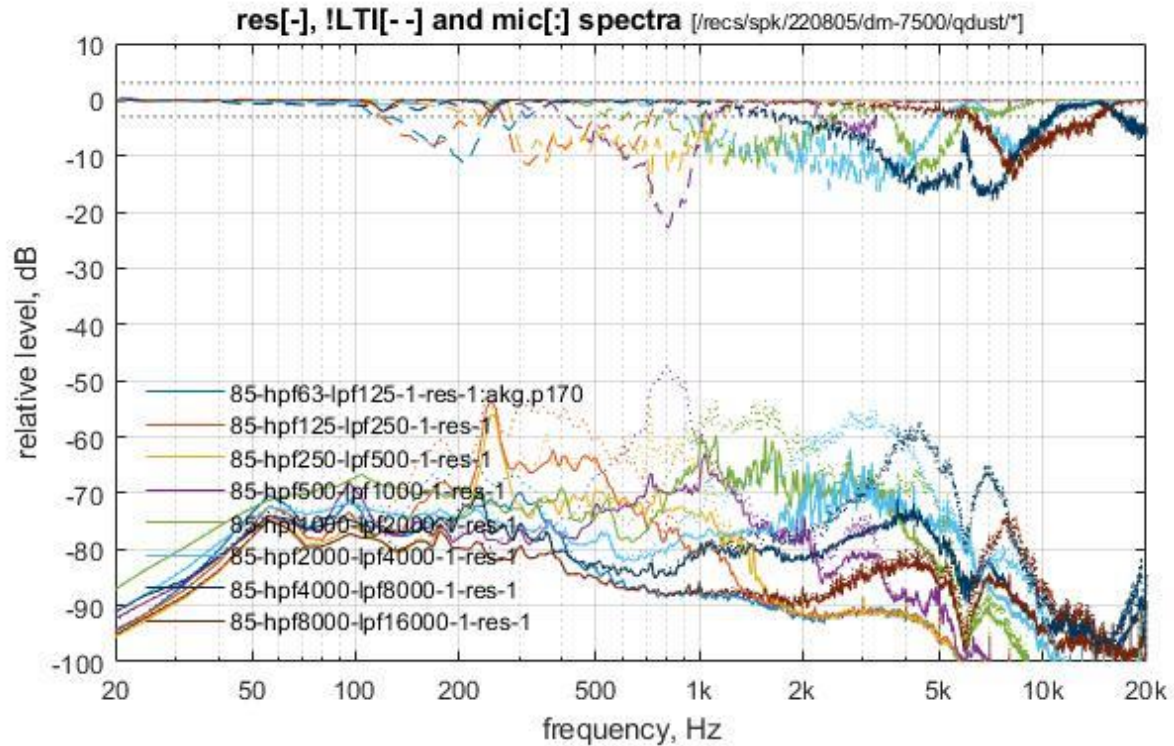
The DM-7500 distortions on voltage drive are lower but the benefits of current drive linearization are lower too.

3.2.5   LTI distortions on a Piano music clip

spk / mic / residual / nse / !LTI spectra [/recs/spk/220805/dm-7500/sbls/hiAX70-hpf250-lpf2500:akg.p170]

Legend:
- 1:spk
- 2:mic
- 3:nse
- 4:res
- 5:SPK+RIR
- 6:arete
- 7:!LTI
- 8:model.err
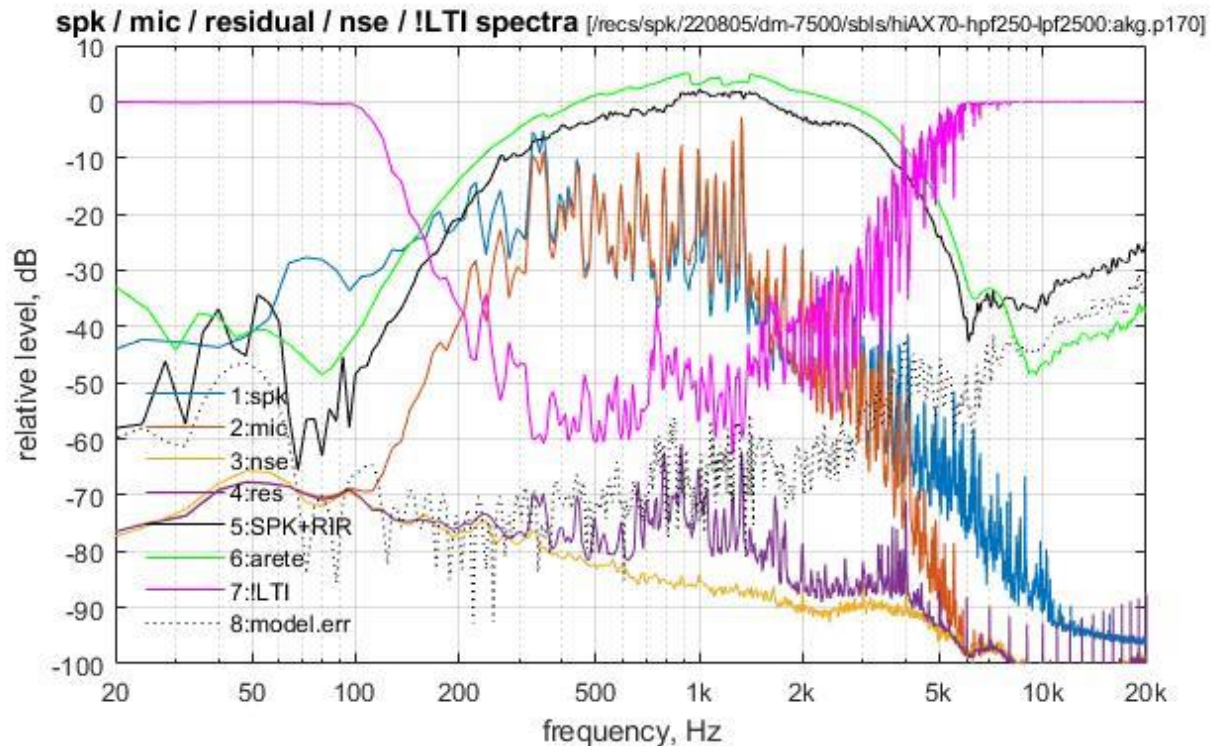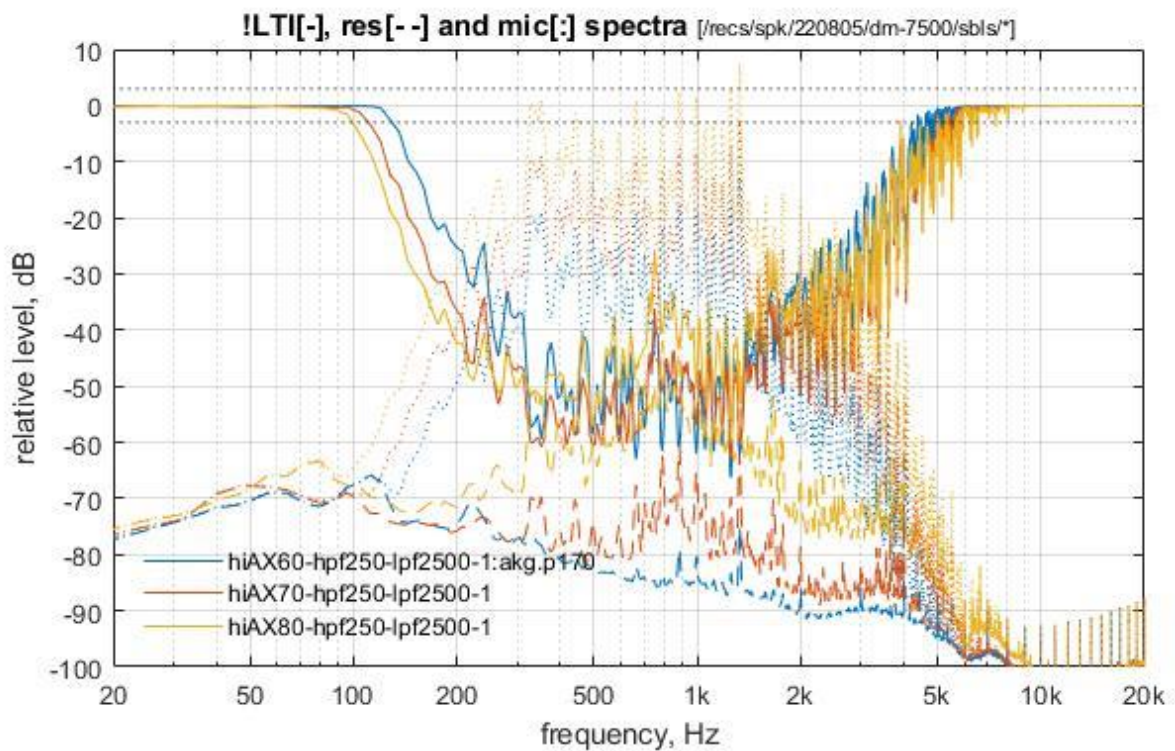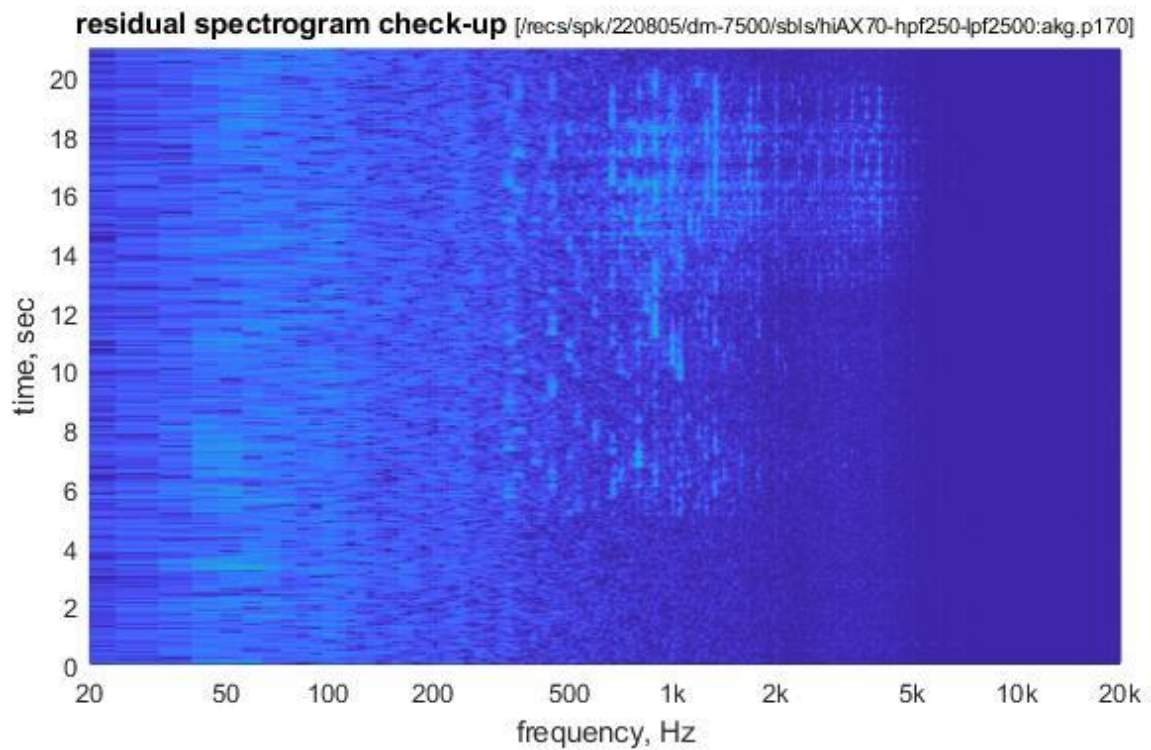
The distortions are lower for low and moderate loudness (≤70dBSPL). DM-7500 would be ok for desktop loudspeakers but not for studio monitors nor for floor standers.

I can not understand the reason for the hump in LTI distortions between 500 and 1000Hz on the levels above 60dBSPL.

3.3   SOUNDERLINK AMT-920 A.K.A. NEO8

A (copy of) flat panel AMT (Air Motion Transformer), 3.5"W * 8"H * 0.5"D, 90dBSPL/W/m. Re=5.9 Ohm, Le=0. Low frequencies can not be applied to such panels. All measurements have been done with 250Hz | 600Hz HPF.

3.3.1   Harmonic distortions
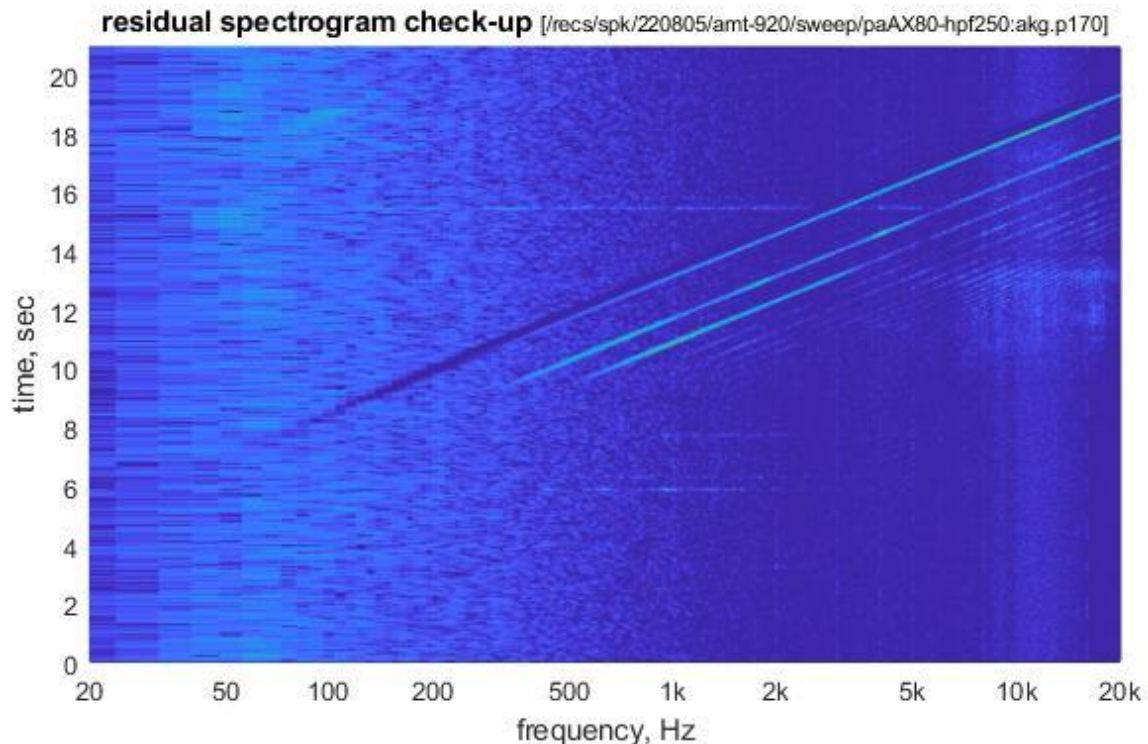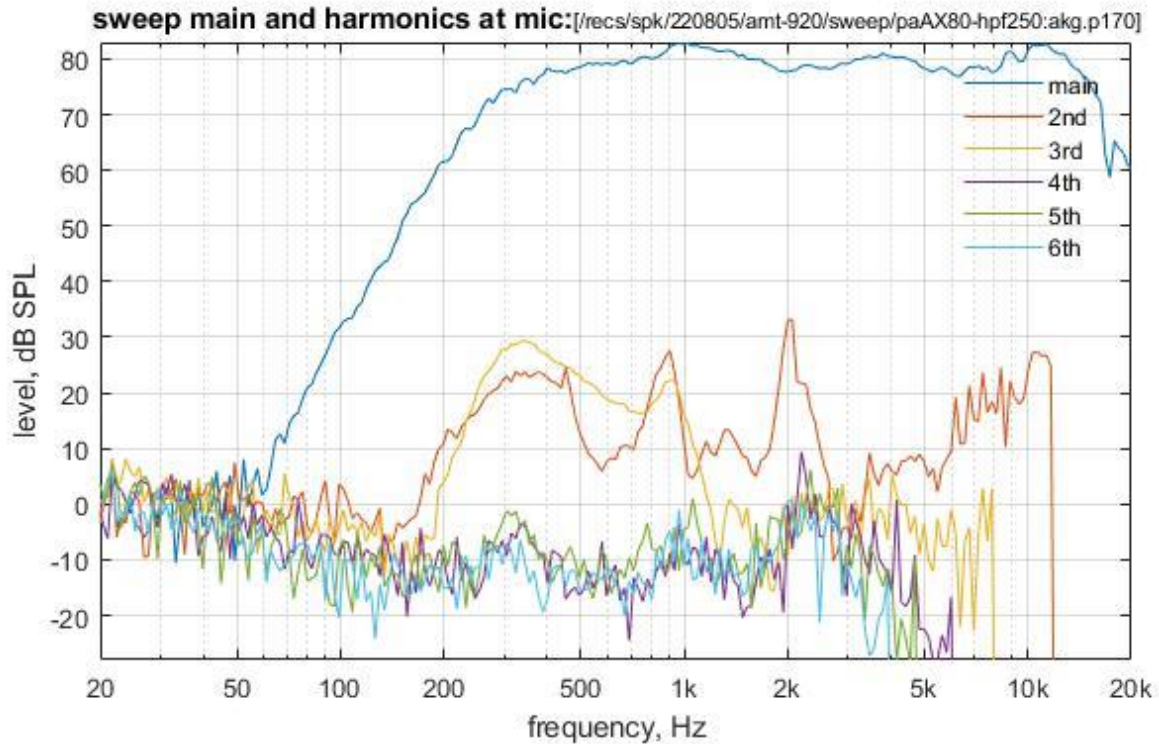
The long tail of distortions above 700Hz announces a discontinuity step in the function itself.

sweep 2 harmonics comparison at akg.p170

sweep 3 harmonics comparison at akg.p170
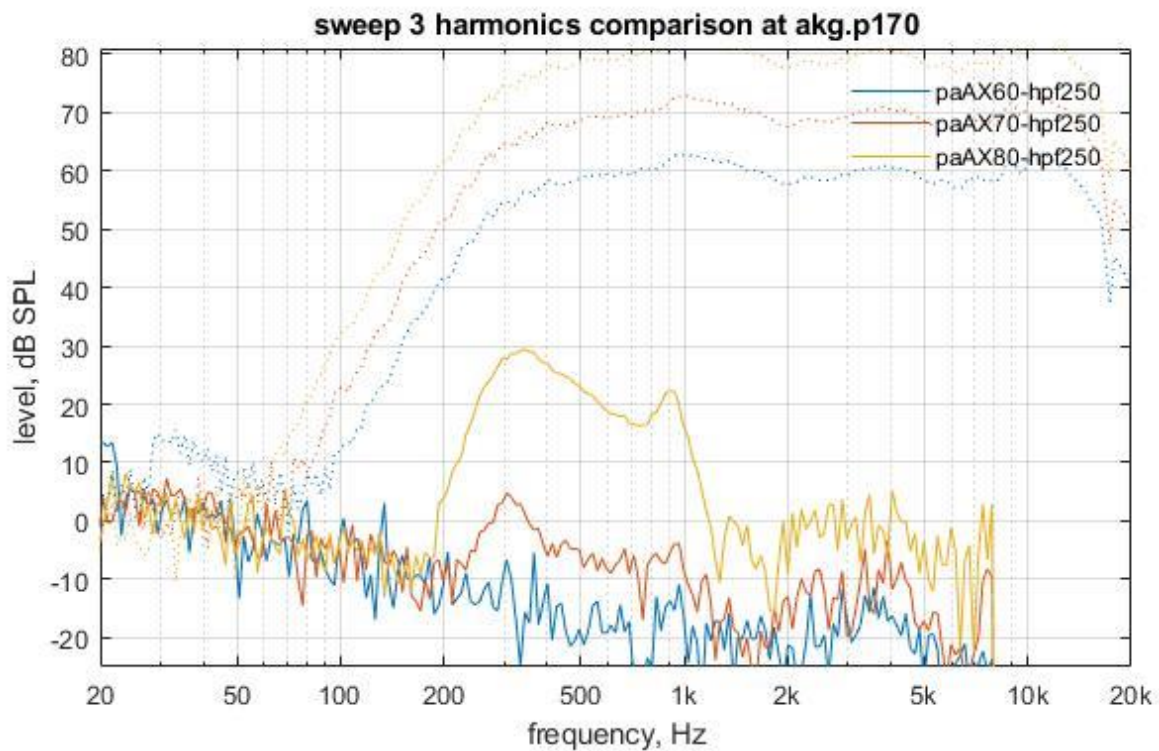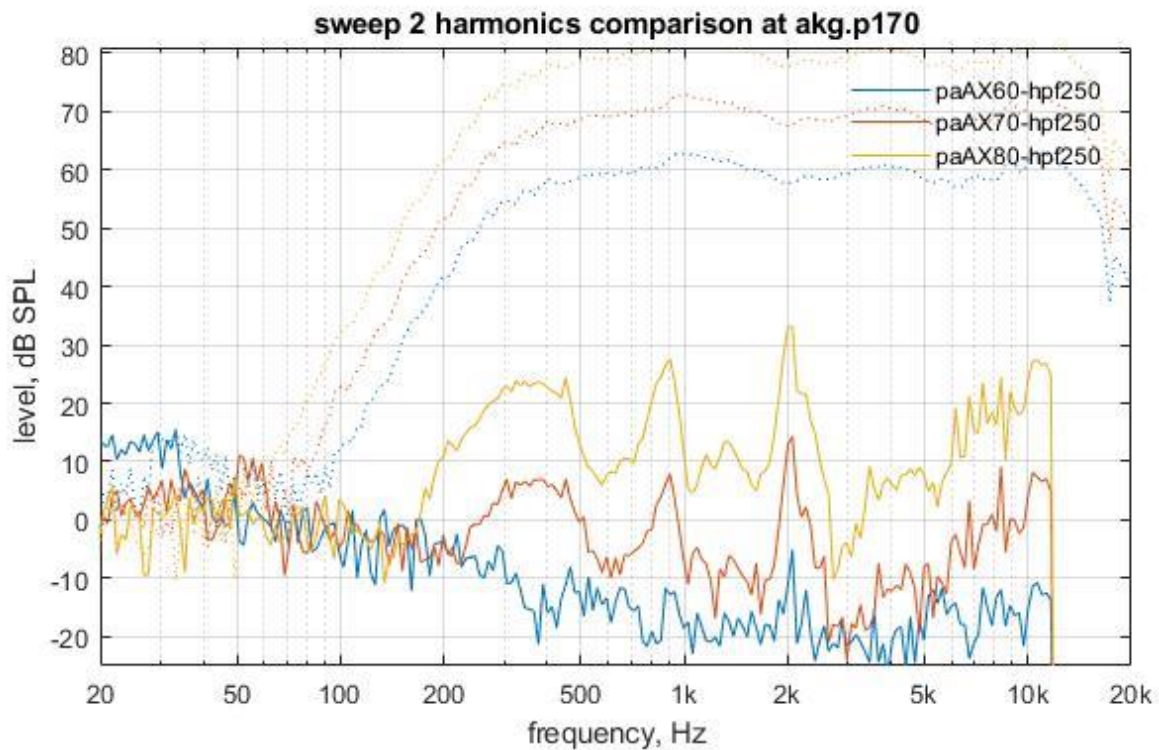
Here we see the "normal" 3rd harmonic, dropping 30dB per 10dB of output decrease but the distortions on 80dBSPL are rather high. All flat panel drivers behave this way[8].

---

[8] Shall we call them softspeakers?

3.3.2   LTI distortions on brown noise

The inclusion (or exclusion of low frequencies) affects the high frequency LTI distortions most. Here HPF = [50 250 500 600] Hz. The performance changes from mediocre to very good:
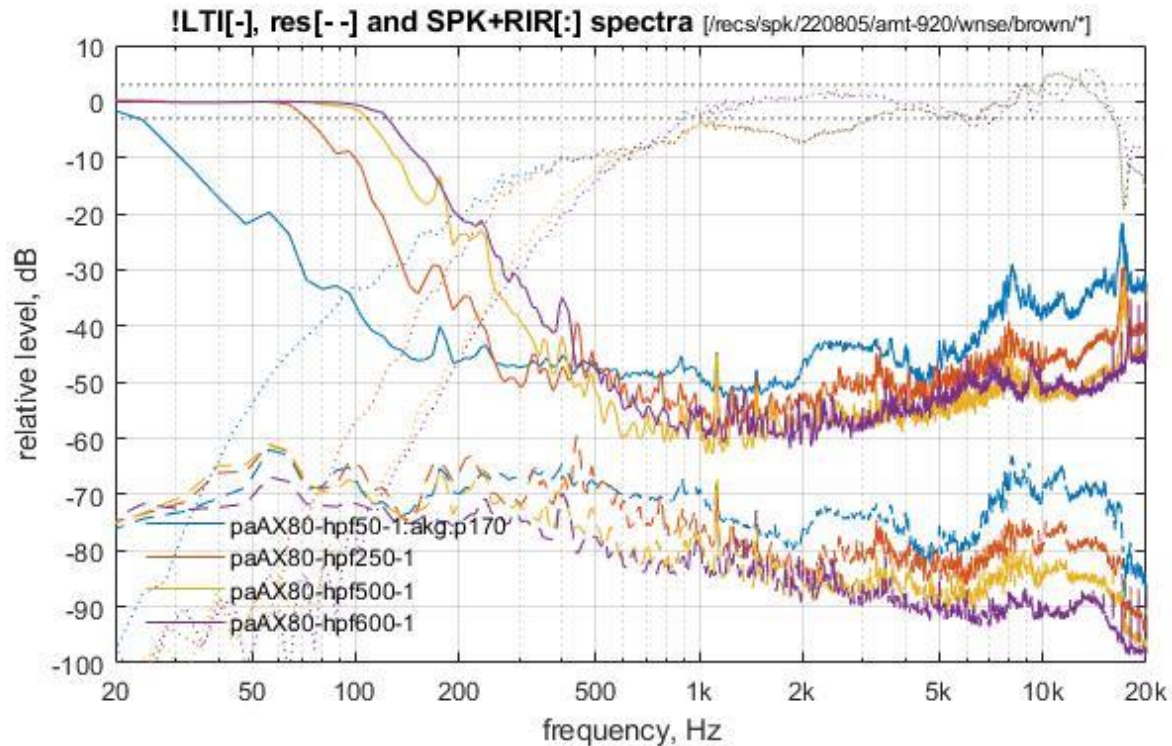
The distortions on 80dBSPL HPF brown noise are much lower than for "traditional" loudspeakers:
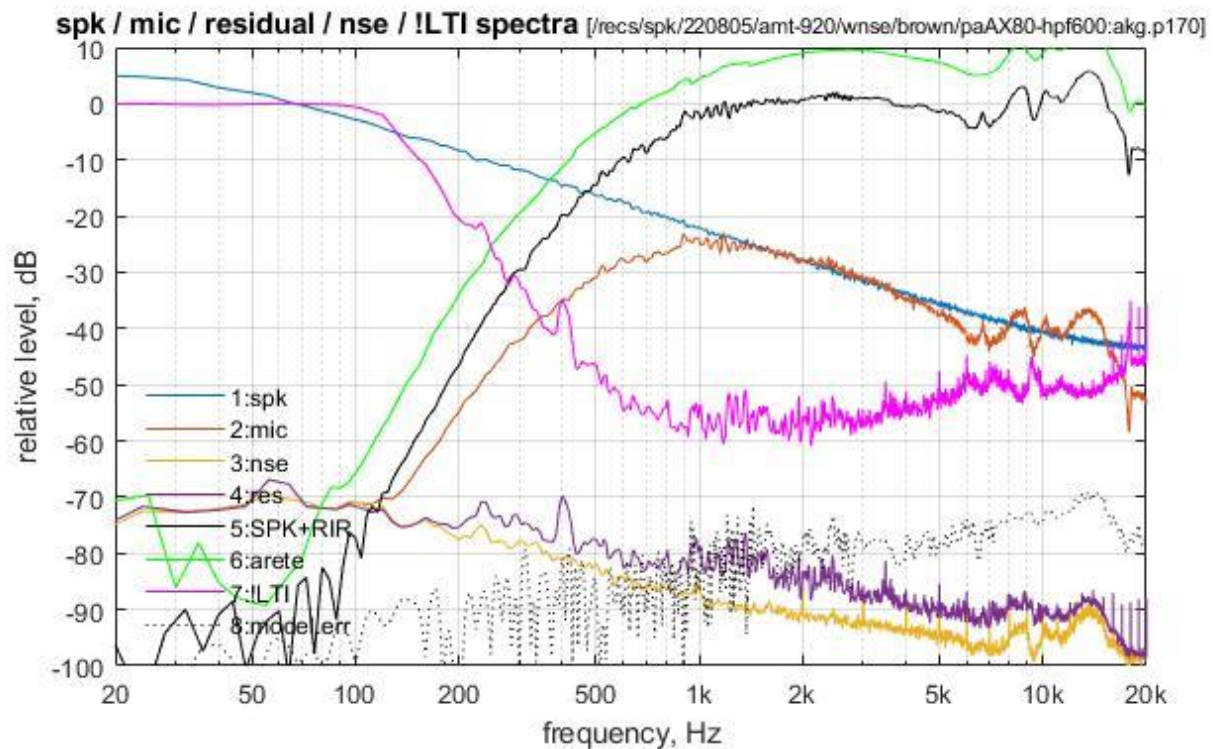
The distortions do not become catastrophic even for 90dBSPL brown noise if AMT-920 is used as mid-range [600...6000]Hz:
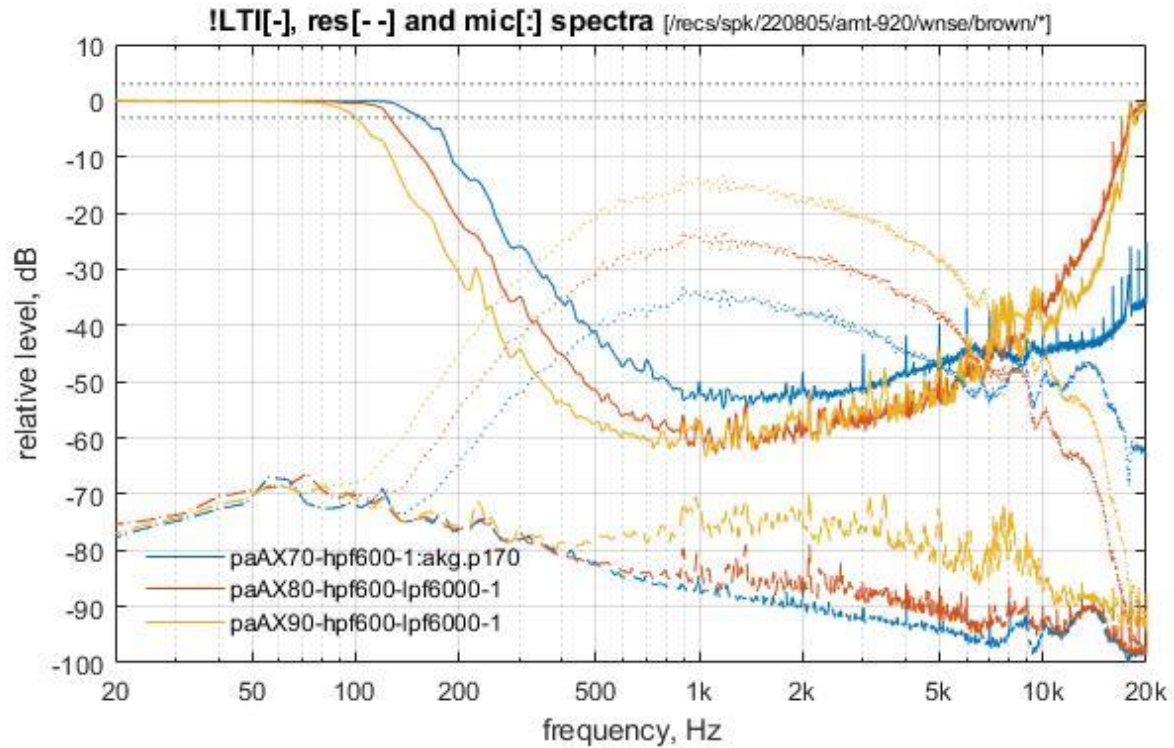
The panel can be used as a tweeter (especially in a horn) if 2500Hz HPF is applied:
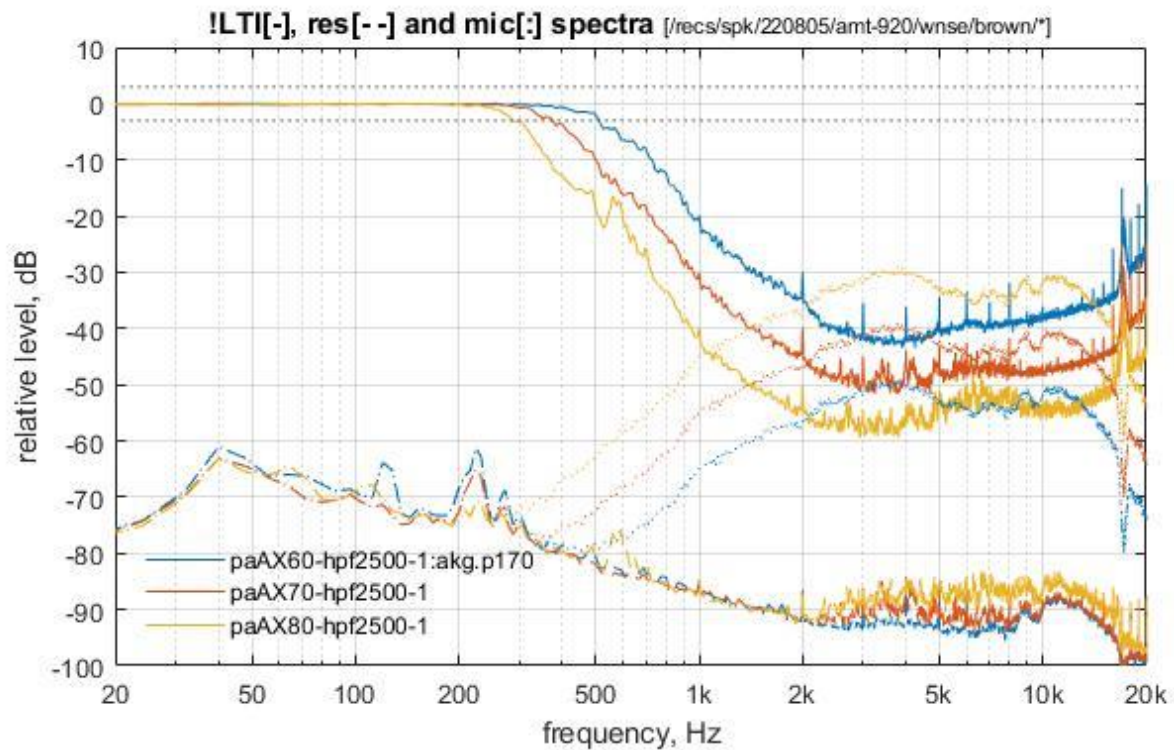
The AMT performance can not be improved by current drive.

3.3.3   LTI distortions on ITU-T P.501 speech

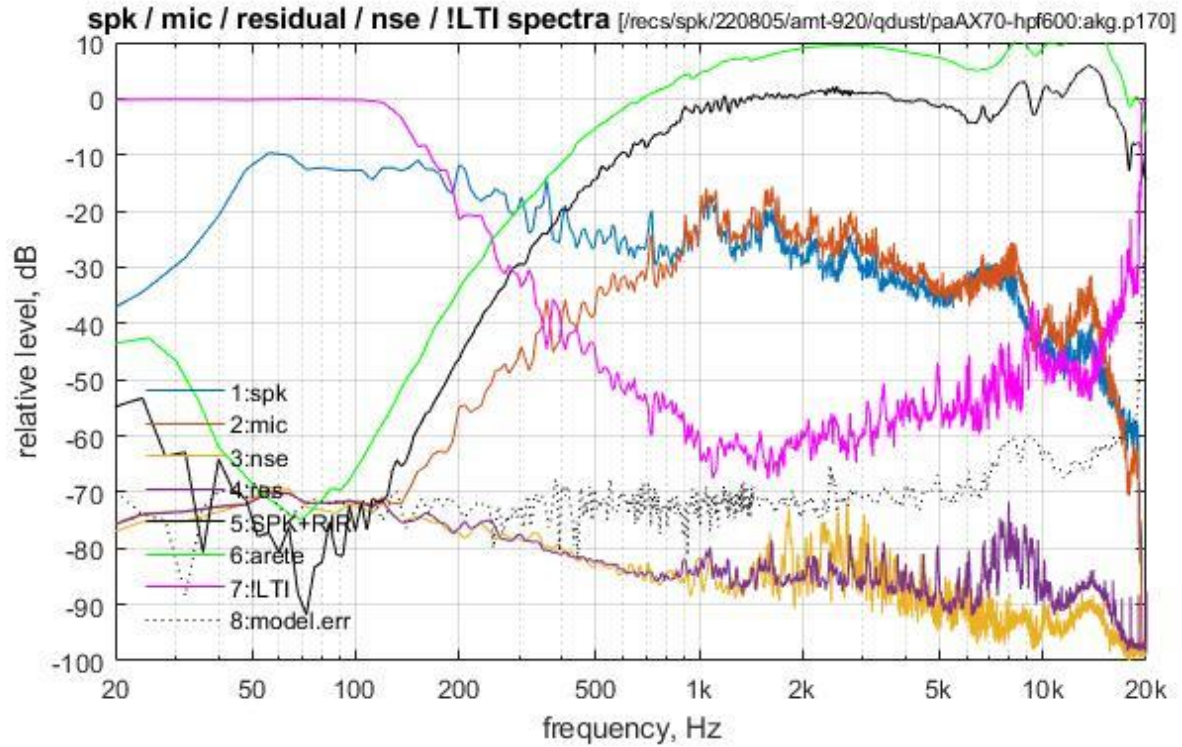The performance on non-stationary signals is less impressive:
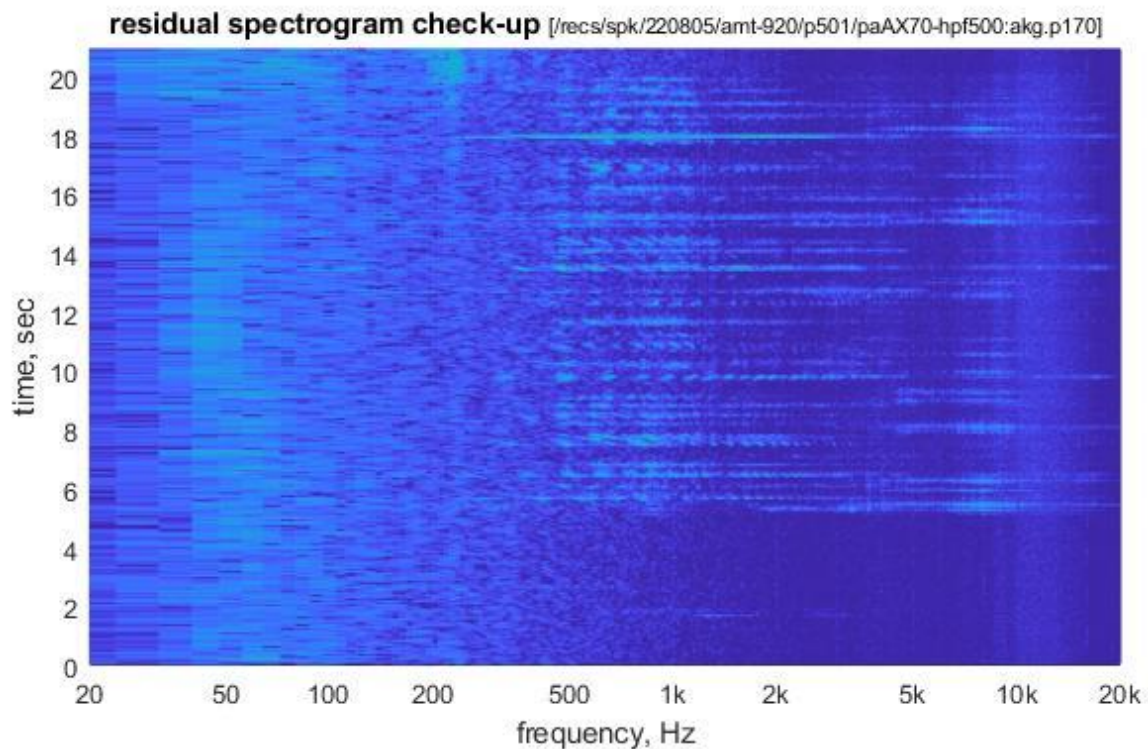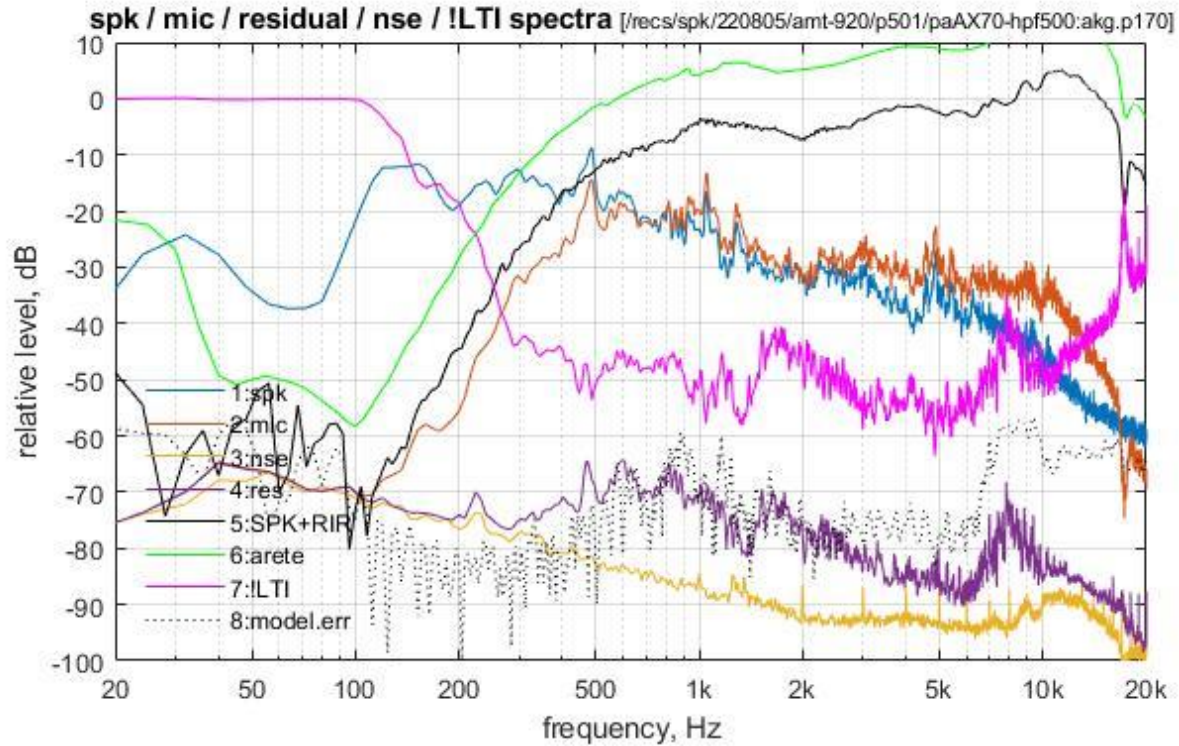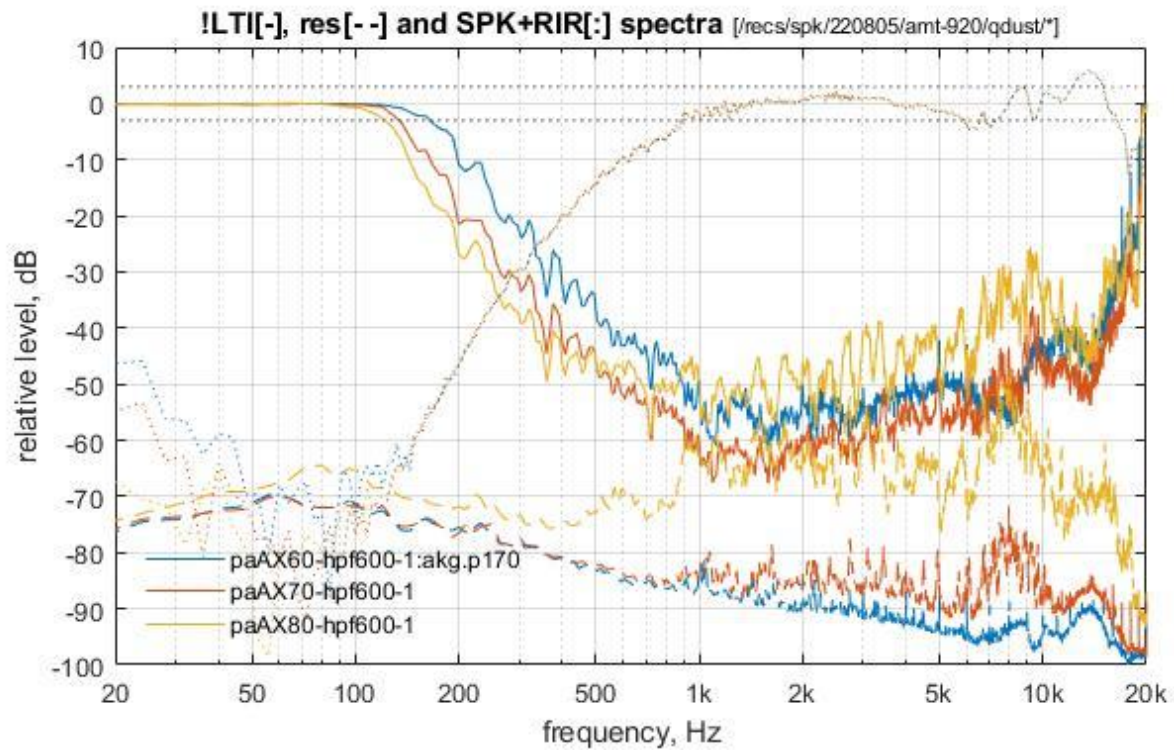

Especially if HPF is higher:
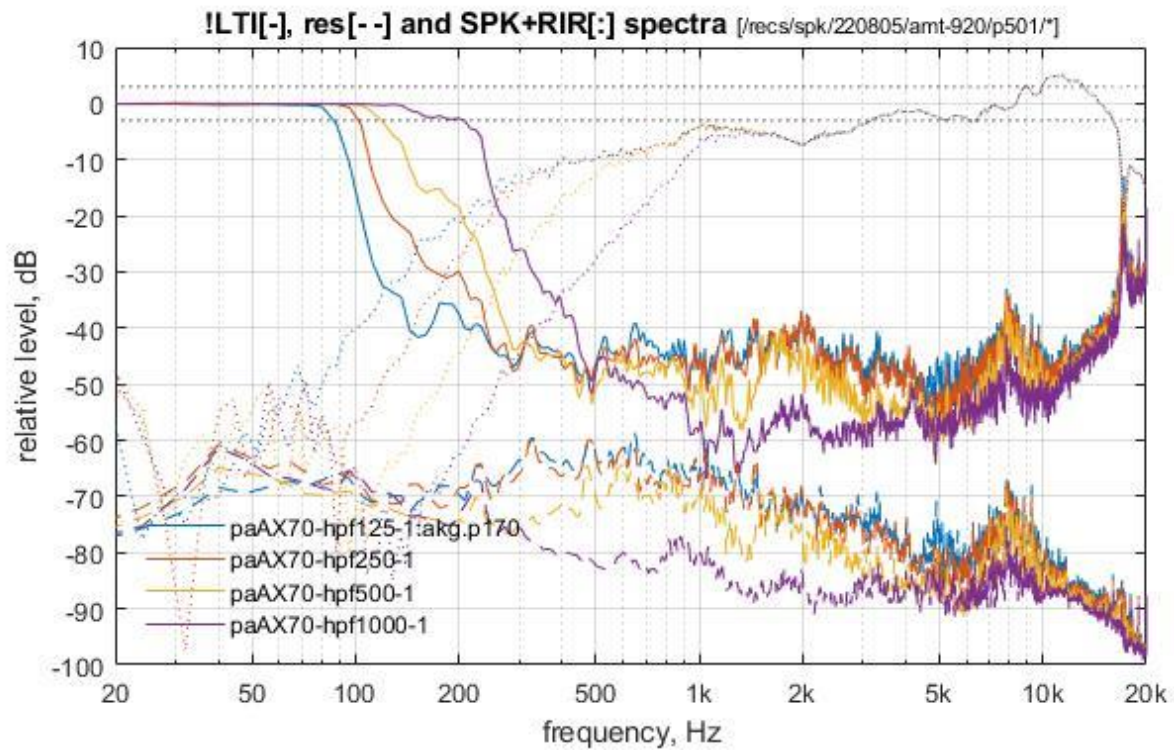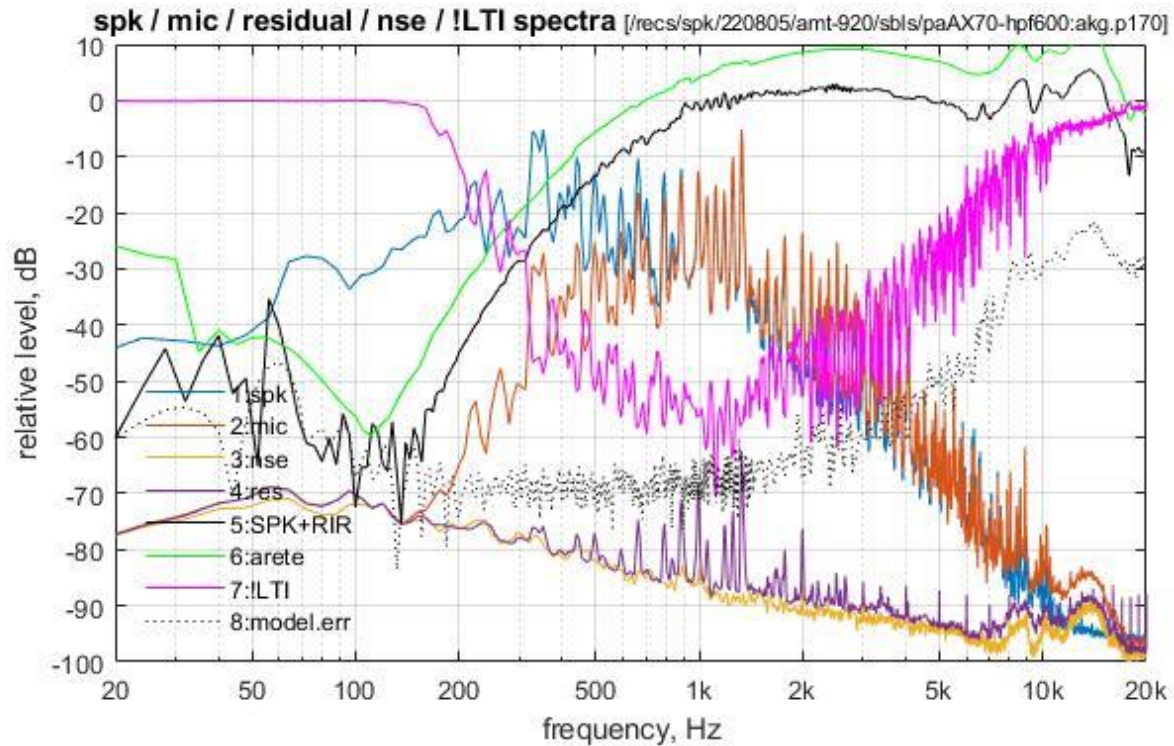
But the performance quickly degrades on louder output:

3.3.4 LTI distortions on a Rock music clip

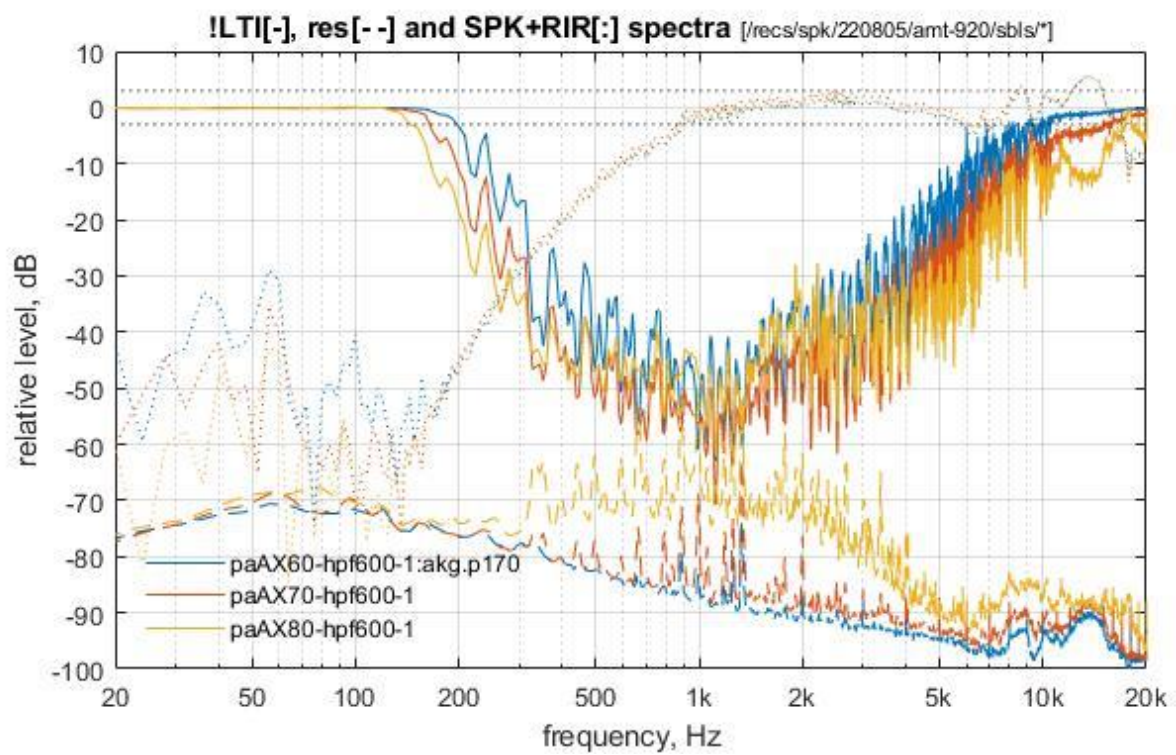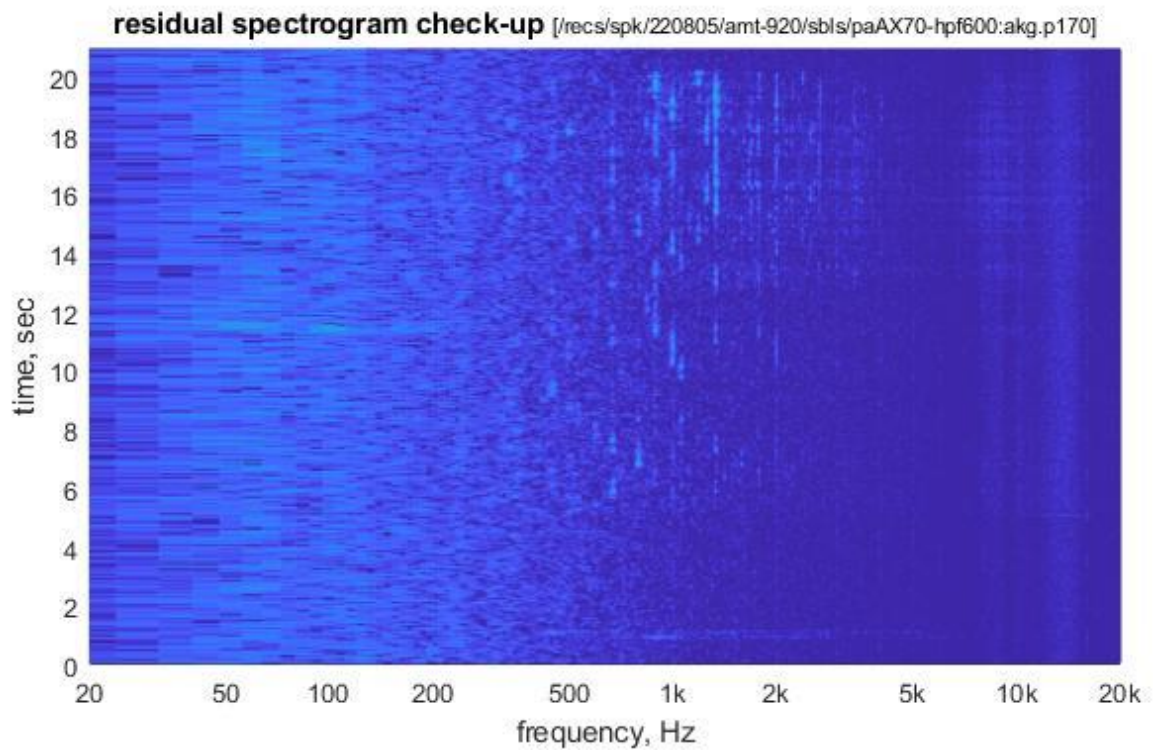spk / mic / residual / nse / !LTI spectra [/recs/spk/220805/amt-920/qdust/paAX70-hpf600:akg.p170]

Legend:
- 1:spk
- 2:mic
- 3:nse
- 4:res
- 5:SPK+RIR
- 6:arete
- 7:!LTI
- 8:model.err

residual spectrogram check-up [/recs/spk/220805/amt-920/qdust/paAX70-hpf600:akg.p170]
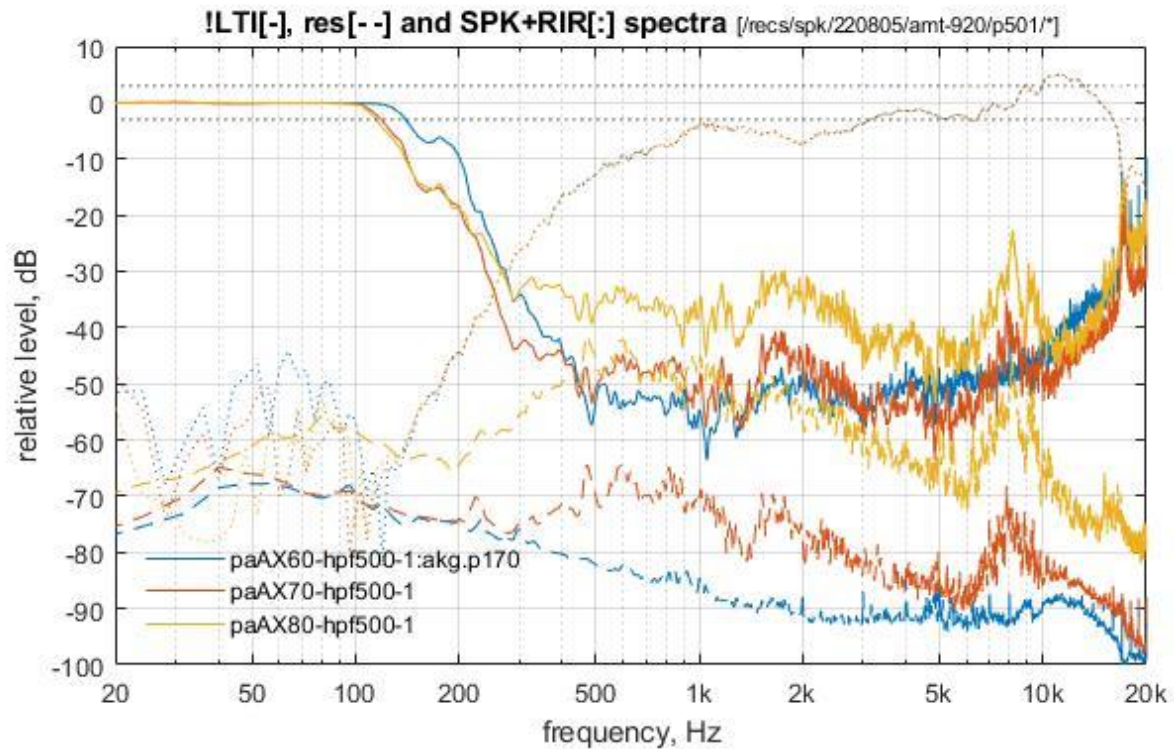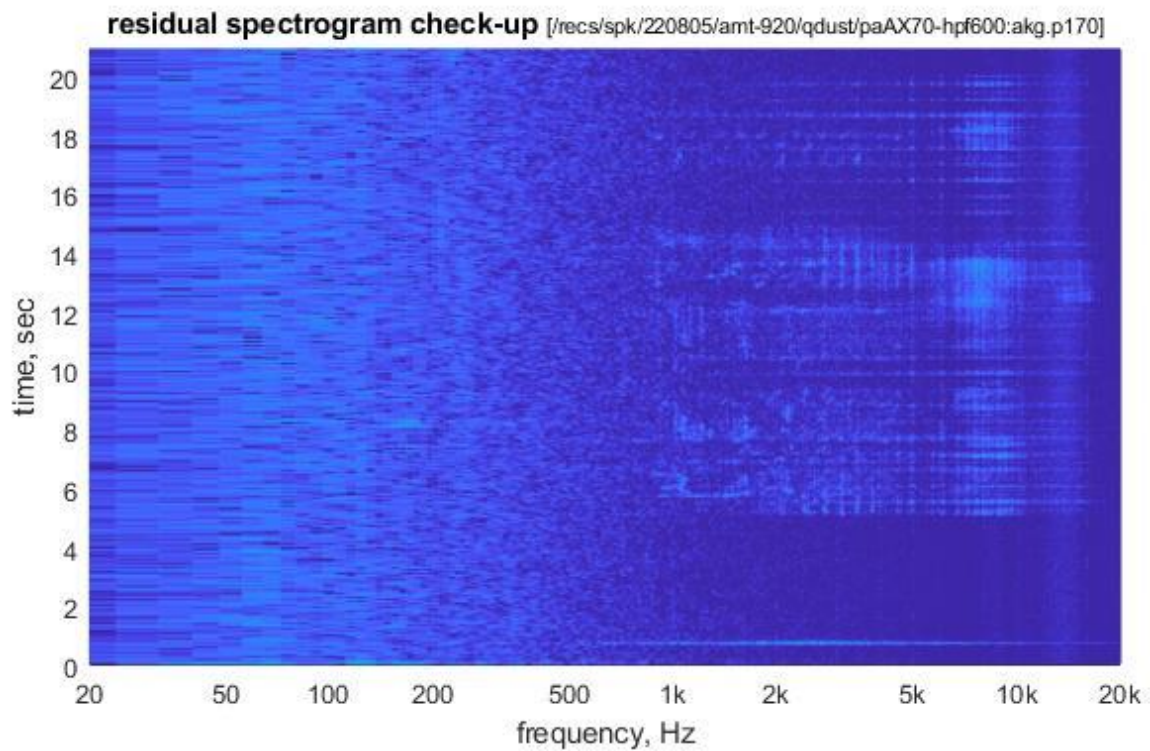
All is well till it breaks which it does:
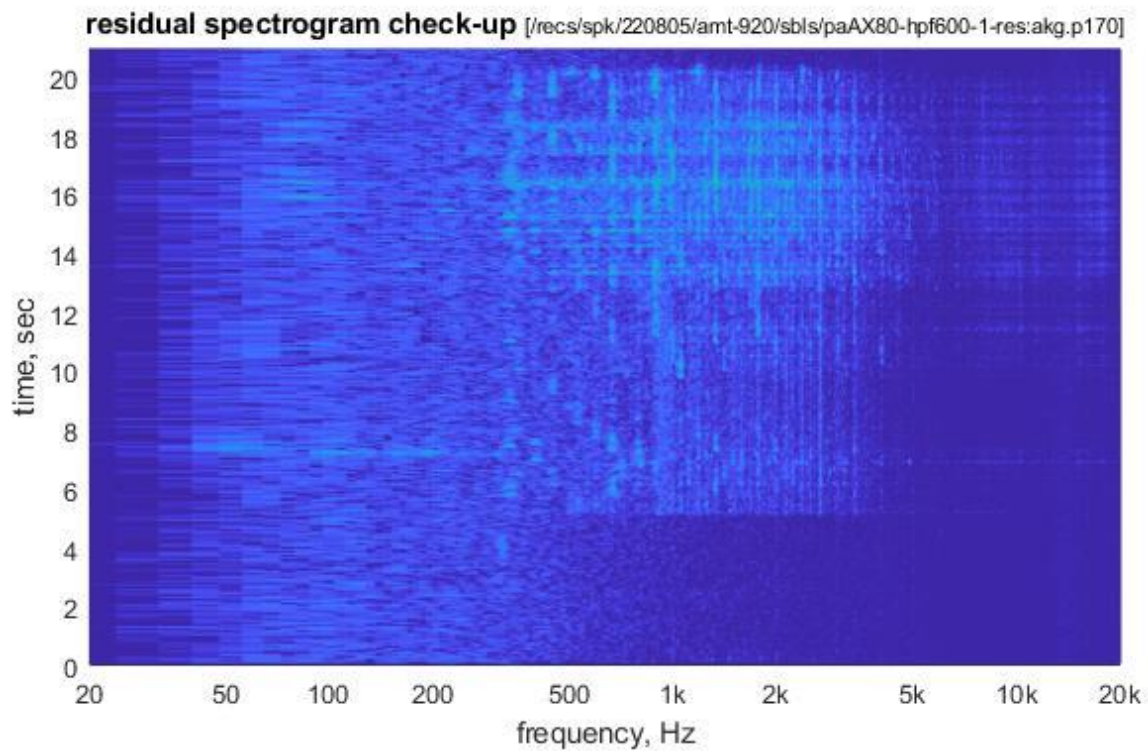
3.3.5   LTI distortions on a Piano music clip

residual spectrogram check-up [/recs/spk/220805/amt-920/sbls/paAX70-hpf600:akg.p170]

!LTI[-], res[- -] and SPK+RIR[:] spectra [/recs/spk/220805/amt-920/sbls/*]

paAX60-hpf600-1:akg.p170
paAX70-hpf600-1
paAX80-hpf600-1

However, unsimilar to the stationary noise, the 80dBSPL Piano LTI distortions…

residual spectrogram check-up [/recs/spk/220805/amt-920/sbls/paAX80-hpf600:akg.p170]
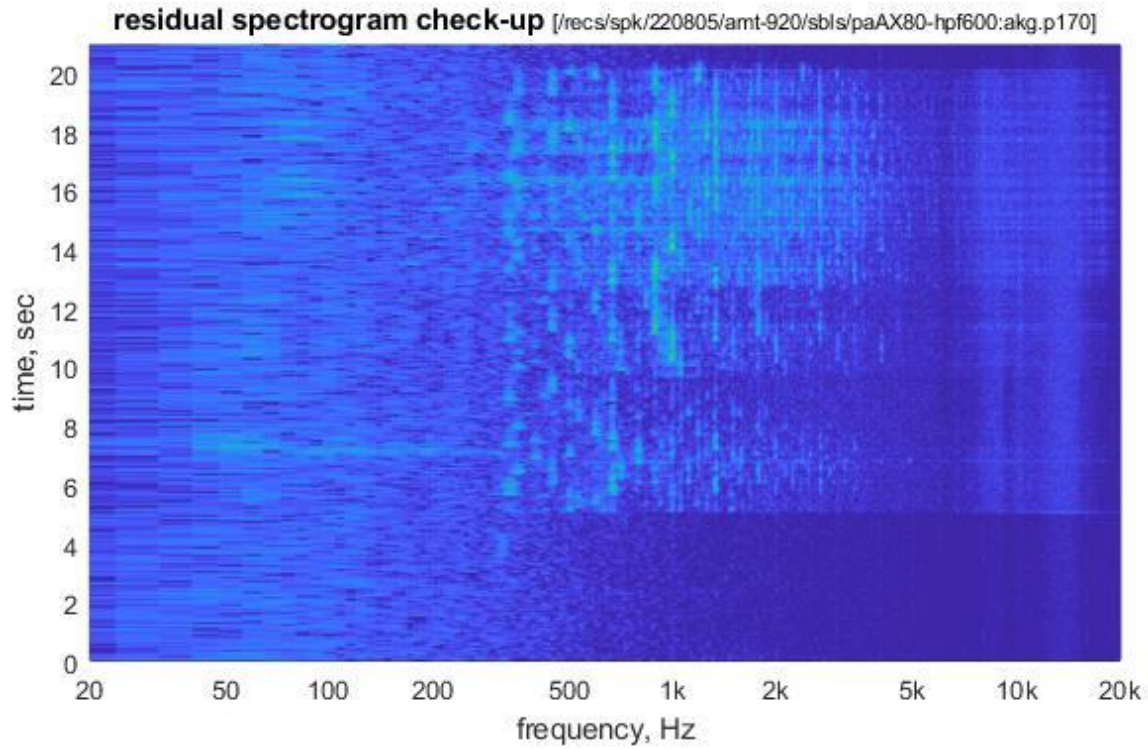
…can be significantly reduced by the current drive linearization:

residual spectrogram check-up [/recs/spk/220805/amt-920/sbls/paAX80-hpf600-1-res:akg.p170]

All peaks are gone, AMT-920 becomes a tamed loudspeaker.

3.4   Vifa (Tymphany) NE95W-04

3″ fullrange driver with aluminium cone, NdFeB magnet and round die cast alu frame. Nominal impedance [Zn]: 4 Ω; Resonance frequency [fs]: 103 Hz; Sensitivity [2.83 V/1m]: 86.1 dB; Equivalent volume [Vas]: 1.11 l; Total Q factor [Qts]: 0.72. Discontinued.
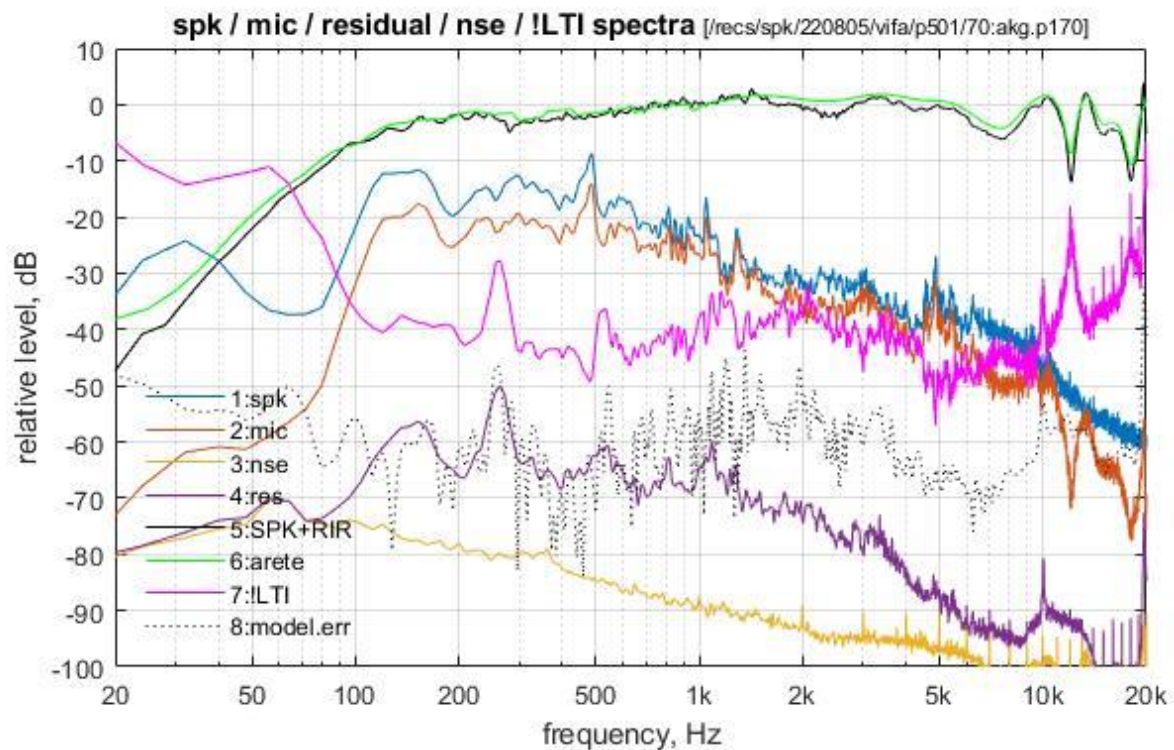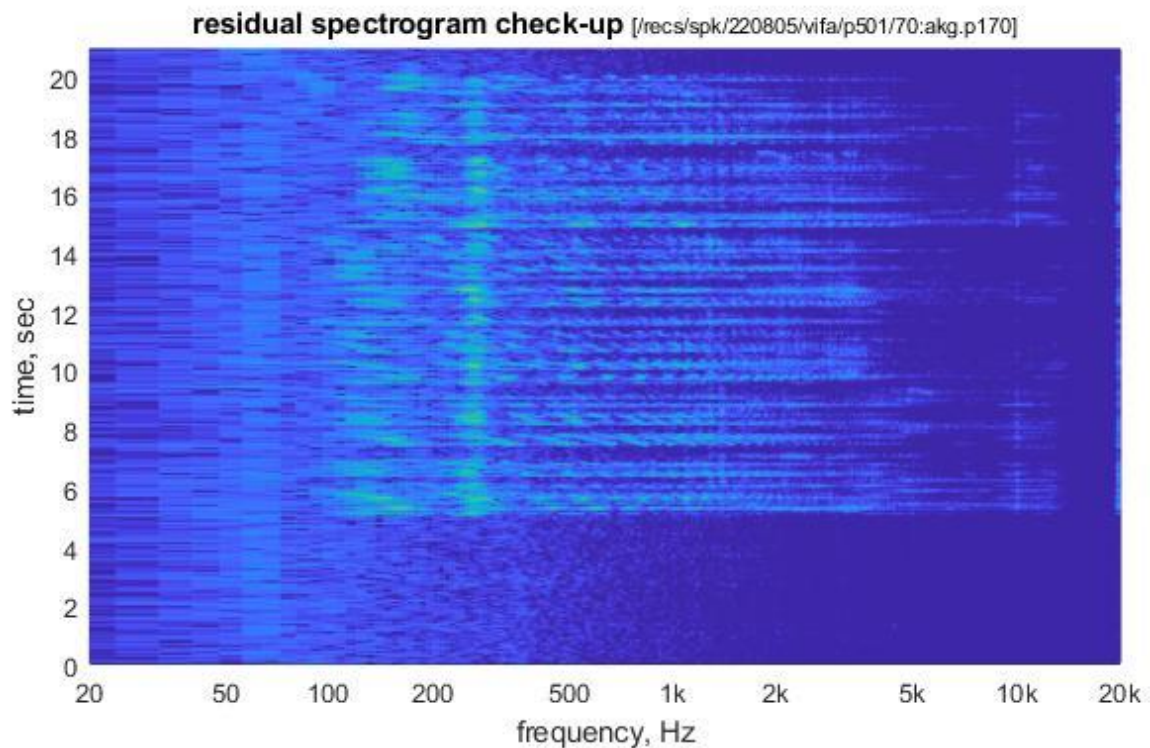
3.4.1   Harmonic distortions
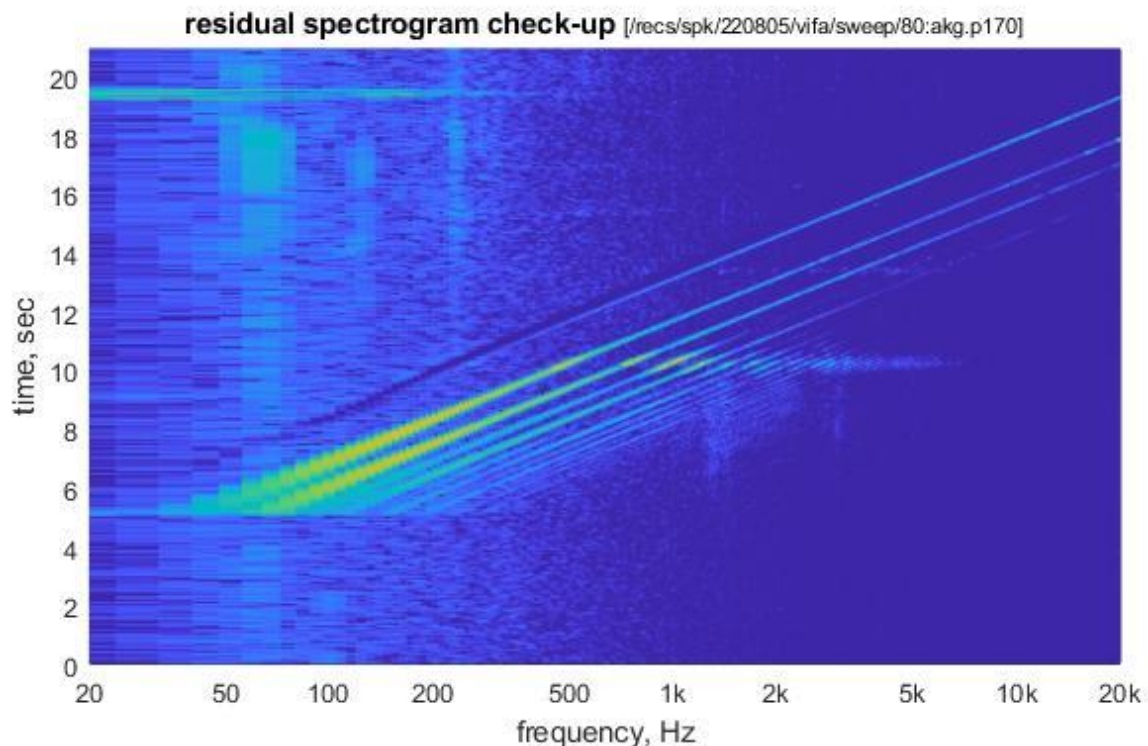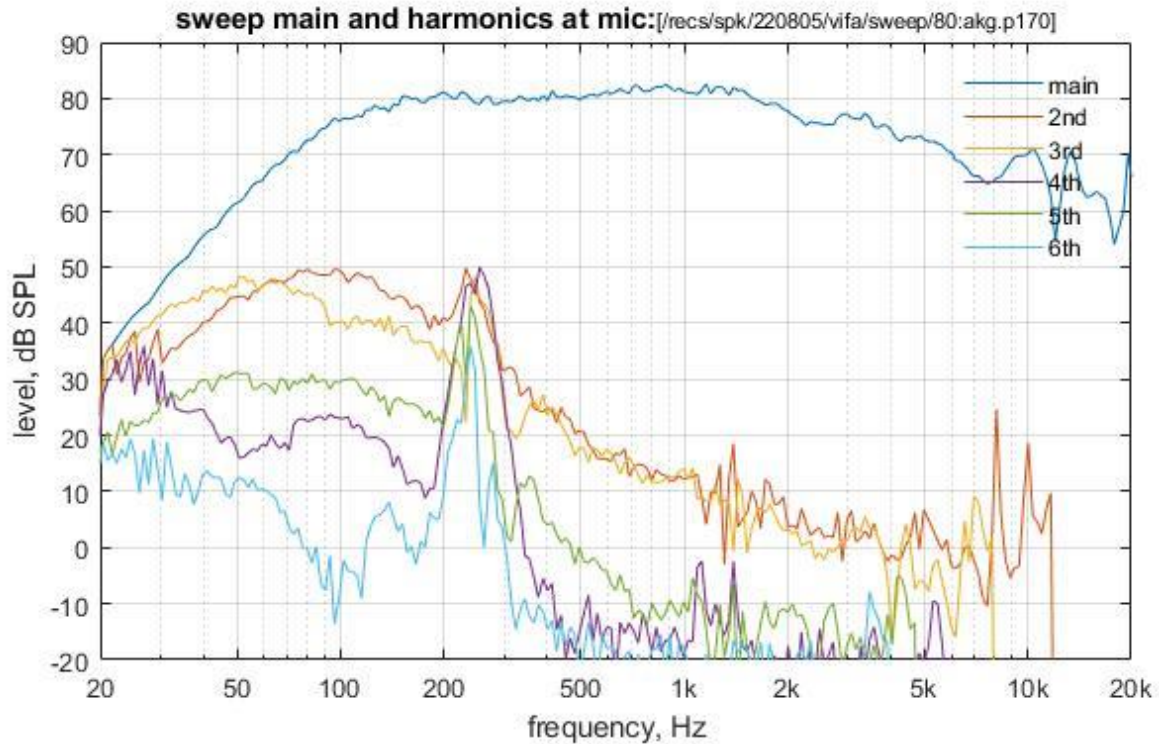
The ~250Hz resonance remains at lower dBSPL with slightly shifted up frequency:
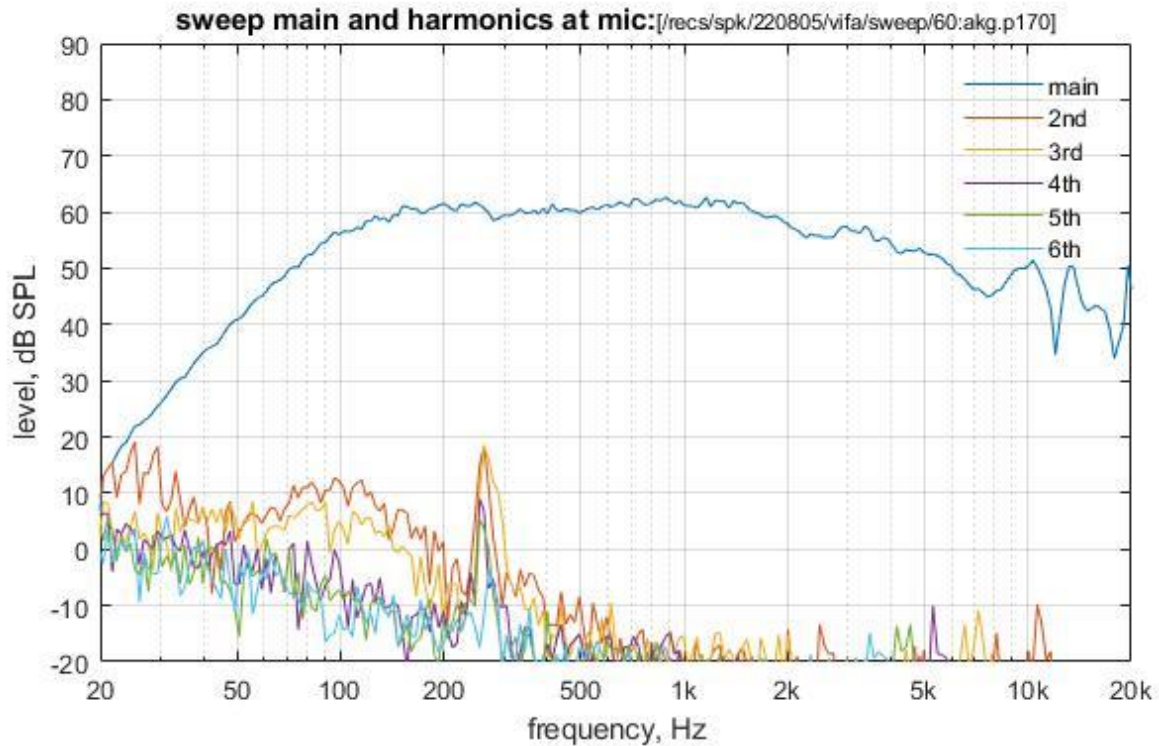
3.4.2   LTI distortions on ITU-T P.501 speech

This is a typical example of a driver with a significant flaw which is totally unpredictable. It does not look as trivial to locate and identify such anomaly in real-time during a conference call.

The spectrogram above shows how difficult is the interpretation of harmonic distortions. Looking at the figures in the previous chapter, we could assume that there will be distortions on resonance harmonics: 500Hz, 750Hz, 1kHz, 1250Hz etc (N*250, for n> 1). But we see only a bump of distortions on the same 250Hz.

The harmonic distortion measurements will be often omitted for next devices.

3.5   OTHER MID-RANGE LOUDSPEAKERS

- adam: X-ART woofer from ADAM F5, "A groundbreaking technological innovation with easily discernible superior performance and no technical flaws, the X-ART driver is clearly unequalled in the annals of audio history… These monitors are a quantum leap forward in musical accuracy, closely approaching the ideal loudspeaker." - X-ART Driver - ADAM Audio (adam-audio.com)
- inf4032: Infinity Reference 4032. "The audiophile-grade tweeters used in Infinity's new Reference speakers deliver smooth, non-fatiguing sound even at high output levels… Plus One™ cones provide … higher sensitivity, increased low-frequency output and accurate music reproduction." Reference 4032cfx | 4" (100mm) coaxial car speaker, 105W (infinityspeakers.com)
- lazo: LABO LB-PS1401D – amazing quality for the price CAD$24/pair on amazon.ca
- mjd4: 4Inch Portable Column Spund Speaker Subwoofer 6Ohm 40W, marked MJD4-6Ω-40 on the back, "Frequency response: 50HZ-20KHZ; Sensitivity: 87 (dB / W); S / N ratio: 83 (dB); Harmonic distortion: 0.1 (TMD%)
- sb12cacs25-4: 4″ SB12CACS25-4 / Ceramic - Sbacoustics
- gw-205-w: Goldwood Sound GW-205/4S Shielded 5.25" Woofer; 70 Watts RMS and 130 Watts Max; 62 - 12,000Hz Frequency Response; 87 dB SPL; 4 Ohm Impedance - Goldwood Sound GW-205/4S Shielded 5.25" Woofer 130 Watt 4ohm Replacement Speaker - Goldwood.com
- aiyima-flat: 4 Inch Audio Portable Subwoofer Speakers 6Ohm 10W; sold by Aiyima; the cheapest driver of the mix (CAD$10 per spk including delivery)
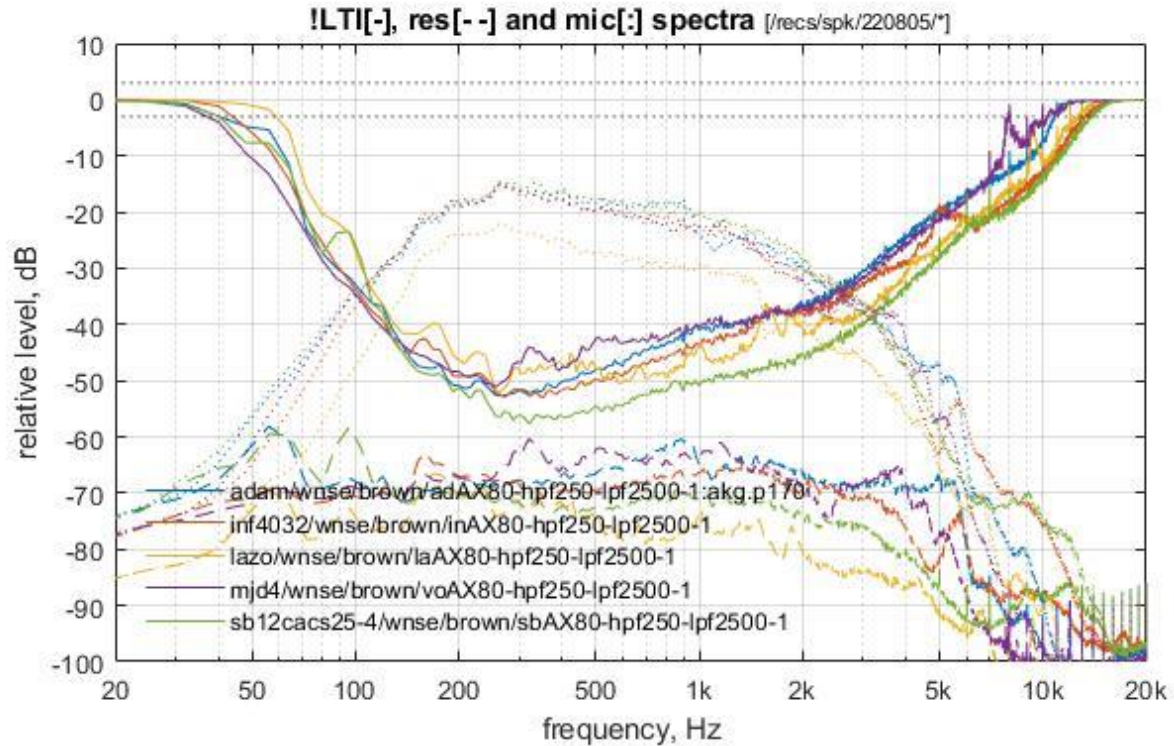- jmlab: JmLab Spectral 908.1 (ca. 1994) as the reference.
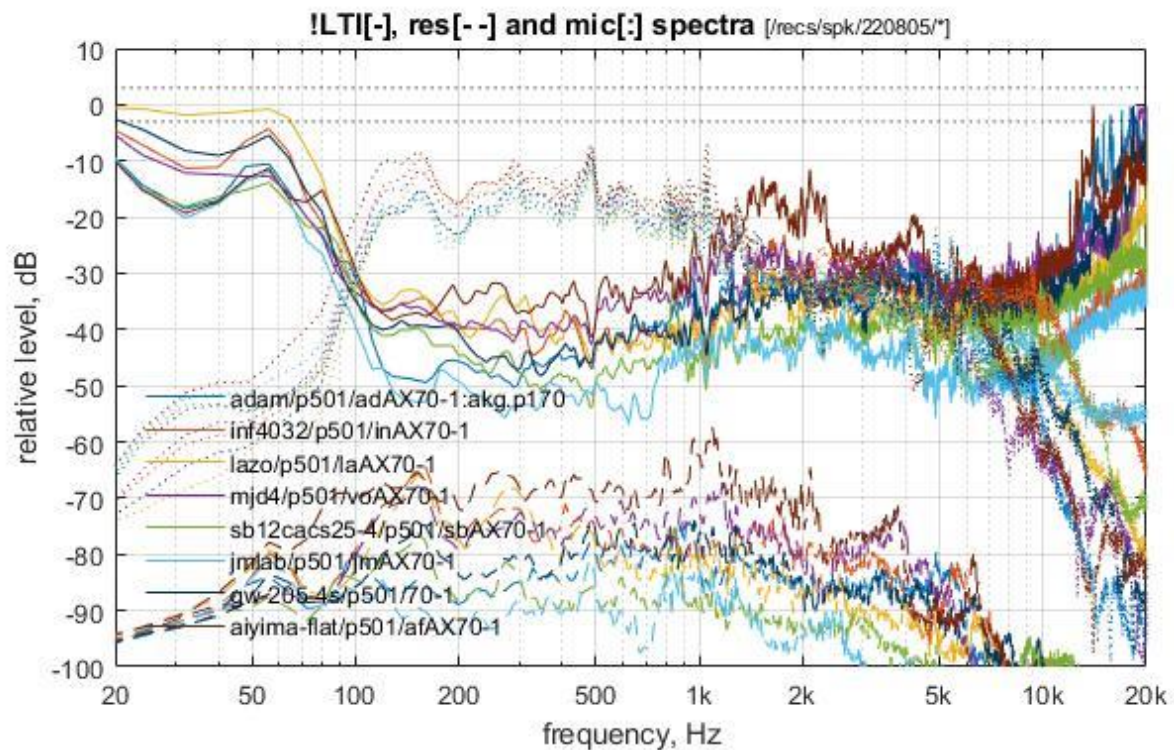
3.5.1   LTI distortions on brown noise
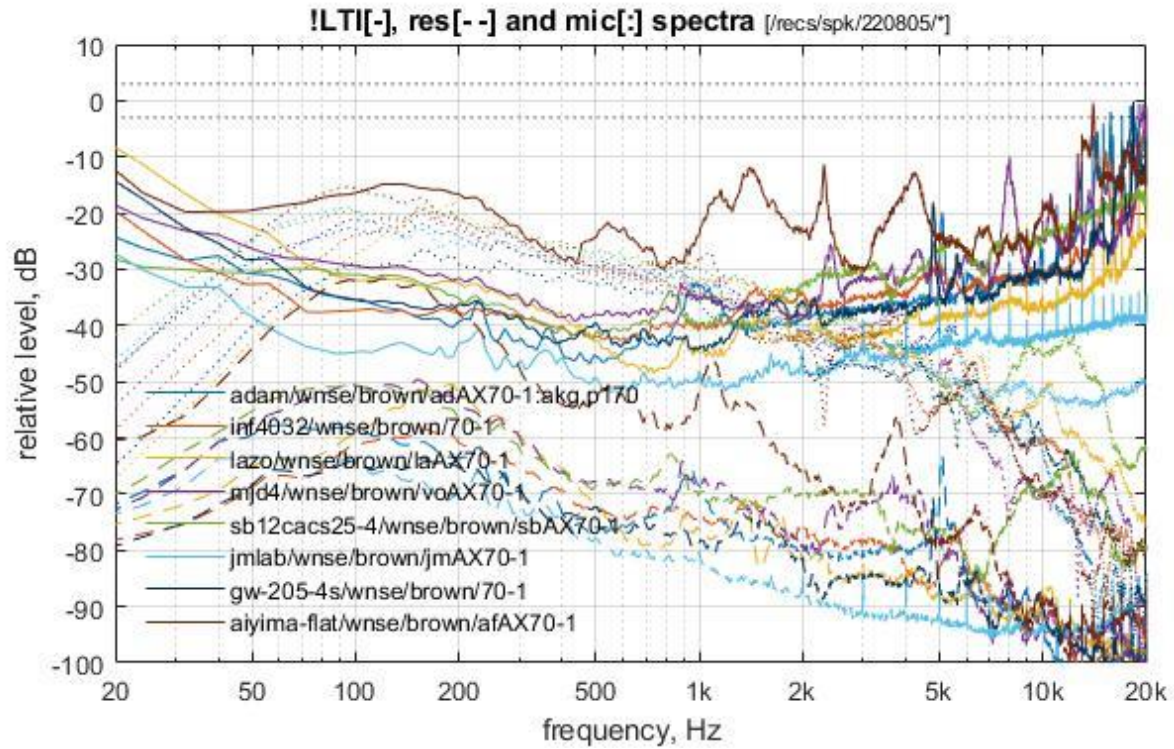
At 70dBSPL. Note that JmLab's distortions are below noise level. Of course, it is superior in all respects, and the louder is the output (80dBSPL below), the more superiority it displays.
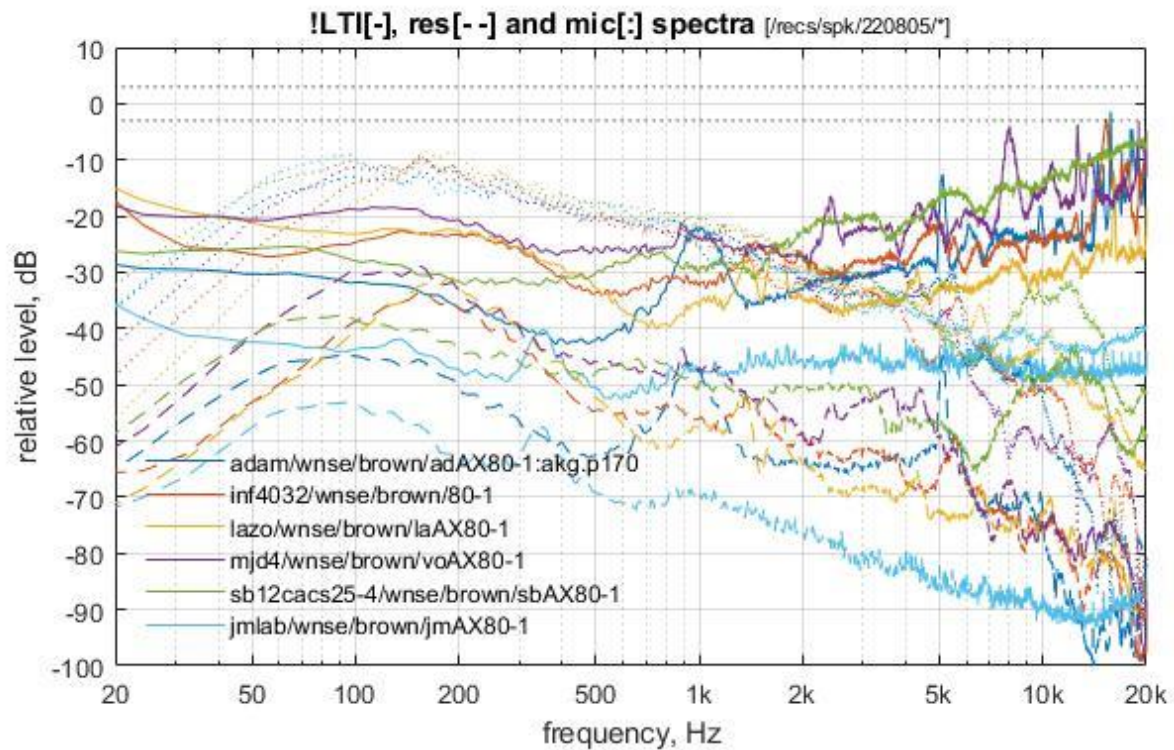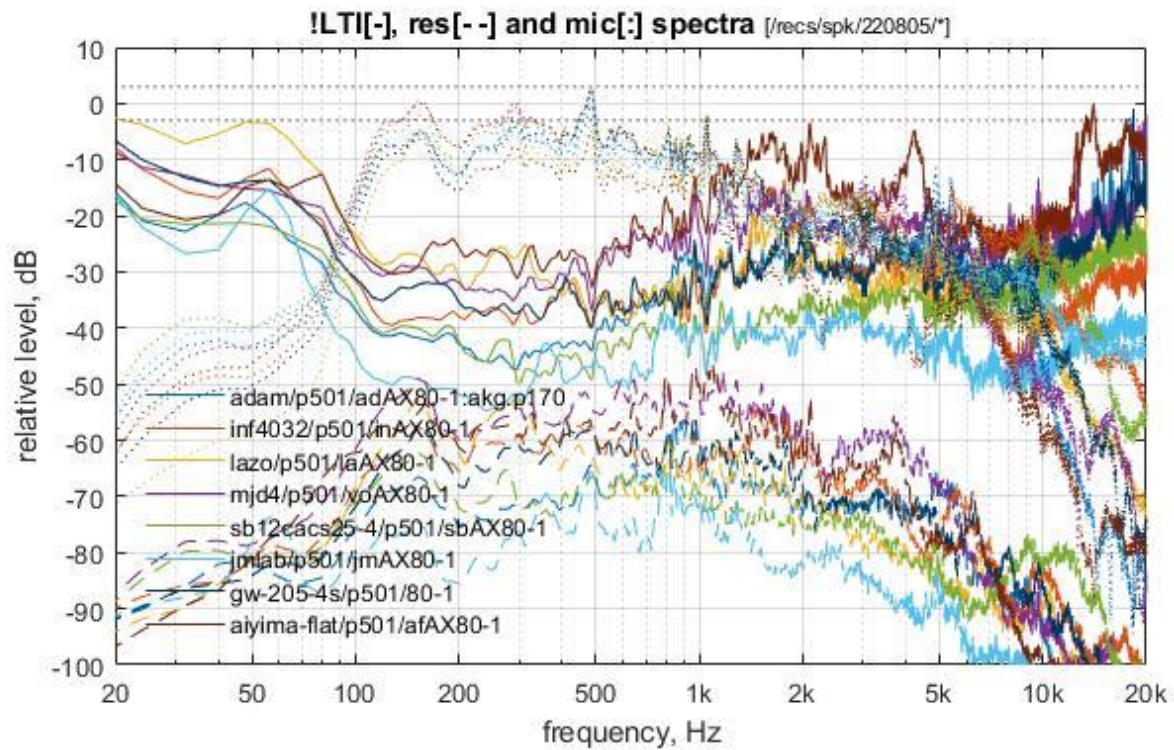
If we band-limit the input to [250…2500] Hz:

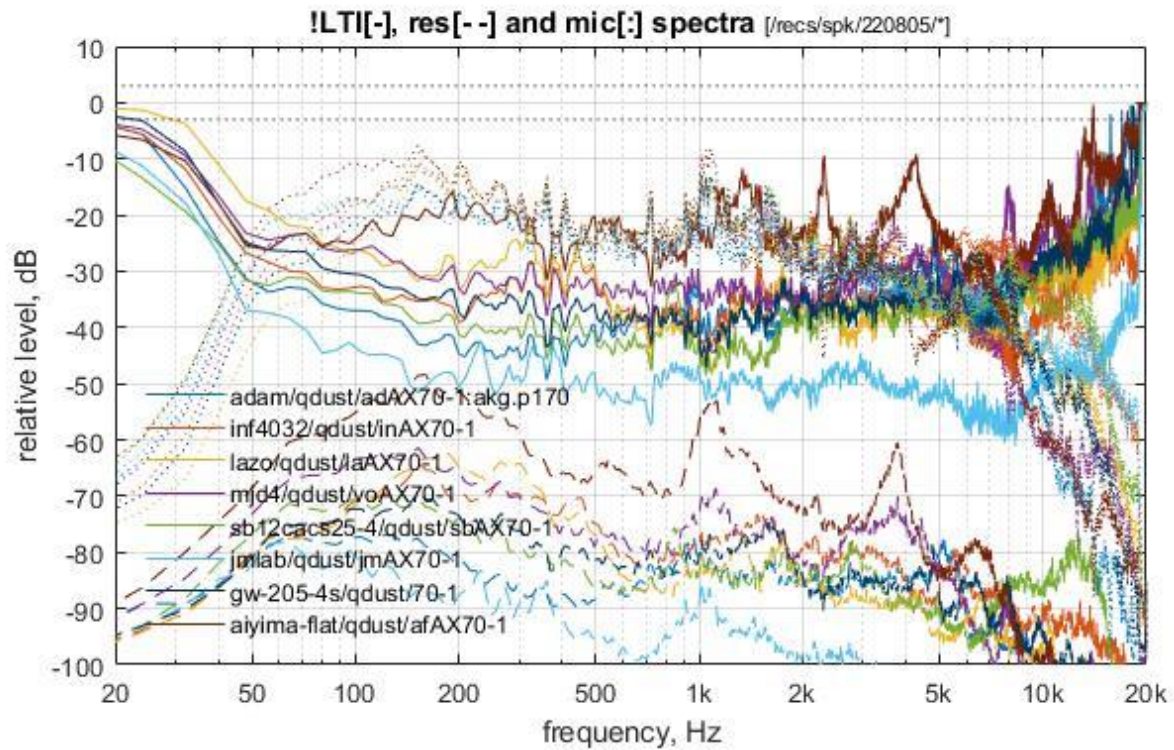### 3.5.2   LTI distortions on ITU-T P.501 speech

The quality differences become more pronounced on the louder output:
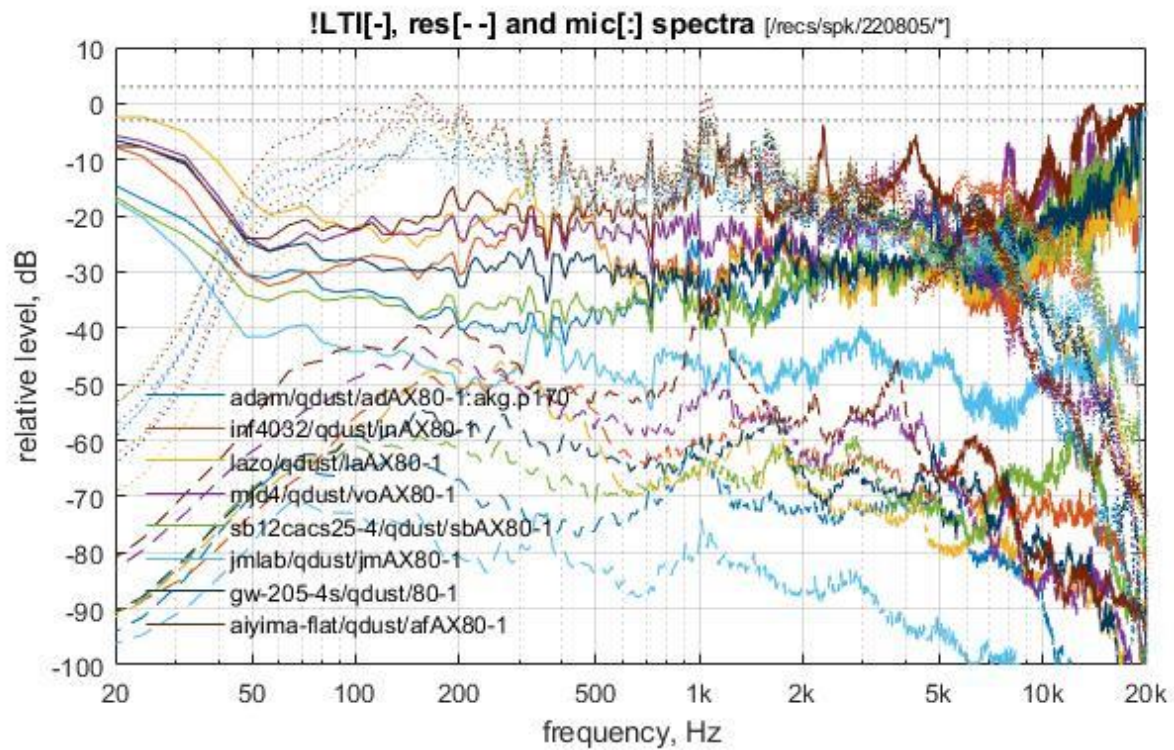
The residual spectra have been shifted down 20dB to improve readability.
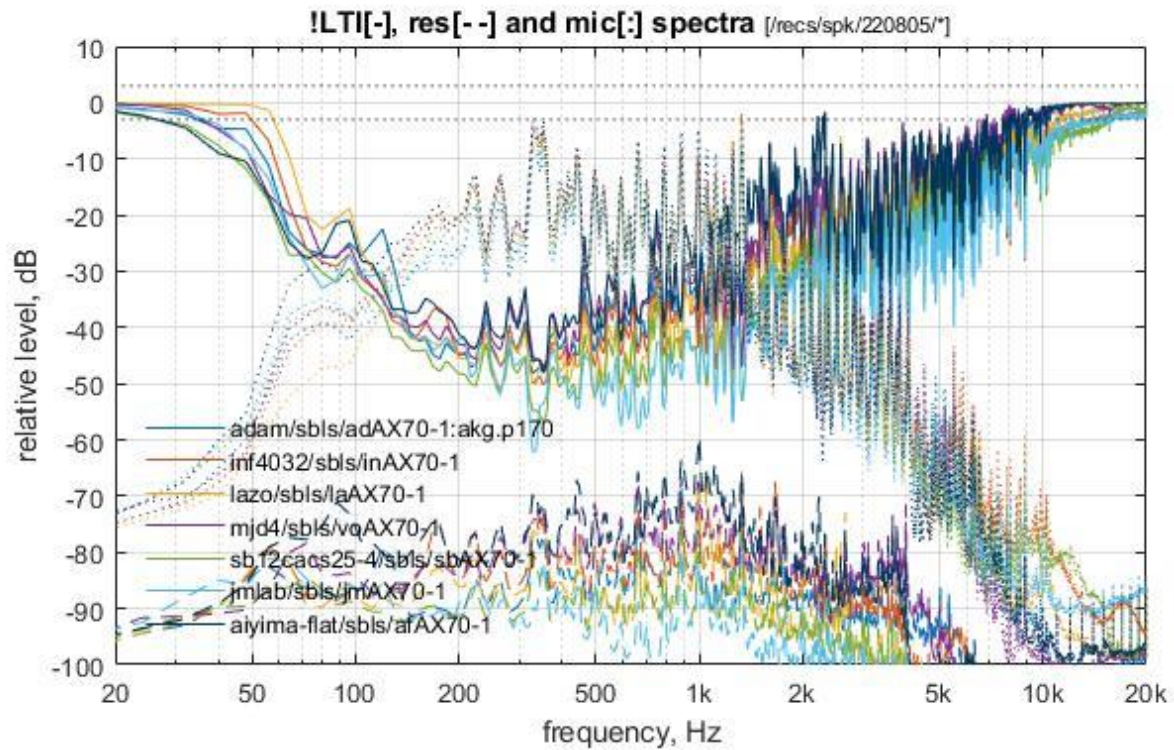
3.5.3   LTI distortions on a Rock music clip

Again, the quality differences become more pronounced on the louder output:
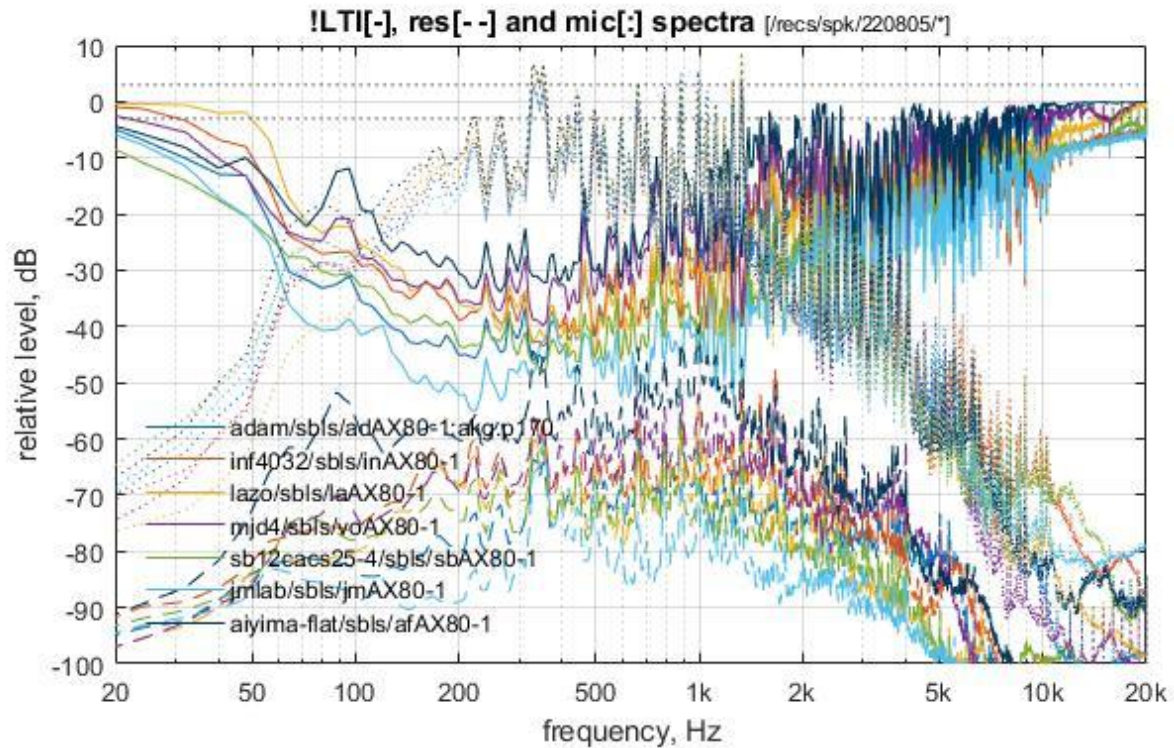
The residual spectra have been shifted down 20dB to increase improve.

### 3.5.4   LTI distortions on a Piano music clip

Same, again:

As observed, each driver (except for the well-polished well-priced JmLab) from a random sample of 3" -
5.25" niche market has its own anomalies and specific distortions which do not seem to fit any reasonably
simple model. The same statement is fully applicable to similarly sampled tweeters.

## 3.6   SB Acoustics (designed by Danesian Audio) SB26ST tweeter

3.6.1   Harmonic distortions
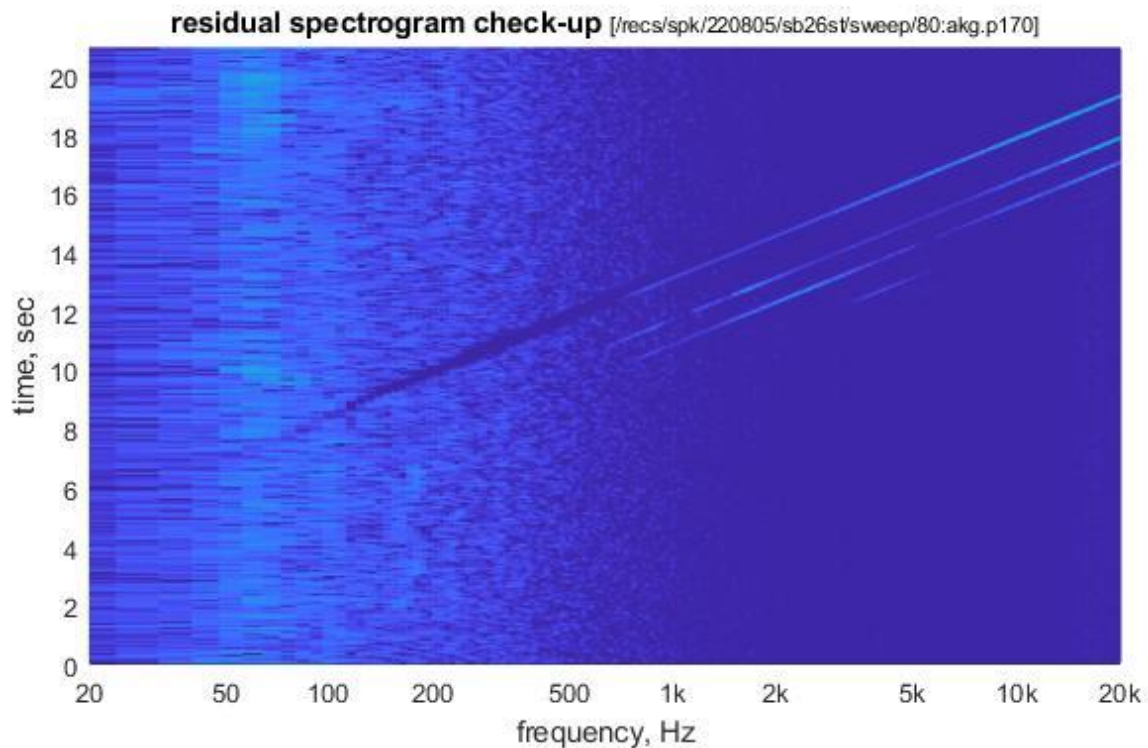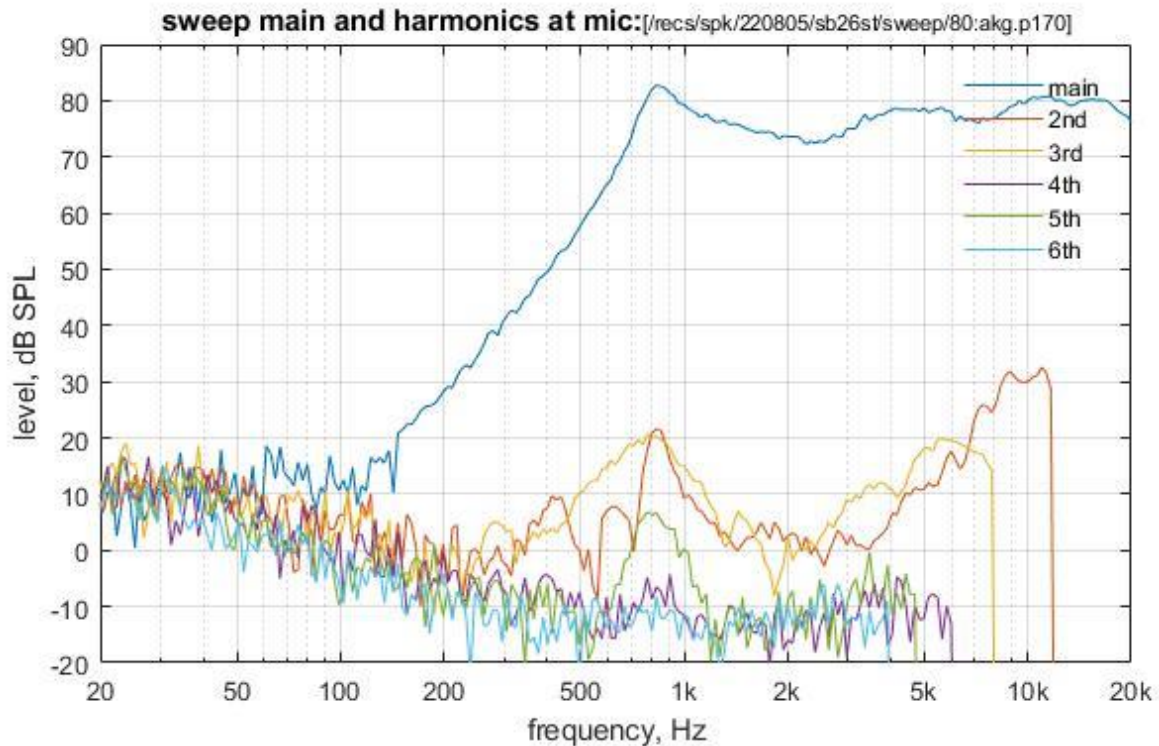
All tweeters have very small and light voice coils. Hence their temperature (and resistance) jumps on stronger input, and most temperamental tweeters display a lot of LTI distortions which are especially annoying on virtuoso piano key attacks. Chirp (or MLS) is a totally useless tool in this regard.
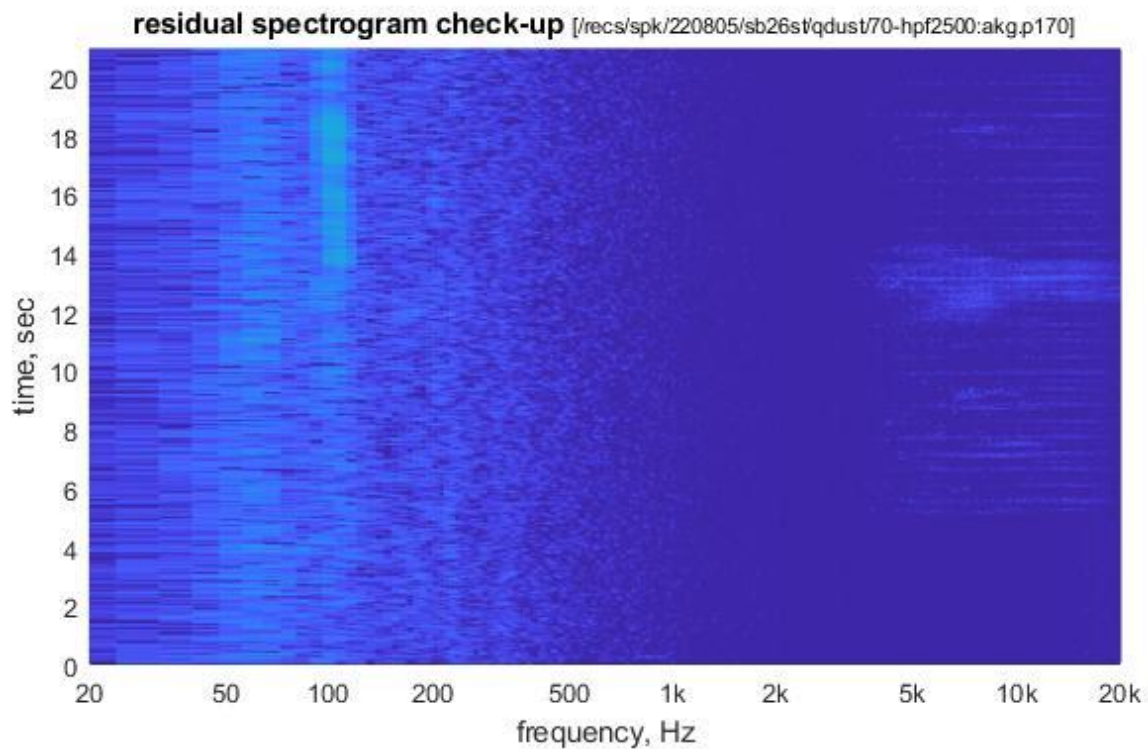
3.6.2   LTI distortions on ITU-T P.501 speech

Input is HPF on 2500Hz:
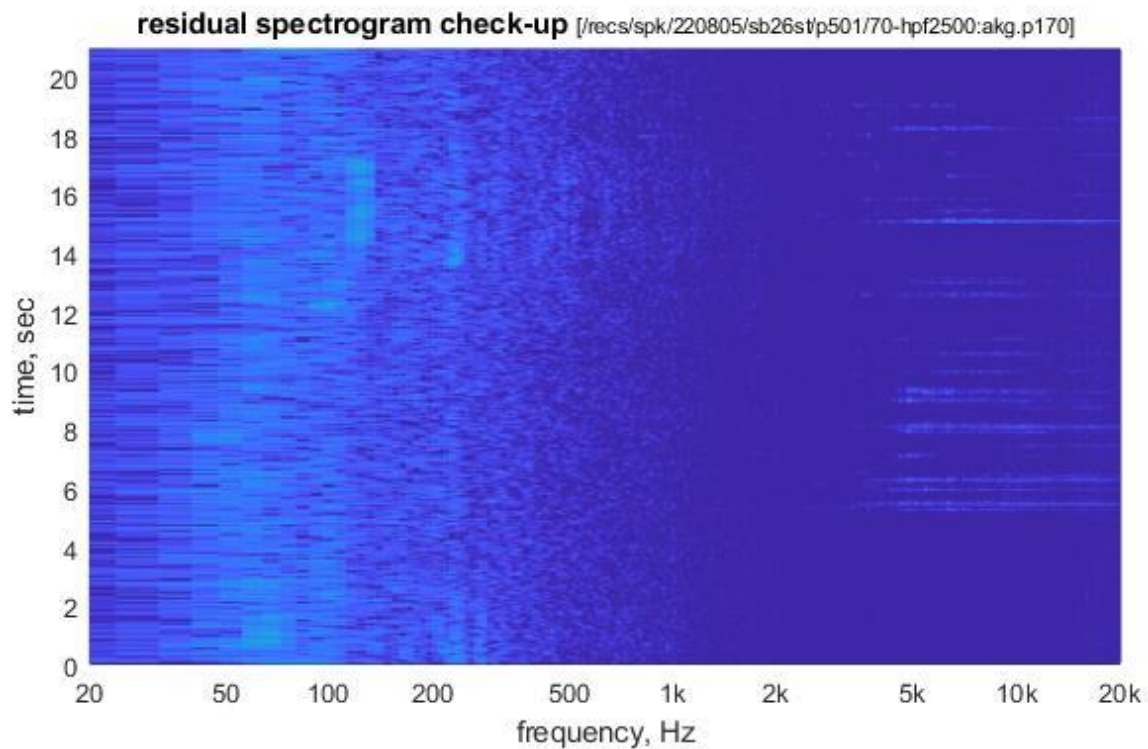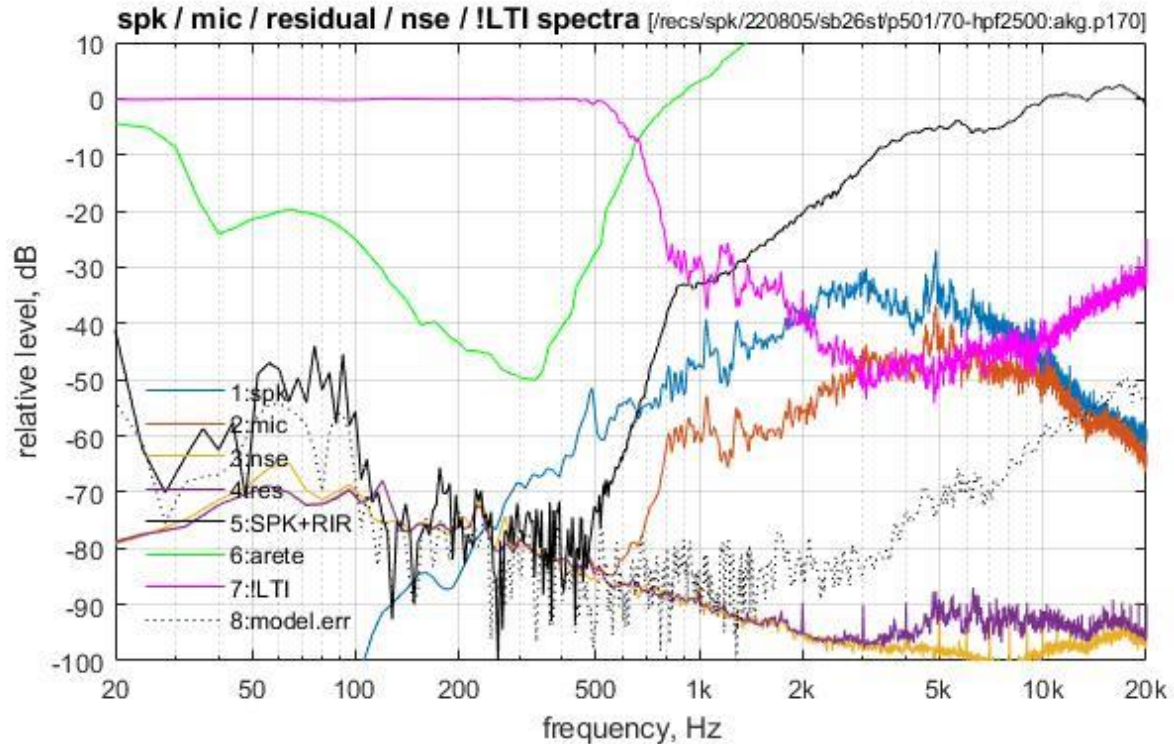
LTI distortions are about -50dB re output, which is on par with the best mid-range drivers and about as good as it gets for traditional drivers. The tweeter FR is easily equalize-able.

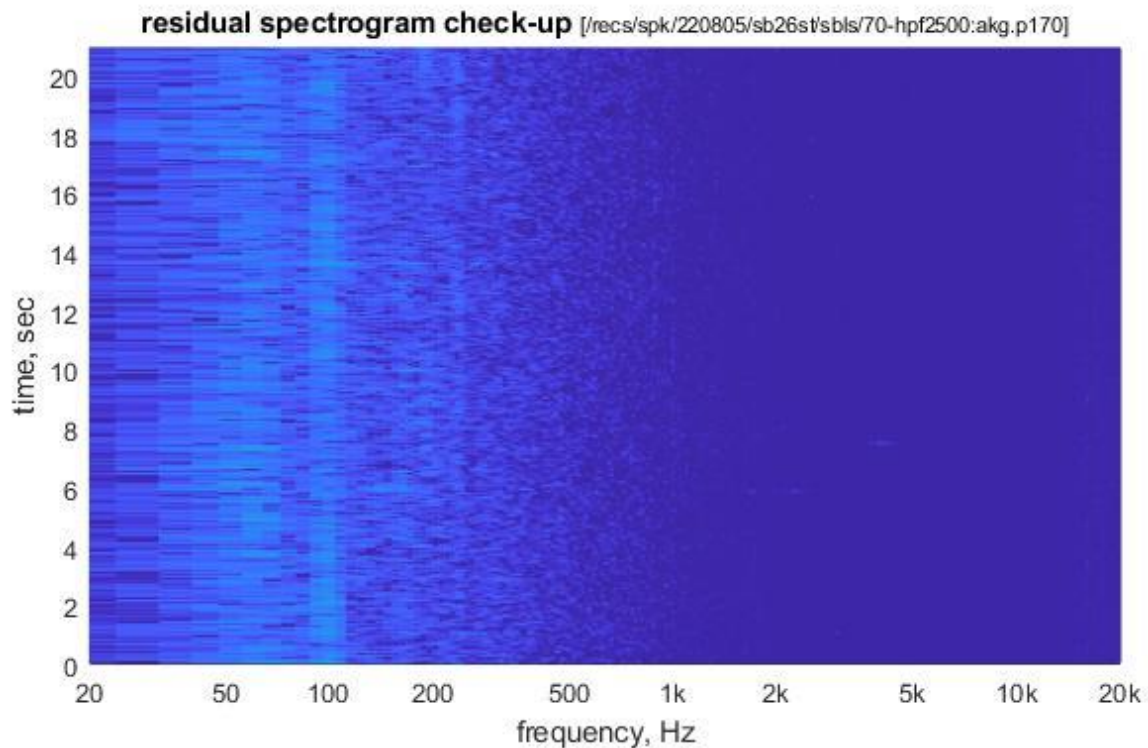### 3.6.3   LTI distortions on a Rock music clip

Input is HPF on 2500Hz:
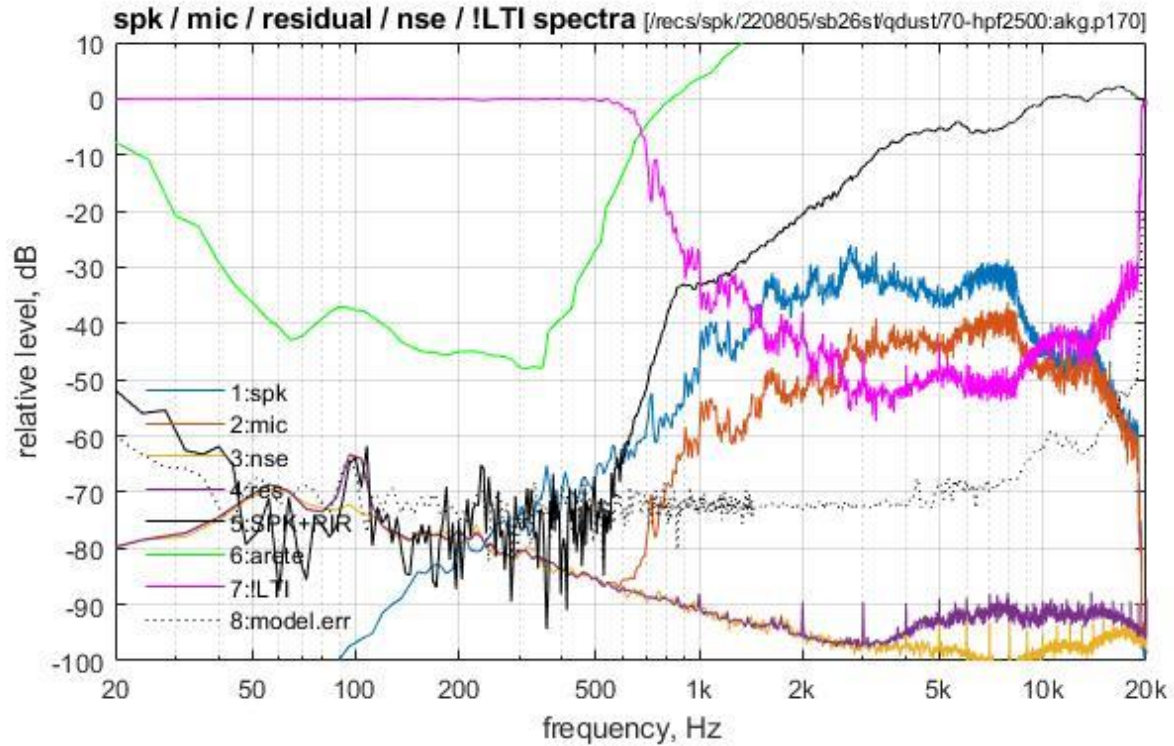
The same behavior.

3.6.4   LTI distortions on a Piano music clip

Input is HPF on 2500Hz:
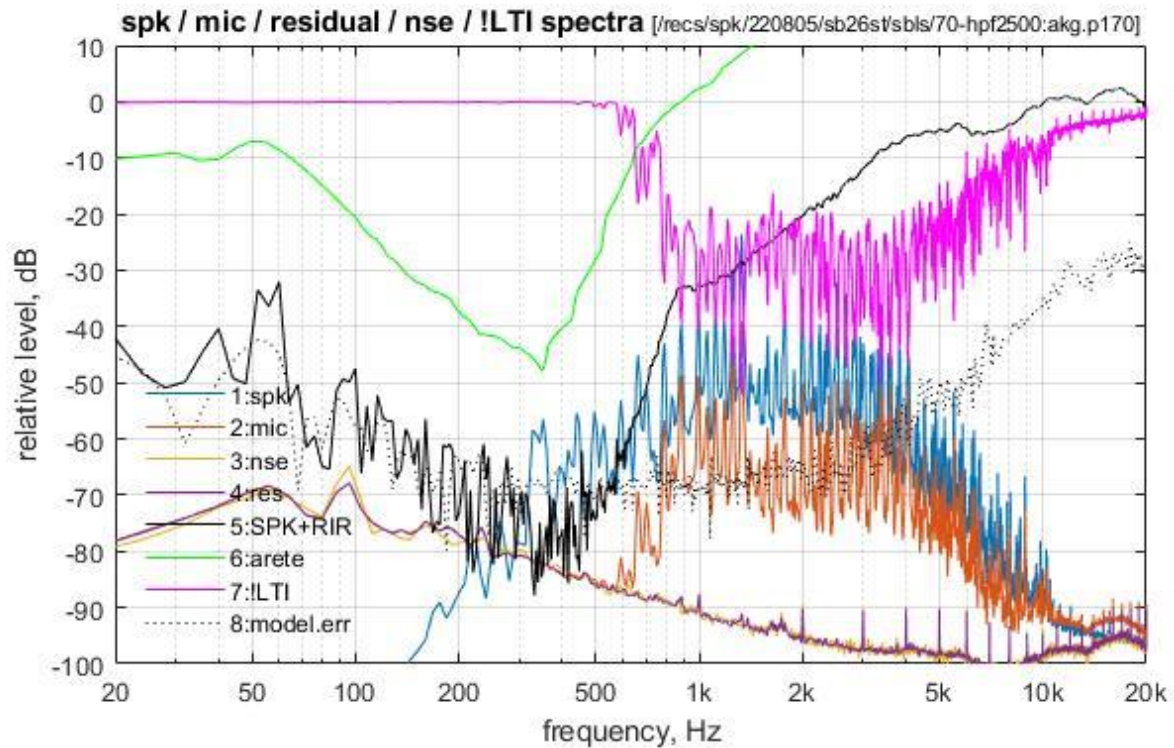
Yes, some tweeters are better than others. This tweeter can be used for AEC.

3.7   OTHER TWEETERS

3.7.1   AMT tweeter

X-ART from ADAM F5: [ADAM Audio - F5 Active Studio Monitor (Nearfield) (adam-audio.com)](): "ADAM Audio has improved upon the Air Motion Transformer utilizing superior geometries and materials to achieve unprecedented audio fidelity." The LTI distortions on ITU-T P.501 speech:
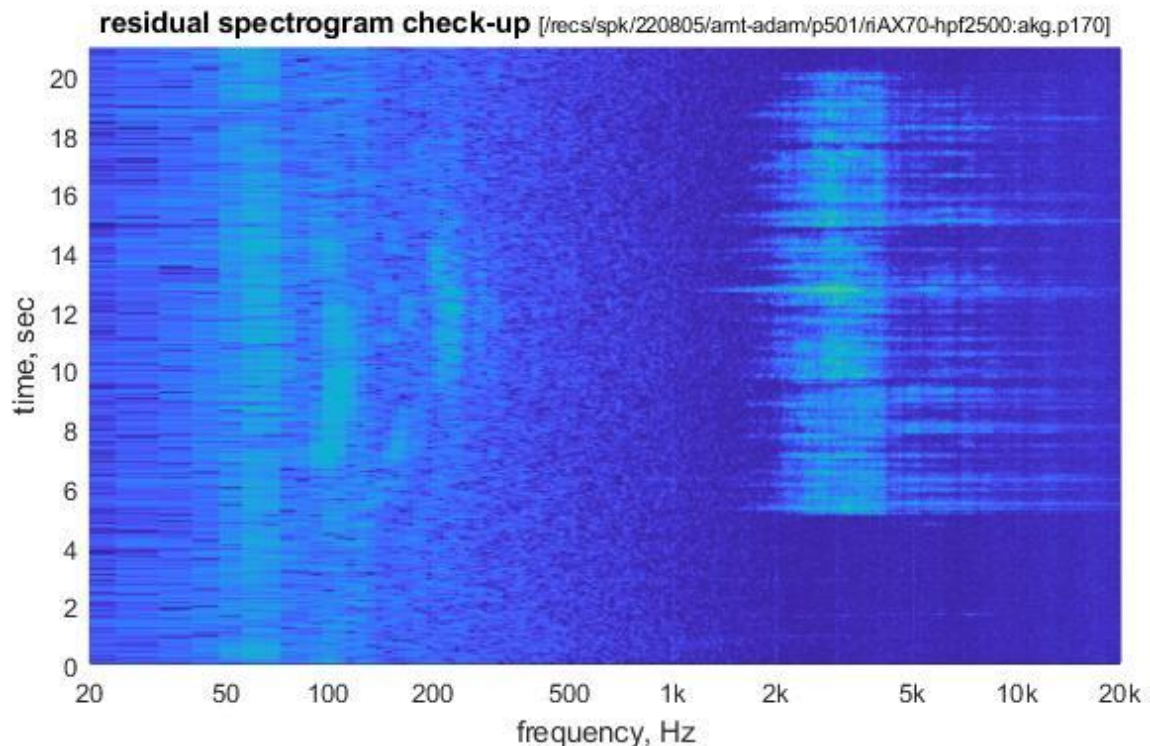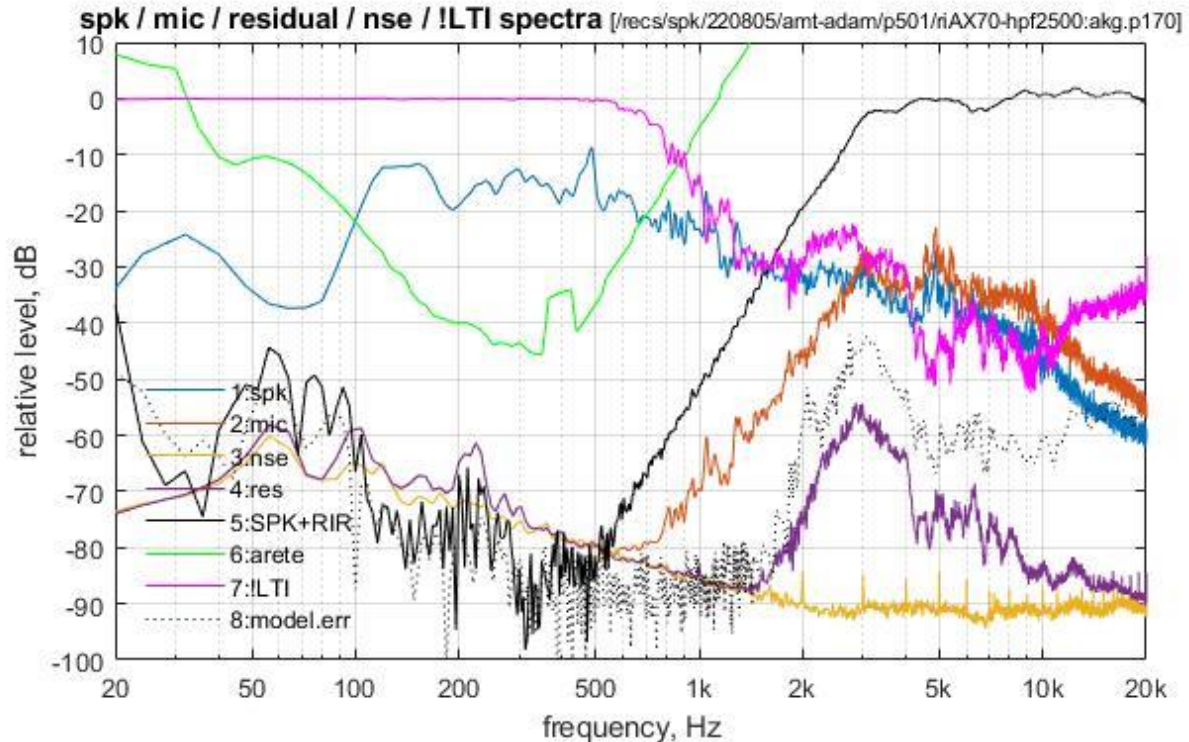
Input is HPF on 2500Hz. Unusable for AEC.

3.7.2   Goldwood Sound GT-501

[Goldwood Sound GT-510 Silk Dome Tweeter 100 Watt 8ohm Replacement Tweeter - Goldwood.com](#), "50 Watts RMS and 100 Watts Max; 3000 - 20,000Hz Frequency Response; 90 dB SPL"
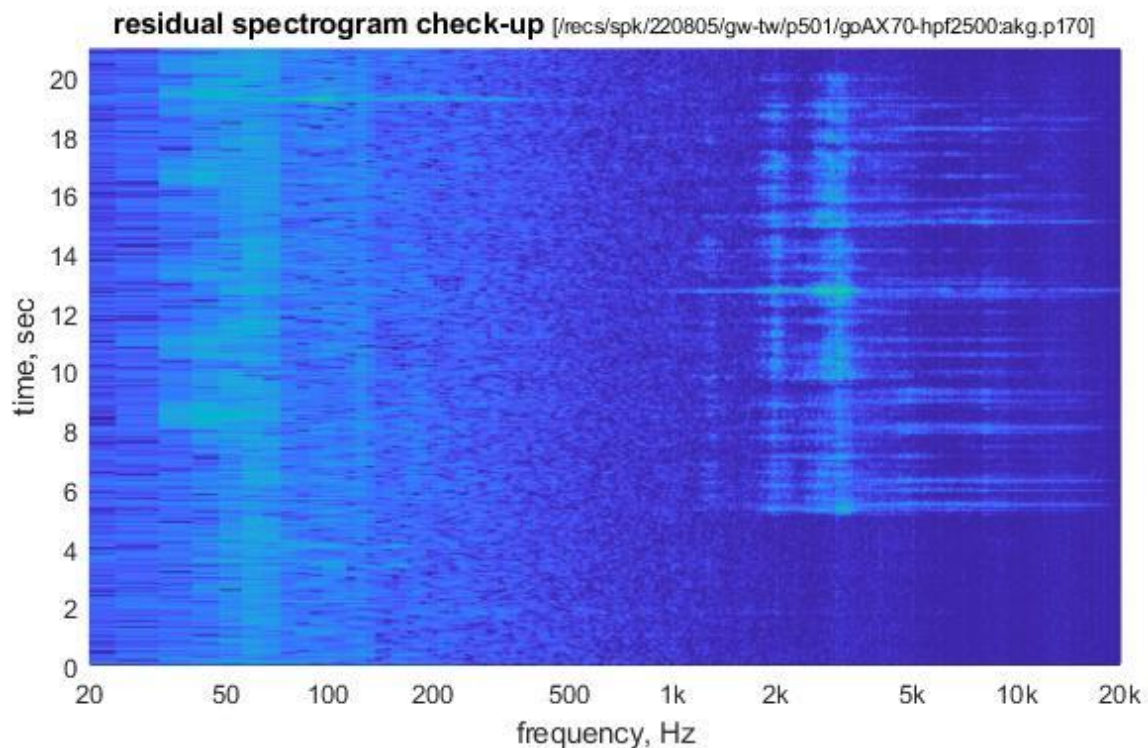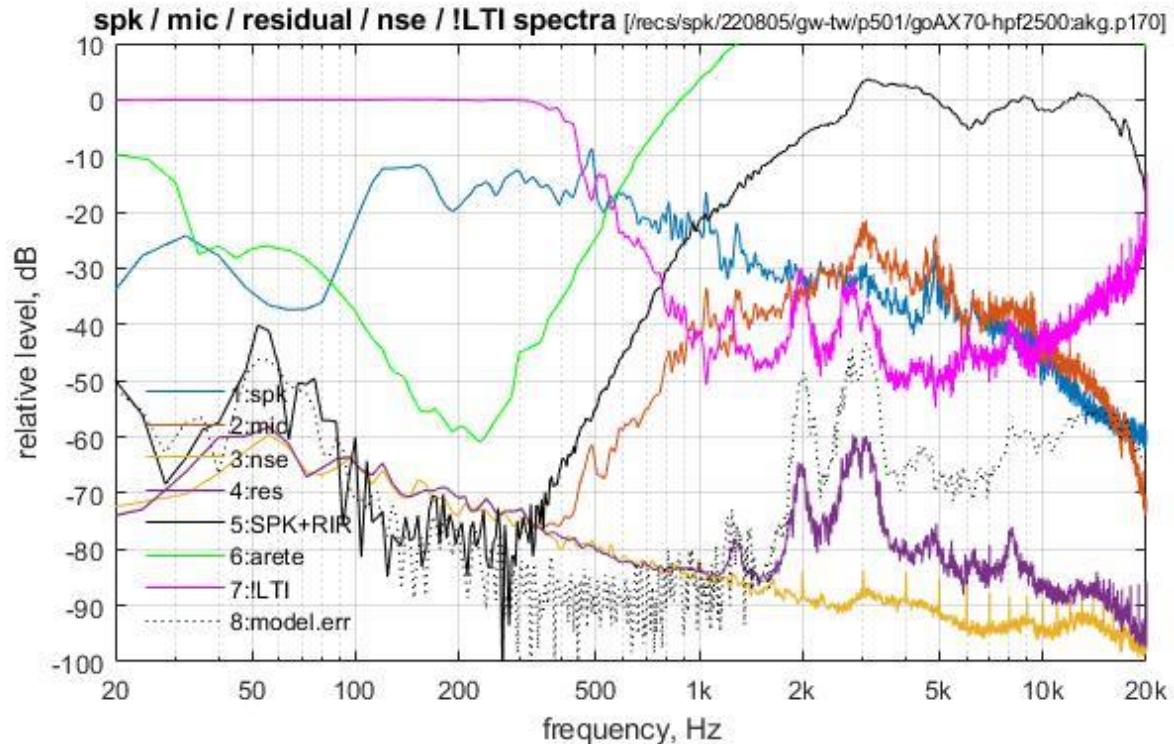
Unusable for AEC.

3.7.3   3.5 inch silk film tweeter

Sold by Aiyima. "Structure: spray paint silver gray panel + high power bare magnetic + silk film;
Impedance: 12 Ohm; Power: 50W; High sensitivity ... sound quality is pure, loud sound is not noisy, wide
applicability."
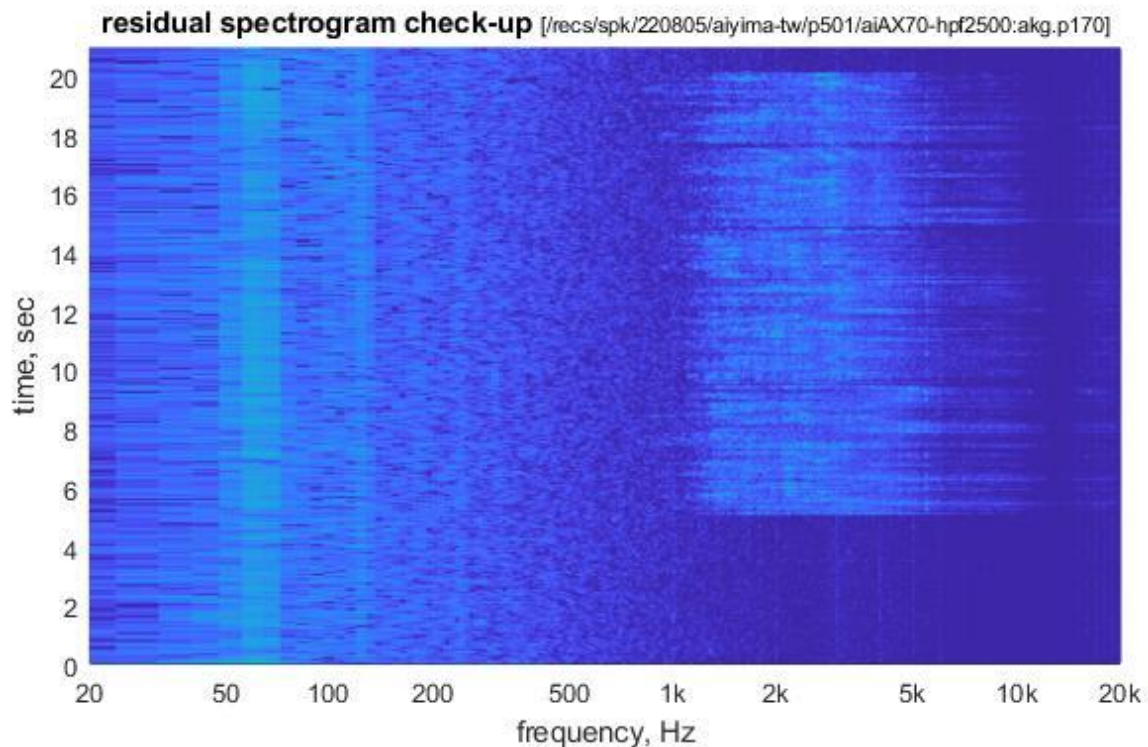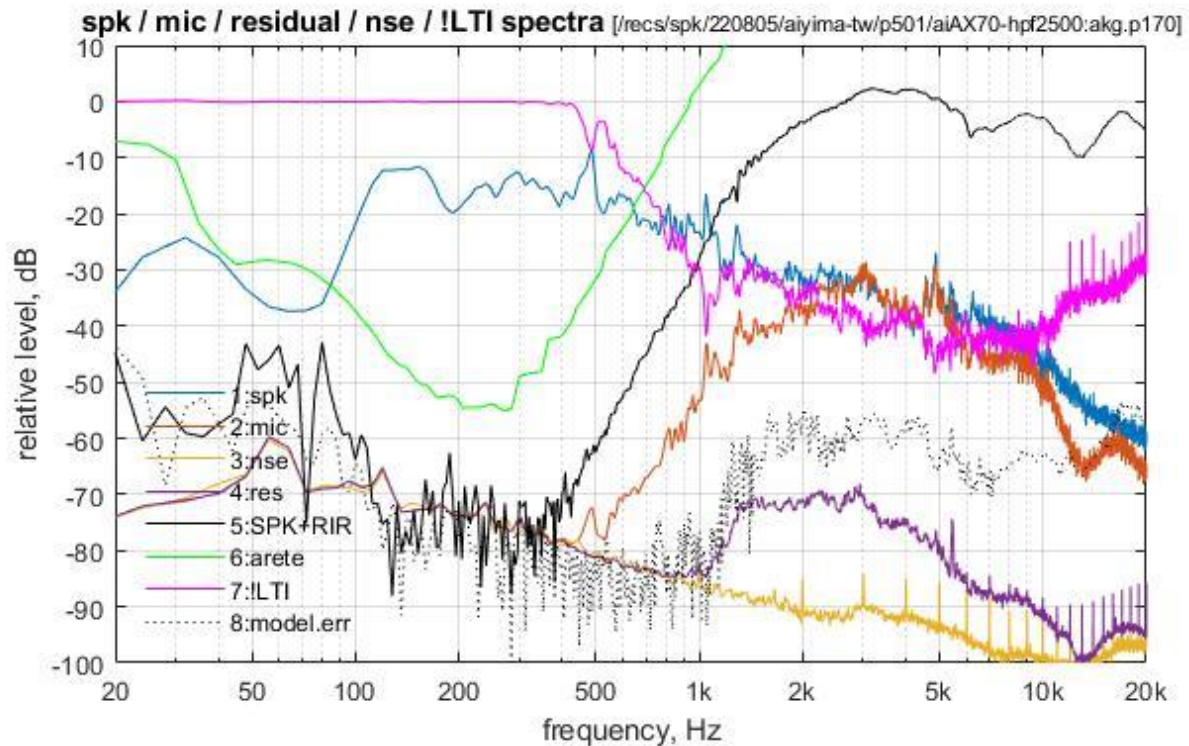
The weirdest distortions I've seen: they start abruptly depending on signal level, jumping from none to -
30dB. Unusable for AEC.

3.7.4   Slim AMT tweeter

"Supper slim flat speaker high power ribbon tweeter planar transducer AMT used on PA HiFi speaker system. RMS Power: 10W; Maximal Power: 20W; Rated Impedance: 6Omega; Sensitivity: 86dB; Frequency Range: 4kHz - 40kHz"
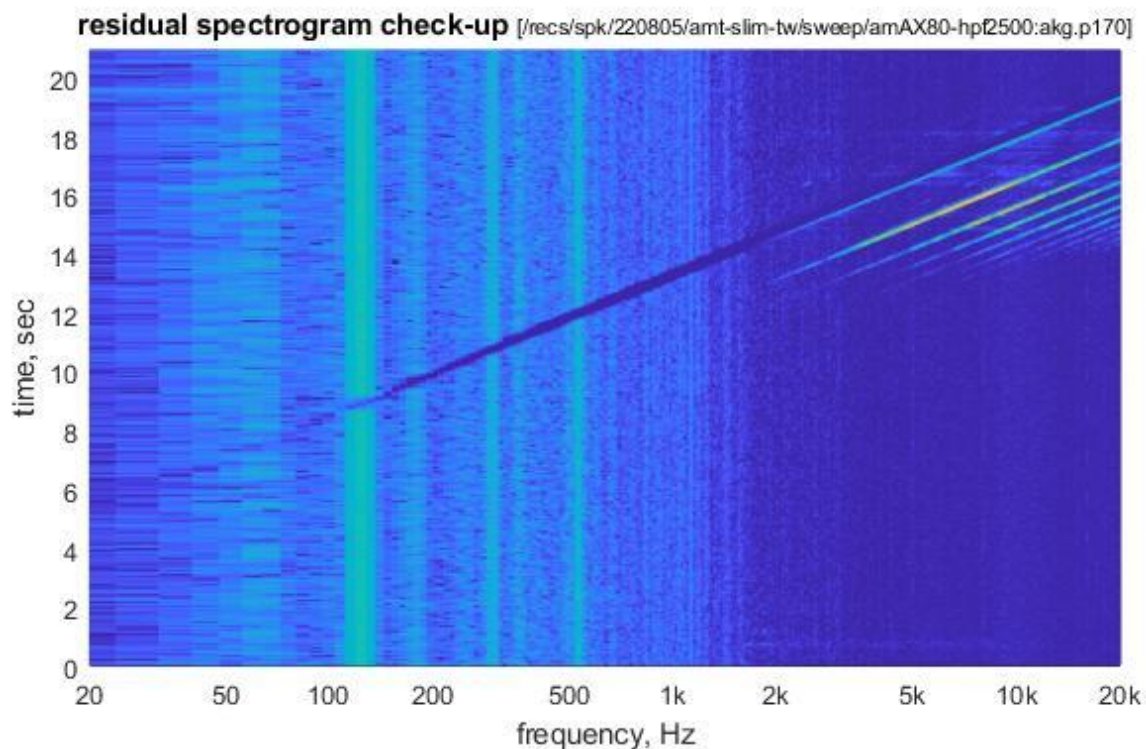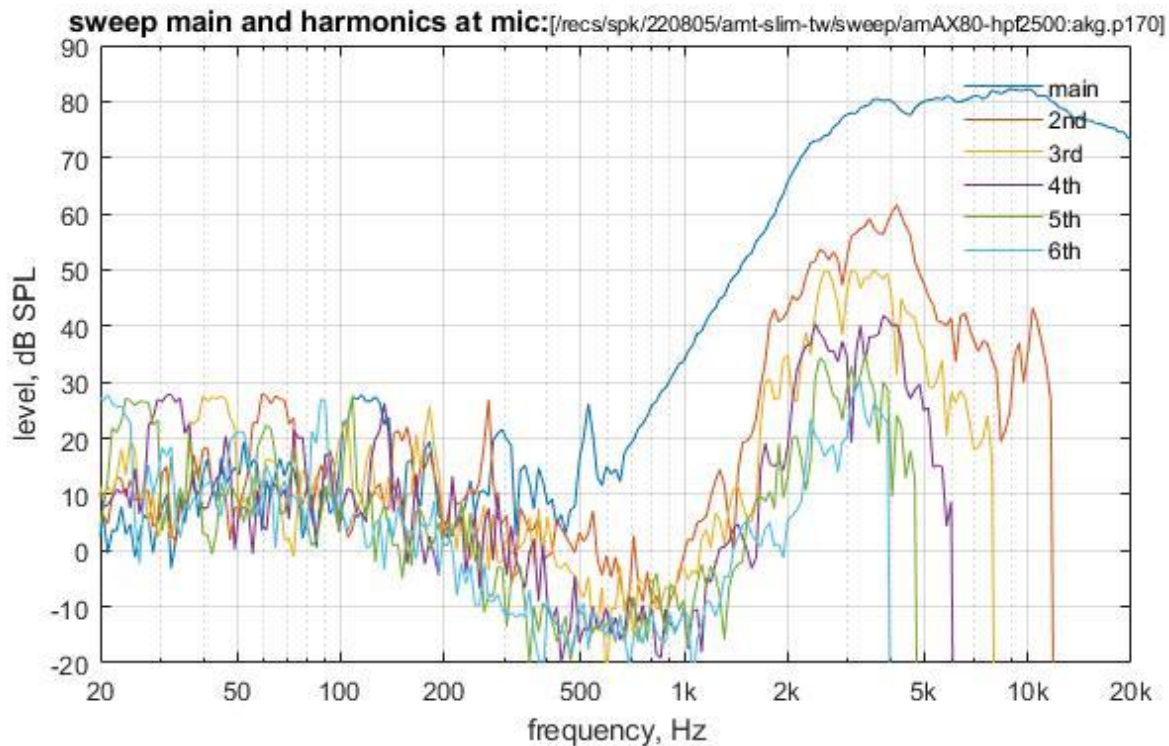
FR is nice and flat but distortions at 80dBSPL (0.25W) are a bit too high. Unusable for AEC.

3.7.5   Pyle Pro PDBT19 Tweeter

PylePro - PDBT19 - Sound and Recording - Tweeters - Horn Drivers (pyleaudio.com):"50 Watts RMS/300 Watts Peak; Sensitivity: 98dB; 2kHz - 25 kHz frequency response. Pyle tweeter drivers are second to none, and feature solid construction and accurate power handling ratings"
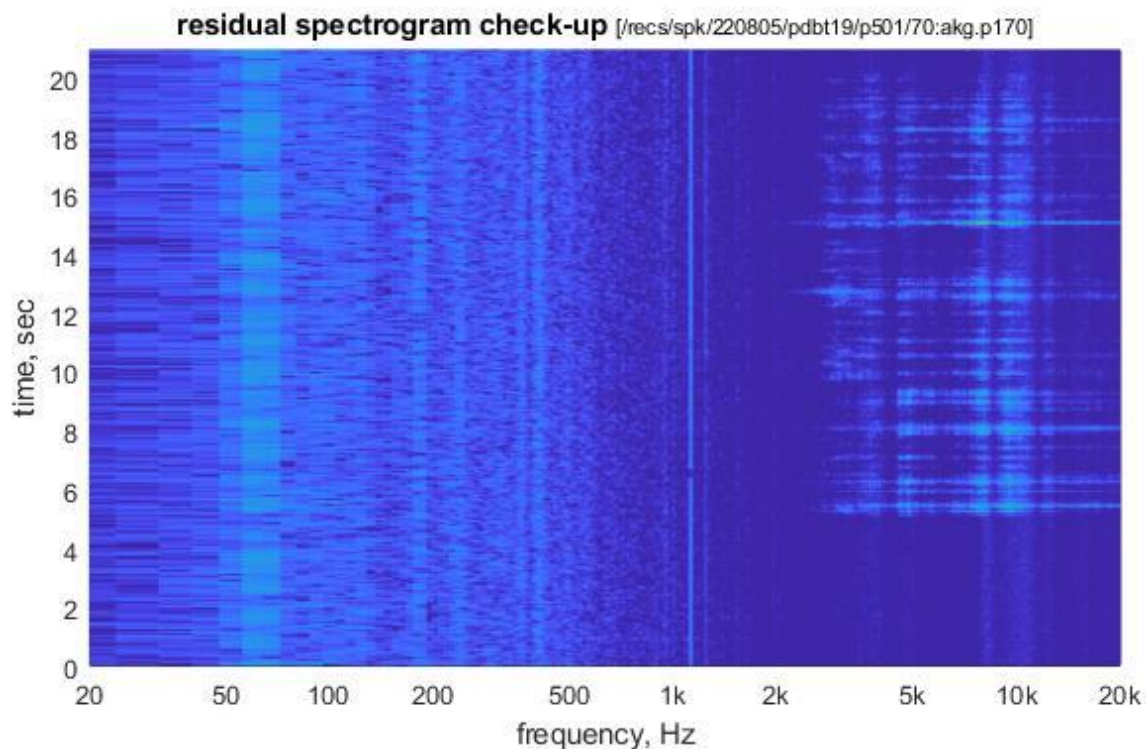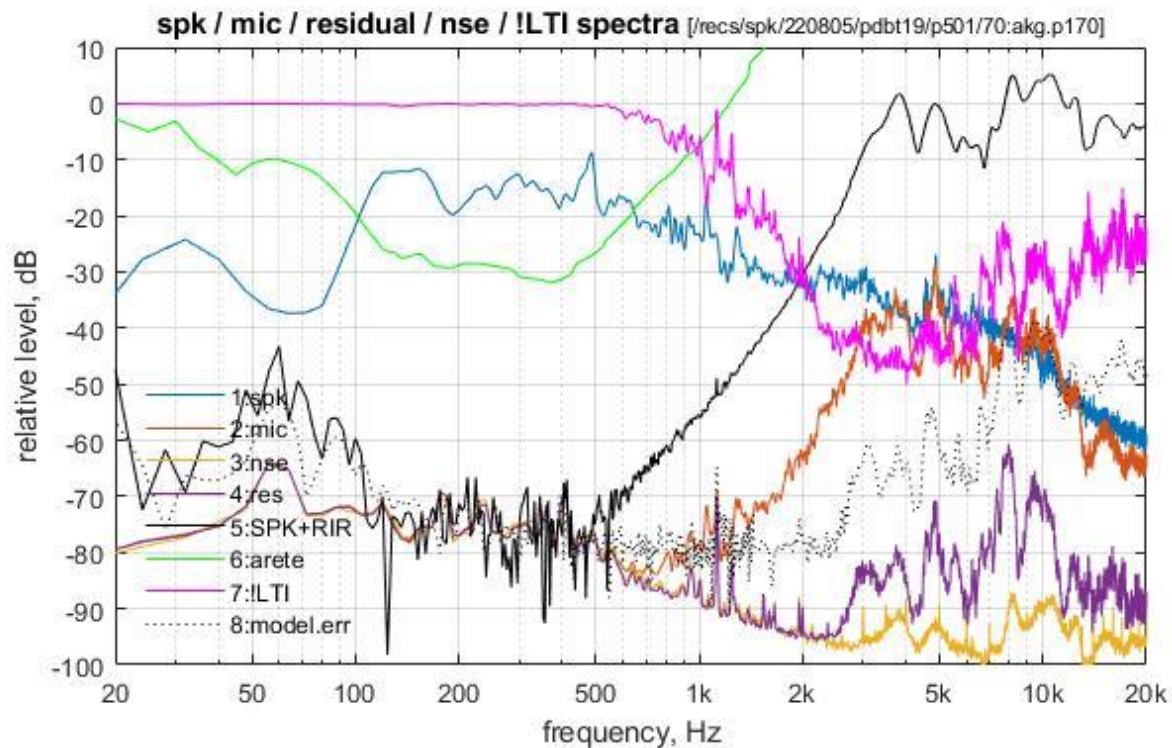
Distortions on Rock clip at 80dBSPL (-18dB re stated 98dB/W/m sensitivity -> 0.016W RMS):
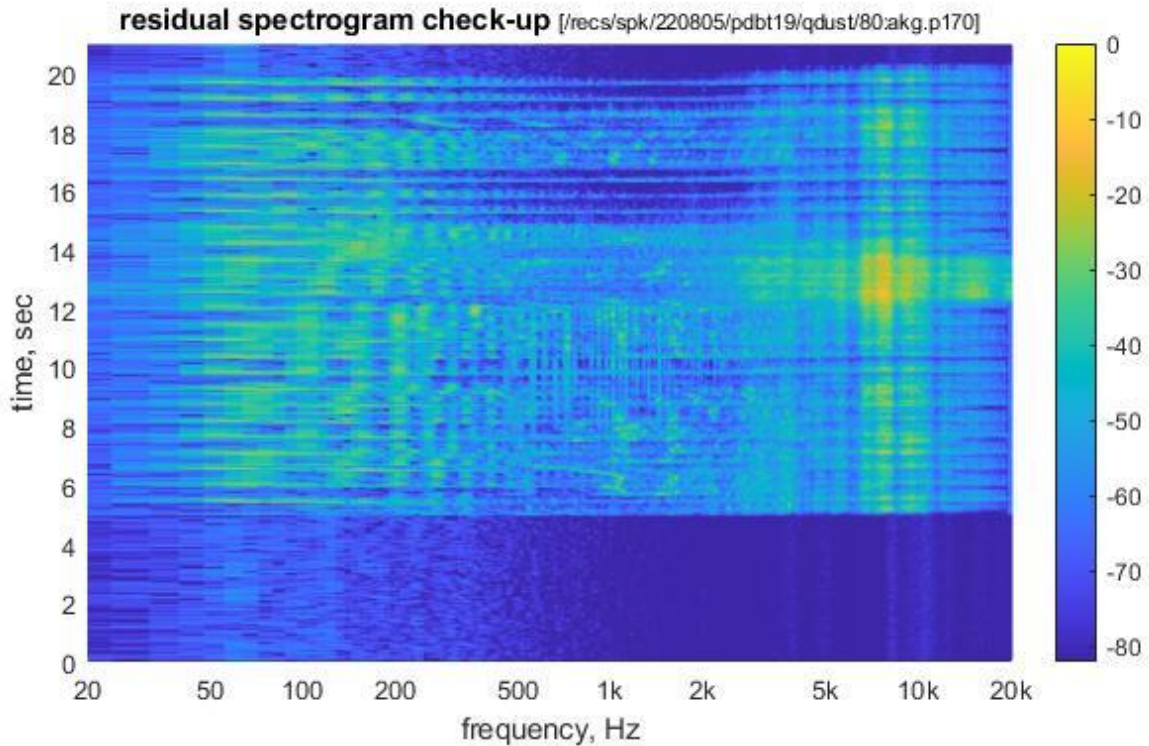
Unusable for AEC.

3.7.6   Bowers & Wilkins CWM 650 tweeter

The 3rd harmonics stays at approximately the same absolute level while output changes from 60dBSPL to 75dBSPL i.e., pure THD (no noise) improves proportionally to the growing signal:
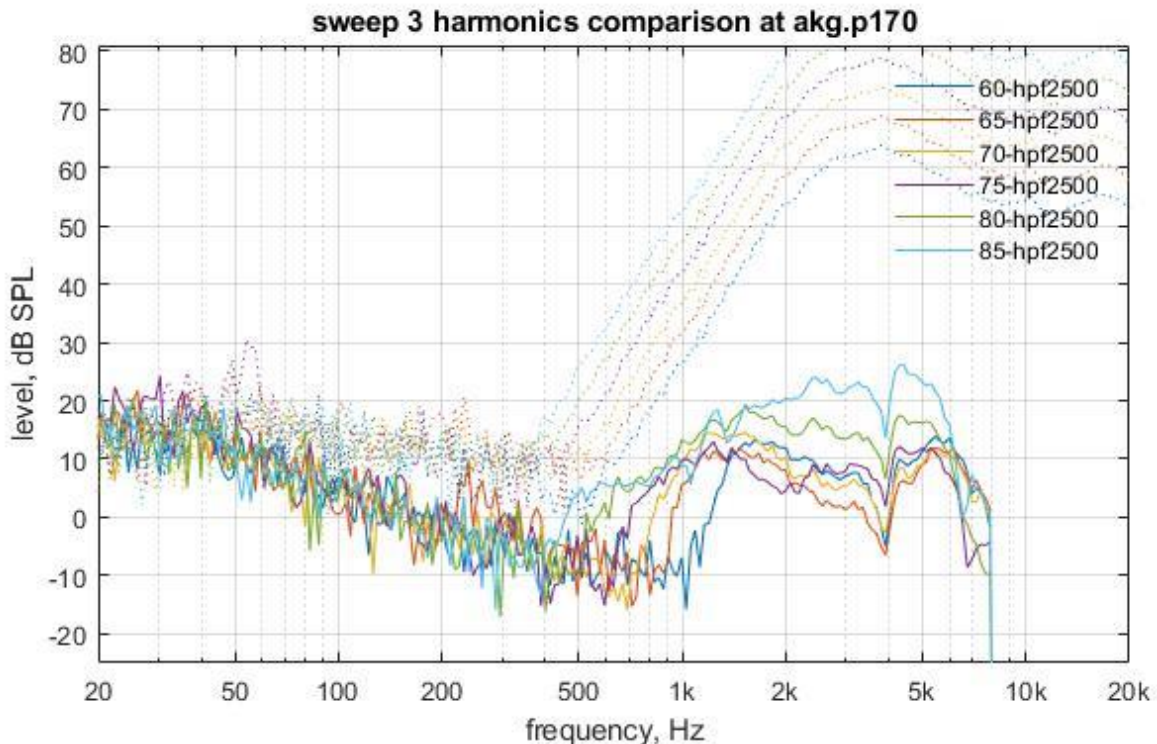
That's not a mistake, the residual on the Rock clip behaves essentially in a same way:
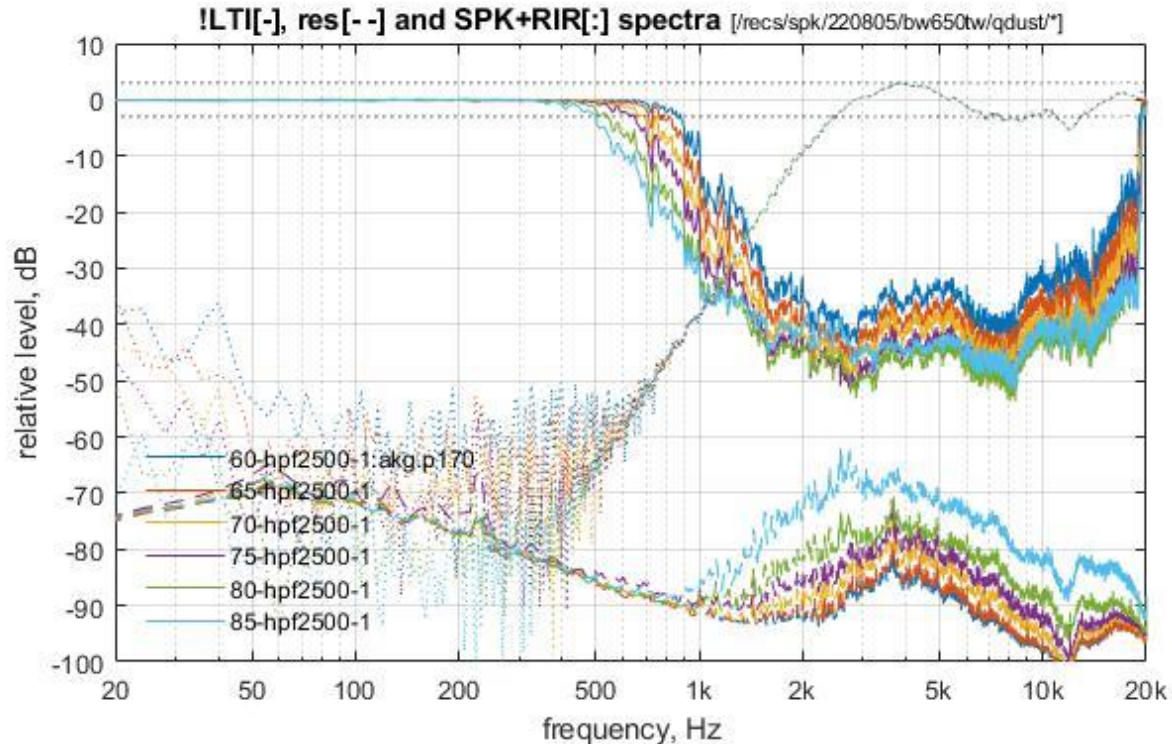
Before this point, you may have decided that you have finally understood how loudspeakers work. Well…Wel…Welcome to the club.

4 DISCUSSION

"Flat panel | AMT | ribbon | electrostatic" loudspeakers do not have shorting rings. Their 3rd (or $2n+1$) harmonic behaves as expected: it increases as x^3 (or x^{2n+1}), which results in a nice sound at low levels and double digits distortions on higher levels[9].

"Traditional" loudspeakers' not-decaying distortions, both in amplitude and frequency, both harmonic and LTI, are due to the "shorting ring" (a.k.a. Faraday ring) which plays a role of internal feedback to reduce distortions, inductance, usually $Le \sim 1/sqrt(f)$ (while partial derivatives shall NOT arise here) but also sensitivity. Unfortunately, the Dr. W. Klippel theory of non-linear distortions in loudspeaker drivers does not see the shorting ring as key feedback loop element. This internal feedback causes 3rd harmonic to be [mostly] of nearly constant level, -40…-50dB, re fundamental, and to dominate over all other distortions in the 50-80 dB SPL range. See page 25 of [10] for explanations why and how this feedback artifact happens.

Loudspeaker looks like a very simple device which it is not. Shorting ring makes the loudspeaker driver a non-Volterra-modellable device in principle. As a rule, such non-linear systems with feedback are… complicated. Exact non-trivial analytic solutions are rarely if ever available, and numerical solutions most often are not robust to the perturbation of boundary conditions. This problem is well beyond the reach of modern science[11], and the attempts of designing a perfect voltage driven passive loudspeaker driver or a

---

[9] May be, that's why Magnepan / AMT and B&W enthusiasts never agree.

[10] Simulation of Crossover Distortions in Class AB 0.00002% THD - File Exchange - MATLAB Central (mathworks.com)

[11] Meaning: solvable only by "audioscientists", internet "experts", and climatologists.

"mirroring" filter are harmless but quintessentially fruitless fun. Full adaptive cancellation of RIR with LTI distortions has the same [absolutely overwhelming] complexity as the loudspeaker linearization which is much easier to approach by "normal" means, in place.

"Normal" is the Lorentz way. Lorenz / reluctance actuators, the workhorses of semiconductor industry, achieved sub-nanometer precision more than 10 years ago, using current control, position, airgap and flux sensors, with nested feedback loops by adaptative control. Alas, loudspeakers did not change much in the last 40 years.

Douglas Self writes in "Audio Power Amplifier Design", 6th edition, on page 351: "...concept of Nested Differentiating Feedback Loops (NDFL), introduced by Edward Cherry in 1982. The original JAES paper is tough going mathematically. ... Cherry lost 99% of his [AES] audience when he launched suddenly into complex algebra and Laplace variables." These are exactly the disciplines required for achieving low LTI distortions - for adaptive feedback control of non-linear systems is neither for amateurs nor tinkerers.

Let's perform a Bayesian test of a hypothesis that audio is mostly a huge snake oil market and "audioscience" is a pile of childish gibberish[12] based on the implicit assumption that all links of entire audio reproduction chain are LTI. "Audioscience" is not capable of assessing LTI distortions adequately, nor of controlling them. Its tools and test methods are based on the same unfalsifiable assumption.

In absence of adequate objective measuring tools (therefore, of feedback & "invisible hand" too), unregulated markets (with naïve customers who have to trust "experts") shall follow the path of least resistance for the "human nature", and the outcome has to be fairly predictable, which it is: [How To Measure Speakers - And Why Many Pros Don't Do It At All - YouTube](). Are there any exceptions? Yes. There are companies who don't lie and occasionally (but not consistently) make decent loudspeakers, but they are few and far between. There are researchers who are competent, with integrity, like Drs. W.Klippel, P.Brunet, and E.Merilainen but again they are few and far between. Yet, even the best today's loudspeakers leave much to be desired. Etc, etc...

Digital cameras are much more complicated than any audio reproductions chain components including loudspeakers. Digital cameras progress has been astonishing, with new stars rising and false pretenders being rounded up in a course of few years. In a large part, this progress was stipulated by timely competent objective review sites with proper measurements, such as dpreview and dxo. The vendors' marketing statements and supplied data became irrelevant. Nobody read them – everyone read the trusted dpreview. If a camera was not reviewed by them – it did not exist. Such reviews served as a market feedback error sensor, and this feedback control loop brought digital cameras from mid-90s infancy into maturity in less than 20 years. Nothing like dpreview have ever existed in the audio field.

5   CONCLUSION

It has been reliably observed that LTI distortion patterns vary wildly between mid-grade drivers, not even talking about [sub]entry-level drivers, and that such thing as a good generic loudspeaker was not observed. Therefore, a good AEC shall be tailored to the LTI distortions of specific driver(s), which have to be measured explicitly: לא נבדק - לא עובד. Therefore, step size control and NLP of a smartphone AEC, with an overdriven loudspeaker mechanically coupled to microphone, will end up to be quite different from a high-end conferencing room AEC. Therefore, *there is no such thing as a good generic AEC*, and there is not much point in improving AEC without improving loudspeakers.

---

[12] Psychology used to be of the same kind before "replication crisis" was admitted.