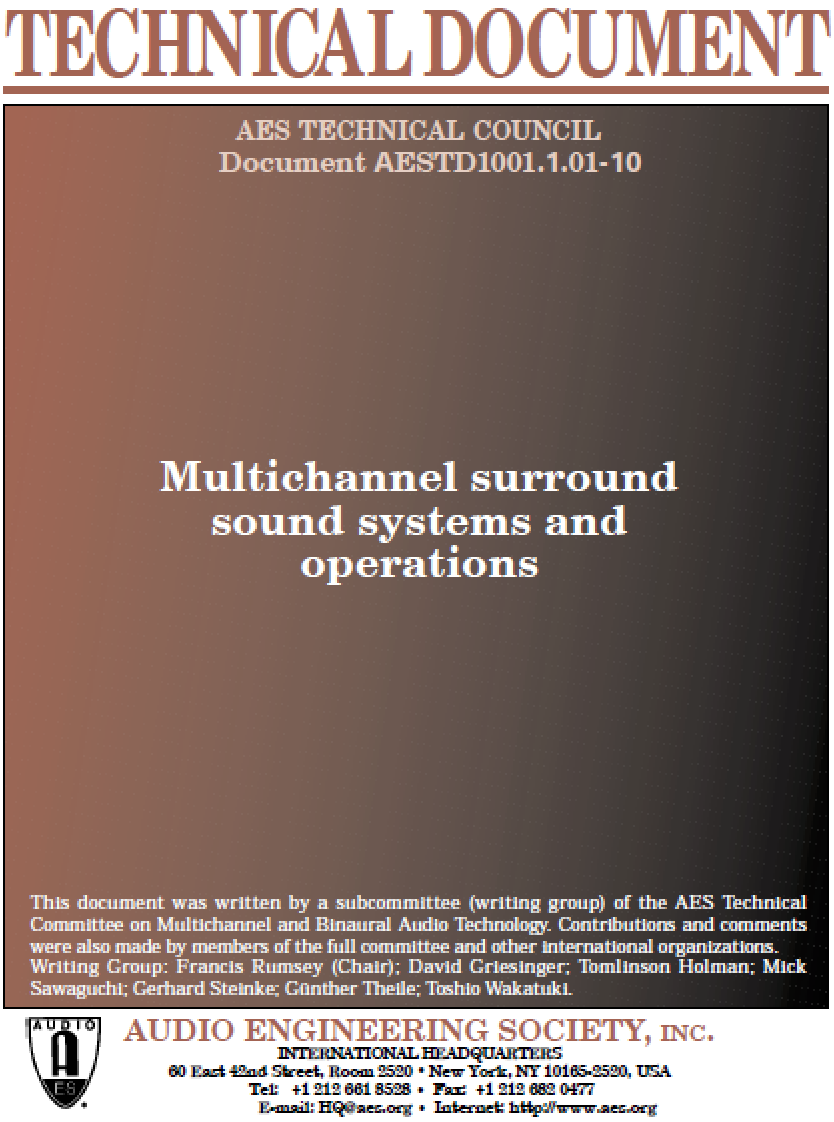
**Comments on AESTD1001.1.01-10**

Source: http://www.aes.org/technical/documents/AESTD1001.pdf


TECHNICAL DOCUMENT

AES TECHNICAL COUNCIL
Document AESTD1001.1.01-10

Multichannel surround sound systems and operations

This document was written by a subcommittee (writing group) of the AES Technical Committee on Multichannel and Binaural Audio Technology. Contributions and comments were also made by members of the full committee and other international organizations. Writing Group: Francis Rumsey (Chair); David Griesinger; Tomlinson Holman; Mick Sawaguchi; Gerhard Steinke; Günther Theile; Toshio Wakatuki.

AUDIO ENGINEERING SOCIETY, INC.
INTERNATIONAL HEADQUARTERS
60 East 42nd Street, Room 2520 • New York, NY 10165-2520, USA
Tel: +1 212 661 8528 • Fax: +1 212 682 0477
E-mail: HQ@aes.org • Internet: http://www.aes.org

The above document makes a number of recommendations, designed to deliver quality audio experience to your ears. Some of the recommendations attracted comments shown below.

The AESTD1001.1.01-10 document does not discuss what is audible or what's not, and I am not doing that either. It simply presents a set of parameters, deemed to be important in audio transduction. You are welcomed to exceed them.

## 1. Frequency Response

Table 3. Suggestions for reference monitor loudspeakers and advice for home loudspeakers.

| ***Parameters*** | ***Units/Conditions*** | ***Value*** |
|---|---|---|
| **Amplitude/frequency response**<br><br><br><br>Difference between front loudspeakers | 40 Hz–16 kHz<br>0°<br>±10°<br>Horizontal ±30°<br><br>In the range<br>>250 Hz to 2 kHz | <br>Tolerance 4 dB<br>Deviation to 0°, 3 dB<br>Deviation to 0°, 4 dB<br><br>0.5 dB |
| **Directivity index** | 250 Hz–16 kHz | 8 dB ±2 dB |
| **Nonlinear distortion attenuation**<br>(SPL = 96 dB) | <100 Hz<br>>100 Hz | –30 dB (=3%)<br>–40 dB (=1%) |
| **Transient fidelity**<br>Decay time $t_s$, for reduction to a level of $1/e$, i.e., 0.37 of output level | $t_s$[s] | <5/$f$ [Hz]<br>(preferably 2.5/$f$) |
| **Time delay**<br>Difference between stereo loudspeakers | ∂t | ≤10 µs |
| **System dynamic range**<br>Maximum operating level (measurement acc. to IEC 60268, § 17.2, referred to 1 m distance) | $L_{\text{eff max}}$ | >112 dB<br>(at IEC 60268 program simulation noise or special condition) |
| **Noise level** | $L_{\text{noise}}$ | ≤10 dBA |

Loudspeaker system designers would typically present and document their products with several critical parameters, suggested by the AES Technical Council. One of the most obvious is the on-axis frequency response of the loudspeaker system. Rather than typical +/-3dB tolerances, the AES TC recommends tighter tolerances of +/-2dB. Also, in order to meet the AESTD1001.1.01-10, the loudspeaker system must provide 40Hz-16kHz frequency response with +/-2dB tolerance.

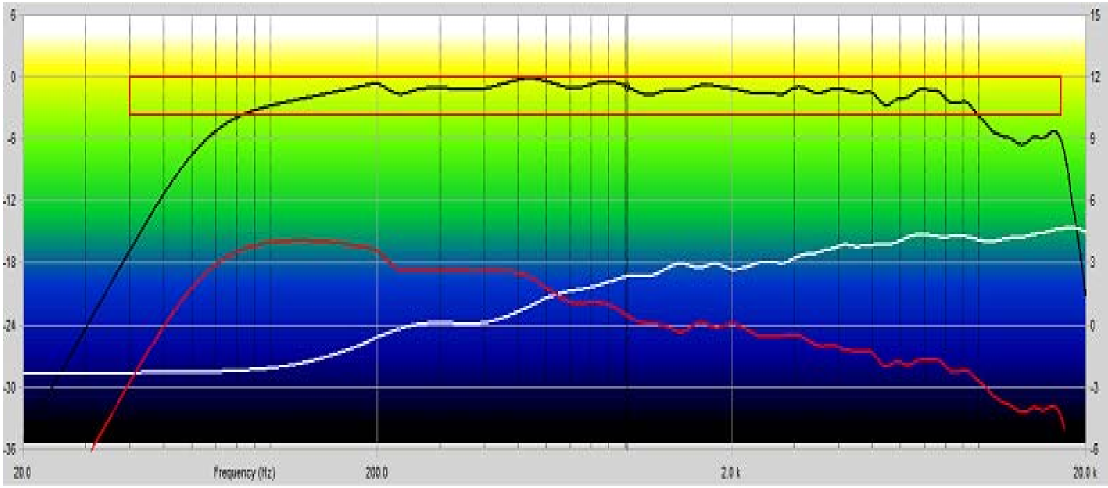



Figure 1. An example of a loudspeaker, which falls outside recommendations of AESTD1001.1.01-10 – red rectangle shows recommended SPL tolerance.

## 2. Difference Between Stereo Loudspeakers / Phase Response

Table 3. Suggestions for reference monitor loudspeakers and advice for home loudspeakers.

| *Parameters* | *Units/Conditions* | *Value* |
|---|---|---|
| **Amplitude/frequency response** | 40 Hz–16 kHz<br>0°<br>±10°<br>Horizontal ±30° | <br>Tolerance 4 dB<br>Deviation to 0°, 3 dB<br>Deviation to 0°, 4 dB |
| Difference between front loudspeakers | In the range >250 Hz to 2 kHz | 0.5 dB |
| **Directivity index** | 250 Hz–16 kHz | 8 dB ±2 dB |
| **Nonlinear distortion attenuation** (SPL = 96 dB) | <100 Hz<br>>100 Hz | –30 dB (=3%)<br>–40 dB (=1%) |
| **Transient fidelity**<br>Decay time $t_s$, for reduction to a level of 1/*e*, i.e., 0.37 of output level | $t_s$[s] | <5/*f* [Hz]<br>(preferably 2.5/*f*) |
| **Time delay**<br>Difference between stereo loudspeakers | ∂t | ≤10 μs |
| **System dynamic range**<br>Maximum operating level (measurement acc. to IEC 60268, § 17.2, referred to 1 m distance) | $L_{\text{eff max}}$ | >112 dB<br>(at IEC 60268 program simulation noise or special condition) |
| **Noise level** | $L_{\text{noise}}$ | ≤10 dBA |

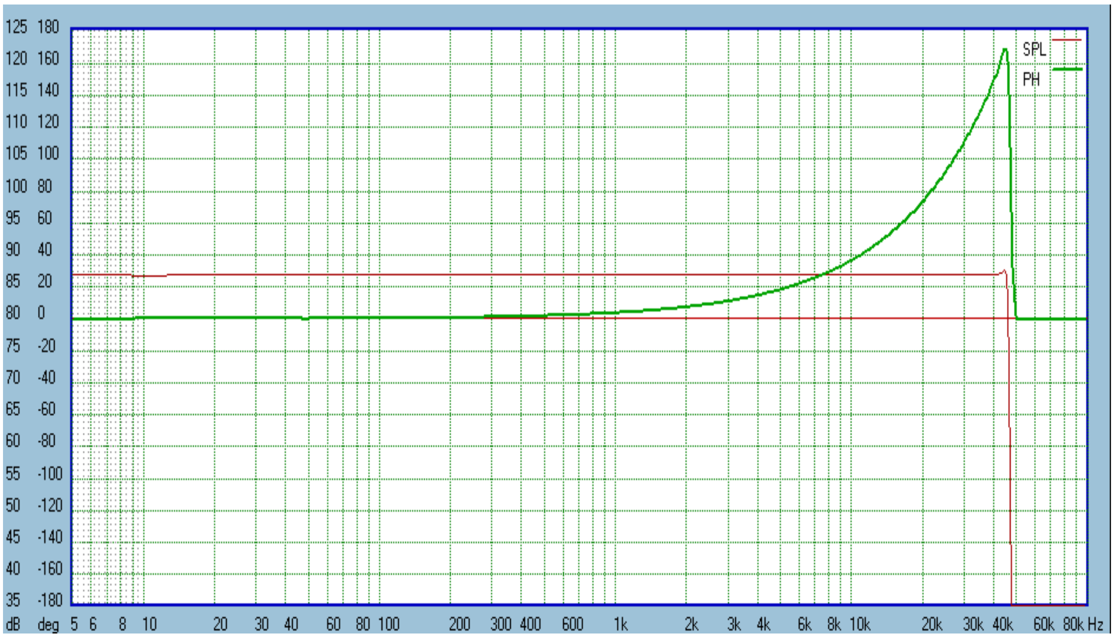



Figure 2. Thick green line shows 10usec tolerance re-calculated as differential phase between the two loudspeakers.

Ideally, in order to put things into perspective, I would have to built 10 loudspeaker systems, using standard, off-the-shelf drivers and crossover components and boxes, and then proceed to measure all the systems under identical conditions, so I would be able to compare all ten SPL and phase responses. Well, this was not going to happen, as I simply do not have resources to do so. I have therefore use simple simulation software example to understand the effect of SPL irregularity on phase compliance with the 10usec requirement.

Here are my two simulated midrange drivers, one has smoother (green) frequency response, and other is more irregular (red). These drivers share identical frequency responses between 40Hz-160Hz and 4.5kHz-18kHz, and are otherwise within +/-3dB from each other. This would not happen in real world, and the SPL differences would be visible across whole spectrum. But for the purpose of a short exercise, this may have to do.

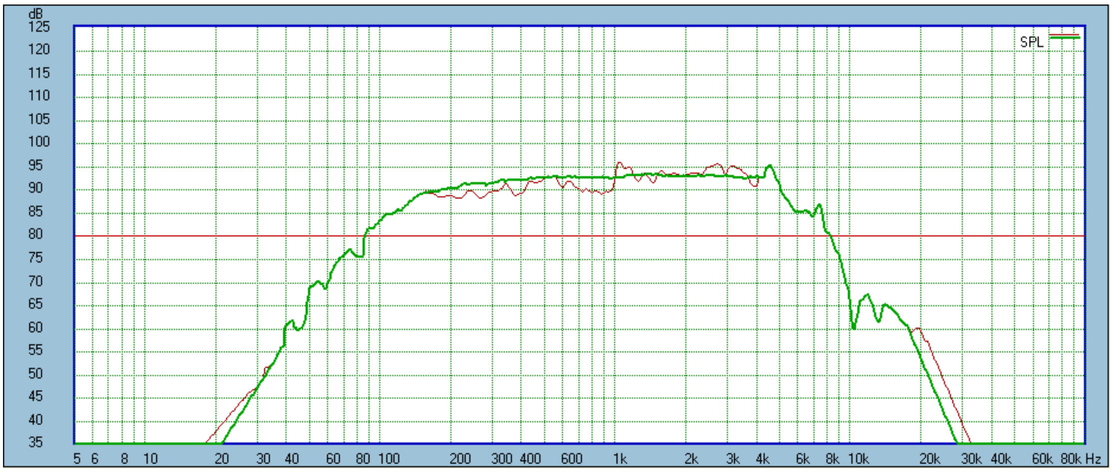



Figure 3. Two midrange drivers (red curve and green curve) with +/-3dB difference in frequency response.

Low-frequency differences in SPL would be attributed to variations in some Thiele/Small parameters, which characterize the driver in this frequency range. But these are the "lumped-element" model elements, which do not represent more modal/resonating character of the cone and other vibrating elements – these would become visible during field measurement. Anyway, suffice to say, that my simulation is simple.


Figure 4. Corresponding difference in phase response.

Figure 5. Group Delay differences 2ms are visible

Corresponding phase responses are displayed on Figure 4. Inspecting Figure 4 one would hazard several conclusions:

2. In order to comply with the AESTD1001.1.01-10, you may actually have to measure the phase response and group delay, as your customers may require this parameter documented for a statistically valid sample size.

3. Using drivers with flatter frequency response leads to much smoother phase responses. The 5dB SPL jump at 1kHz leads to 20deg phase spike on Figure 4.

4. Partially overlapping frequency responses do not guarantee partially overlapping phase responses. Since the 40Hz-160Hz SPL curves overlap exactly, you may be tempted to expect, that phase responses will overlap too. This may clearly not happen.

5. Even though the SPL responses are similar to within +/-3dB, or even partially-identical, the phase response difference fails the AESTD1001.1.01-10 in the example above.

6. Below 1kHz, the maximum allowed differential phase is 5deg, becoming very close to nil for the frequencies below 300Hz. If the set of measurements from (2.1) can not confirm the compliance, one option may be to introduce phase adjustment mechanism into your design.

Next, I attempt to examine two hypothetical subwoofer drivers. One has more jugged frequency response (green SPL) and one has more irregular frequency response ( red SPL). Still, they are within +/-3dB from each other.


Figure 6. Simulated SPL of two subwoofers

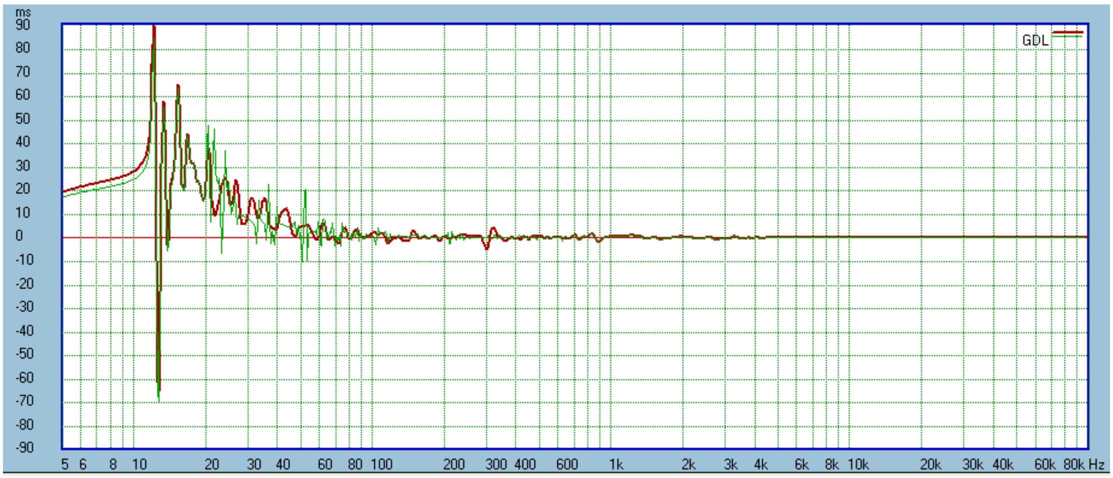



Figure 7. Corresponding GD differences – 10Hz – 100Hz. There are visible spike differences up to 20ms.

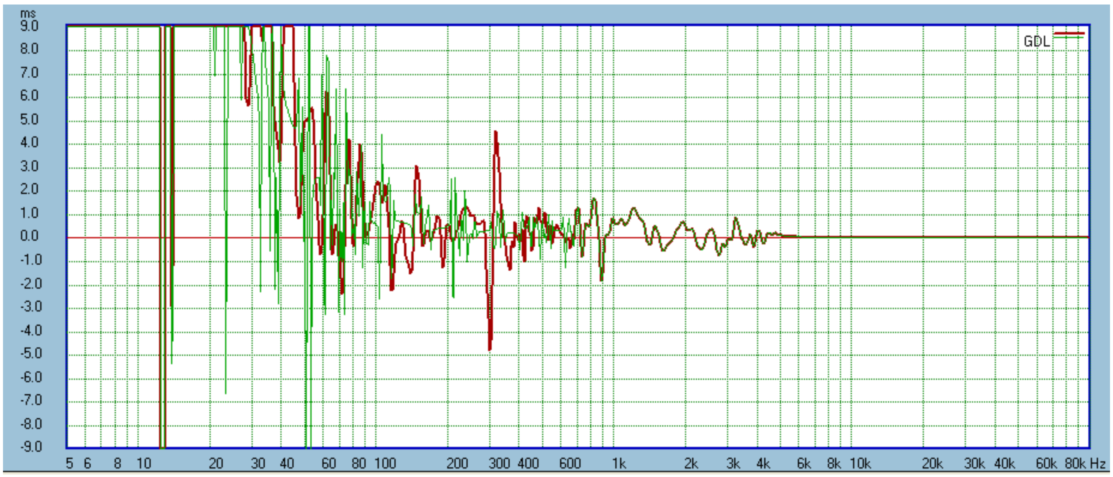
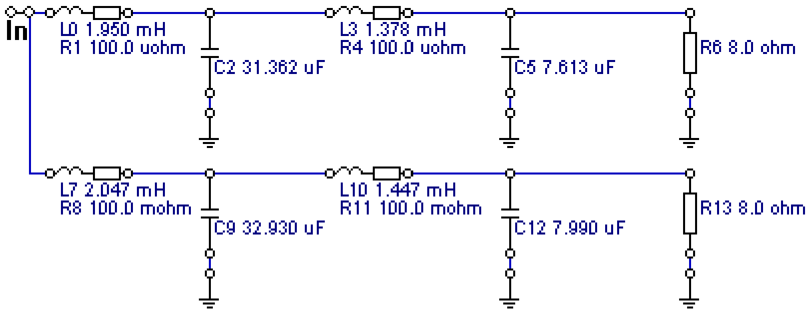



Figure 8. Corresponding GD differences – 100Hz – 1000Hz. There are visible differences up to 4.5ms.

There may be other sources of differential phase problem between free-standing loudspeaker systems. Three examples are shown below.

## Contribution of crossover components

Phase response difference between two 24dB/oct Butterworth 1kHz, LP filters.




Filter 1: L1 = 1.95mH, L2 = 1.378uH,  C1 = 31.362uF, C2 = 7.613uF
Filter 2: L1 = L1+5%, L2 = L2+5%, C1 = C1+5%, C2 = C2+5%.

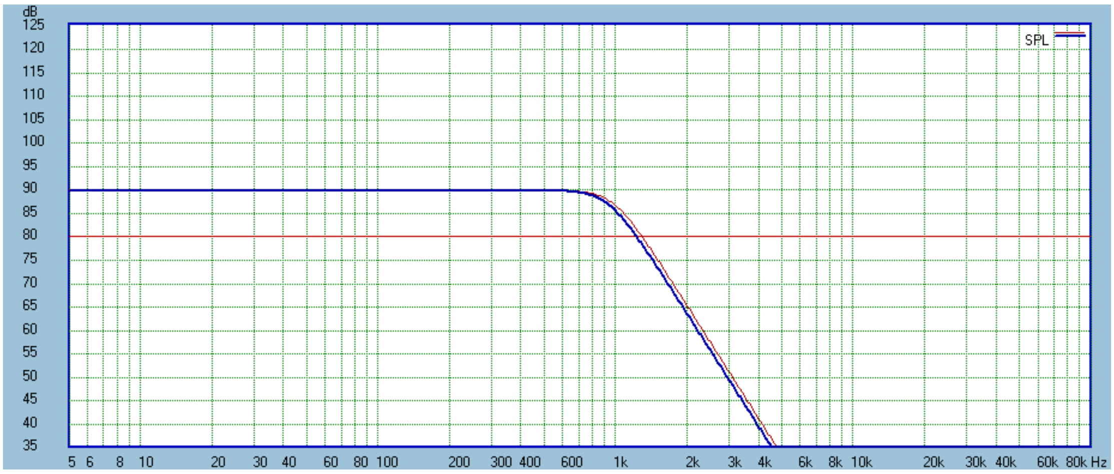



Figure 9. Frequency response of both filters.

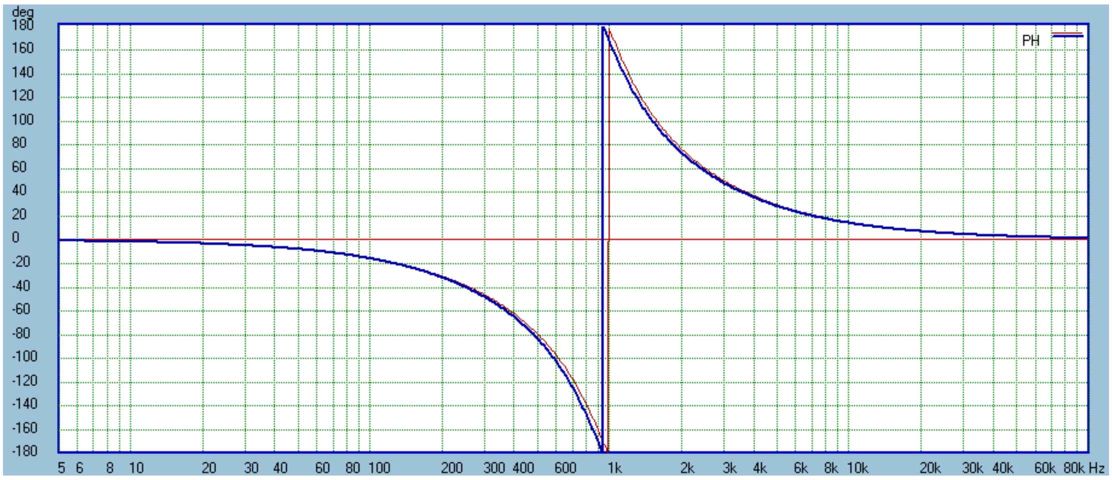



Figure 10. Phase response of both filters.

This was only a very simple circuit, but it can be concluded, that crossover's component tolerance may deteriorate the relative phase response between two loudspeakers. The LP filters with 5% tolerance components examined above, provide nearly 20deg phase shift difference in a frequency range centred at 1kHz. **The 20deg at 1kHz equates to about 55usec timing difference.**

At lower frequencies, the situation gets worse. Here is an example of -24dB/oct LP 100Hz Butterworth filter, together with it's companion filter, with all components shifted by +5%.

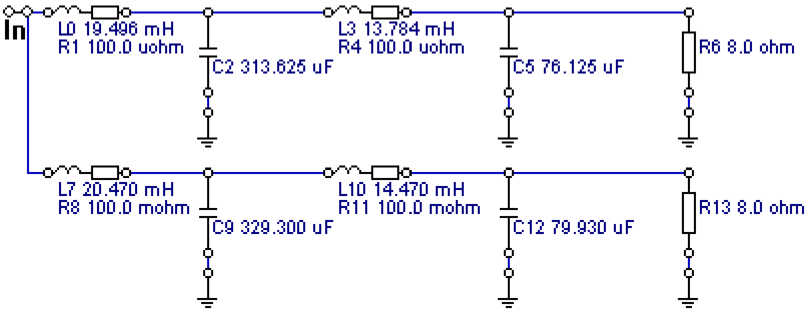



Figure 11. GD of the two filters.

At 50Hz, GD1 = 4.73ms, GD2 = 5.06ms.

The difference between group delays is now: GD2 – GD1 = 5.060 – 4.730 = **330usec**.

## **Contribution of out-of-synch clocks in digital ADC devices**.

Source: http://en.wikipedia.org/wiki/Audio_system_measurements

"… Sample accuracy/synchronization

> Not as much a specification as an ability. Since independent digital audio devices are each run by their own crystal oscillator, and no two crystals are exactly the same, the sample rate will be slightly different. This will cause the devices to drift apart over time. The effects of this can vary. If one digital device is used to monitor another digital device, this will cause dropouts or

distortion in the audio, as one device will be producing more or less data than the other per unit time. If two independent devices record at the same time, one will lag the other more and more over time. This effect can be circumvented with a wordclock synchronization. It can also be corrected in the digital domain using a drift correction algorithm. Such an algorithm compares the relative rates of two or more devices and drops or adds samples from the streams of any devices that drift too far from the master device. Sample rate will also vary slightly over time, as crystals change in temperature, etc. See also clock recovery……"

Example below shows the recommended limit of phase difference between stereo loudspeakers (green curve) and the phase difference between two digital devices sampling at 48kHz, and differing by one sample of **20.38usec** (blue curve). It is observable in this example, that the "out-of-synch" condition alone, will result in failing of the AESTD1001.1.01-10 recommendations.

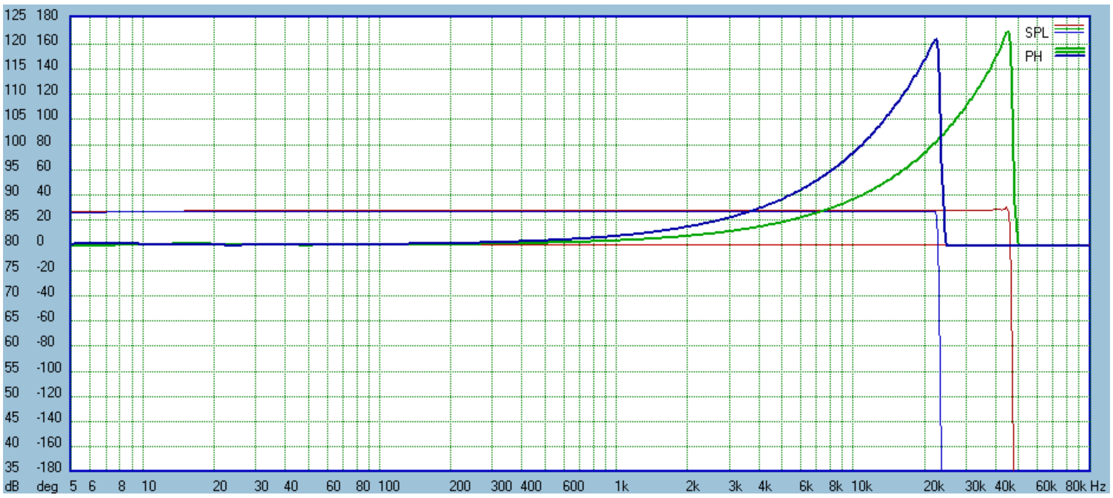



Figure 12. Sampling error (blue) of two ADC devices sampling with 48kHz.

If the digital devices are mot synchronized, their clocks will drift apart in time and create a beat-pattern between them. The frequency of the beat-pattern will depend on the difference in frequencies of their clocks. The phantom image, created by stereo system using unsynchronized clocks will oscillate in the lateral plane accordingly.

An example of a digital loudspeaker system with synchronization is Meridian DSP8000.

http://www.meridian-audio.com/download/Handbooks/Loudspeakers/DSP8000%20\User%20Guide.pdf

There is one "master loudspeaker" or "controller" – typically the centre channel and the rest are "slave loudspeakers". It is understood, this is related to the necessity of single clock source running the whole system.

### Variability of TS parameters

Variability of TS parameters is a well known issue. I have simulated several vented box SPL frequency responses, while varying some of the TS parameters within 10% of their original value.

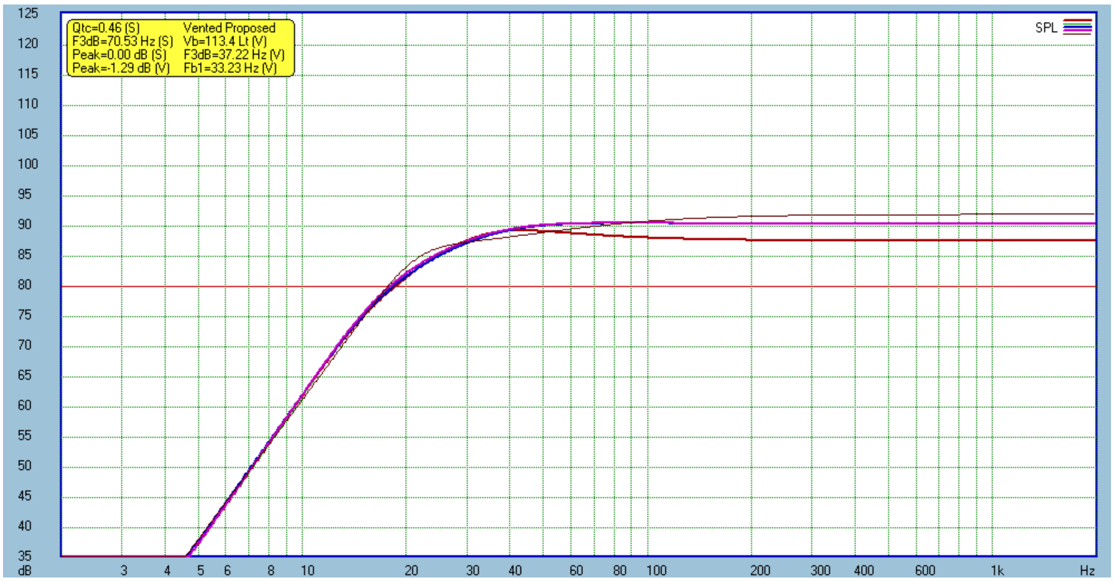



Figure 13. SPL curves of a typical woofer – varying TS parameters.


Figure 14. Corresponding GD for the woofer above.

The above example shows, that group delay will experience changes as high as 2ms within the operating 3dB bandwidth of the loudspeaker – in the case above, it is 30Hz and above.

## 3. Transient Fidelity

Table 3. Suggestions for reference monitor loudspeakers and advice for home loudspeakers.

| *Parameters* | *Units/Conditions* | *Value* |
|---|---|---|
| **Amplitude/frequency response** | 40 Hz–16 kHz<br>0°<br>±10°<br>Horizontal ±30° | <br>Tolerance 4 dB<br>Deviation to 0°, 3 dB<br>Deviation to 0°, 4 dB |
| Difference between front loudspeakers | In the range<br>>250 Hz to 2 kHz | 0.5 dB |
| **Directivity index** | 250 Hz–16 kHz | 8 dB ±2 dB |
| **Nonlinear distortion attenuation**<br>(SPL = 96 dB) | <100 Hz<br>>100 Hz | –30 dB (=3%)<br>–40 dB (=1%) |
| **Transient fidelity**<br>Decay time $t_s$, for reduction to a level of 1/*e*, i.e., 0.37 of output level | $t_s$[s] | <5/*f* [Hz]<br>(preferably 2.5/*f*) |
| **Time delay**<br>Difference between stereo loudspeakers | ∂t | ≤10 µs |
| **System dynamic range**<br>Maximum operating level<br>(measurement acc. to IEC 60268, § 17.2, referred to 1 m distance) | $L_{\text{eff max}}$ | >112 dB<br>(at IEC 60268 program simulation noise or special condition) |
| **Noise level** | $L_{\text{noise}}$ | ≤10 dBA |

## Examples

The following example comes from a paper "On the Use of a Non-Enviromental Control Room As a 5.1 Surround Listening Room" by Torres-Guijarro, Pena and Sobreire-Seoane. Telecomunicacion Universdale de Vigo, Spain.

http://www.lomg.net/media/noticias/id/artigos/19th%20International%20Congress%20on%20Acoustics%202007%20-%20On%20the%20use%20of%20a%20non-environmental%20control%20room%20as%20a%205.1%20sorround%20listening%20room.pdf

Their measurement shows basically, that the speakers are almost compliant with ( 5/f) transient fidelity. The authors discuss room influence on the results and conclude, that more accurate measurements need to be performed – see Figure below

For a sake of an exercise I took a liberty of adding ( 2.5/f ) limiting reference, as preferred by the AESTD1001.1.01-10. This loudspeaker would clearly fail this new limits.

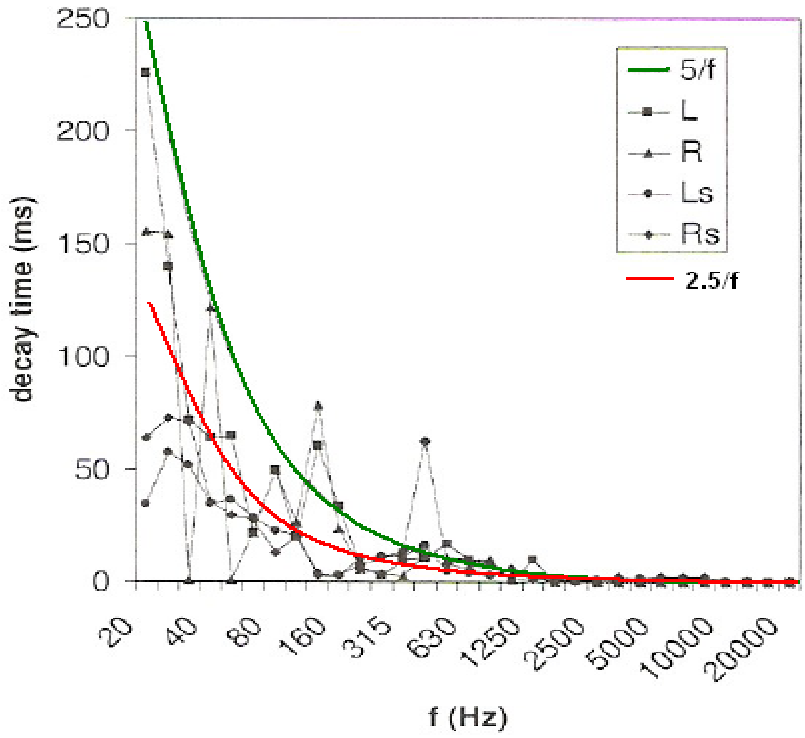



Figure 3: transient fidelity

In the next paper: "Correlation of Transient Measurements on Loudspeakers **w**ith Listening Tests" by M. Corrington, published in JAES, JANUARY 1955, VOLUME 3, NUMBER 1, we find an interesting measurement method, allowing for separation of the "overhang transient" – see below

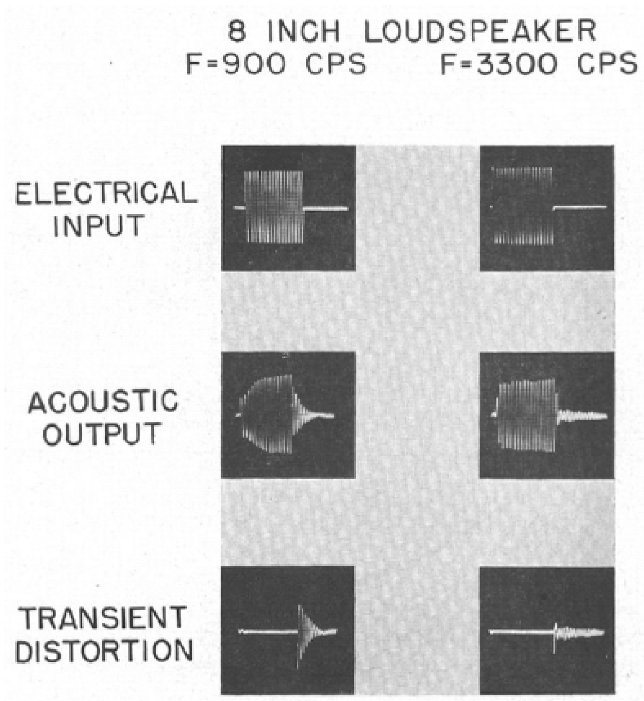



FIG. 3. Transient distortion for 8-in. loudspeaker.

The paper reads well, and has the following interesting conclusion:

"….This information supplements the steady-state sound pressure measurements. We have never found any system with low transient distortion that did not also have a smooth sound-pressure curve; on the other hand, we have measured systems with fairly sharp and small peaks in the sound-pressure response that produced objectionable transient distortion.

**There is very good correlation between transient distortion and subjective listening tests. Whenever there are peaks in the transient distortion, one can be sure that the listening tests will reveal unpleasant distortion, even though the sound-pressure curve is quite smooth….**

Extensive measurements show that for a high-quality audio system the sound-pressure curve must be smooth and properly shaped, and that the transient distortion should be down at least 18dB throughout the range. One can then be fairly certain that the system will pass very careful listening tests…."

### Measurement examples

The ( 5/f ) limit may appear fairly relaxed, as it translates into 5ms decay for 1000Hz tone, for example. Not everybody is in agreement with this limit level. Martin Colloms in his book "High Performance Loudspeakers, 4th edition", makes a comment:

"…Speech and music are largely asymmetric in waveform structure, and theoretically they demand an accurate pulse or transient response from the loudspeaker. If the standards often applied to the other components in the audio chain were applied to loudspeakers, then the acoustic output of a loudspeaker should decay to negligible levels 0.025ms after the cessation of a narrow impulse excitation….."

Another audio luminary, Sigfried Linkwitz, has also commented on his website www.linkwitzlab.com:

"…4 - Transient fidelity

This is a very peculiar looking spec. 5 ms decay time at 1 kHz is at least by an order of magnitude too large, before it even might begin to have some meaning…."

The recommended transient fidelity is for the acoustic output from a loudspeaker to drop to 1/e = 0.37 = -8.64dB within 5/f (or preferably 2.5/f ) after the signal was switched off. So, for instance the sound pressure would have to drop by -8.64dB within 2.5ms for 1000Hz tone, if the preferable limit was applied.

Anyway, the method of applying a large number of sine-wave tone bursts is rather laborious. It is therefore no surprise, that other, faster substitute methods are being examined. One of them is the well-known Cumulative Spectral Decay (CSD) method.

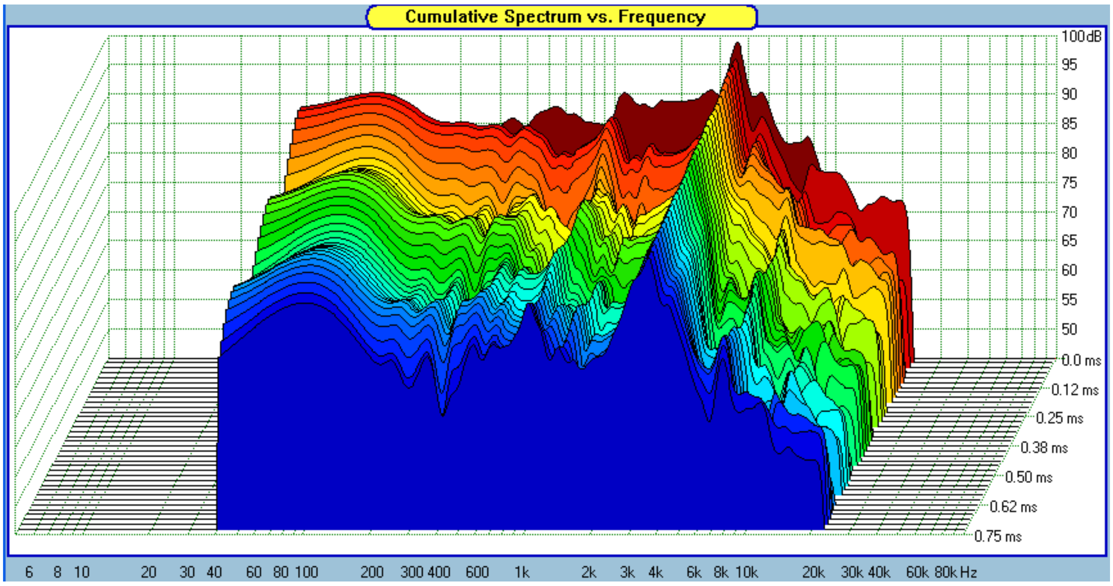



Figure 15. 12” guitar loudspeaker.

The loudspeaker shown above (12”, high efficiency guitar speaker) would fail the transient fidelity test at 3.5kHz. The requirement is for the output to drop by -8.64dB by 0.714ms. This can be easily demonstrated on the picture below.

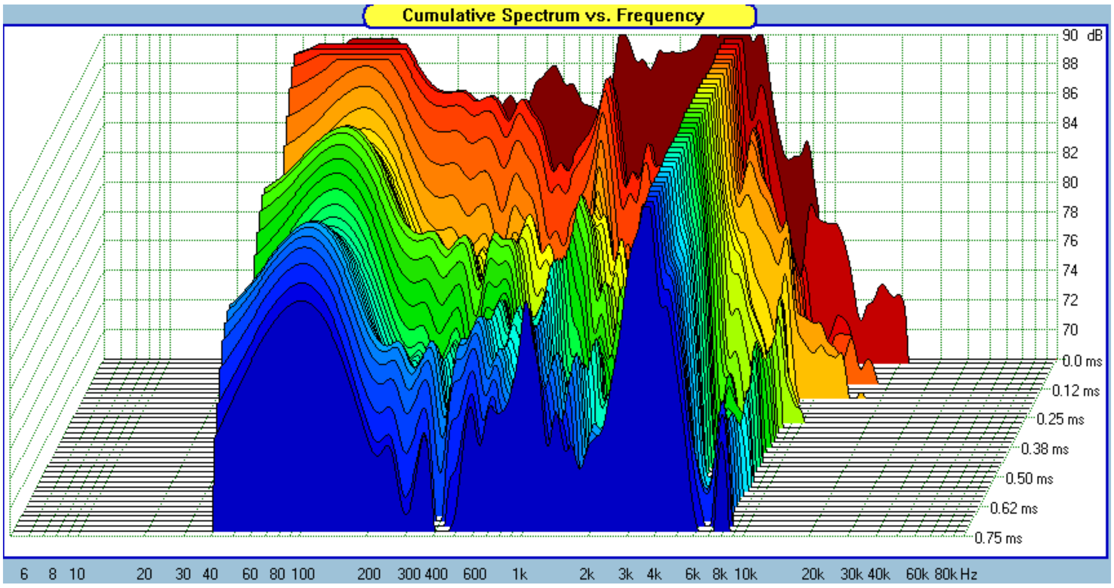



Figure 15.b. The same woofer, scaled for transient fidelity inspection.

We can observe, that at 3.5kHz the SPL of this loudspeaker is about 98.5dB. Given the 8.6dB required drop in output, I can scale the display plot such way, that the top of the screen is now 90dB ( this comes from 98.5dB – 8.6dB = (appr)90dB ). I can take an advantage of the software clipping feature, which will clip SPL above the top legend of 90dB, regardless of the timing trace.

Now, all I have to do, is to look for a “flat-top” SPL curves. Such curves must have exceeded the allowed transient fidelity level.

Indeed, the “flat-topped” SPL curves are seen right down to 0.75ms at 3.5kHz – clear failure.

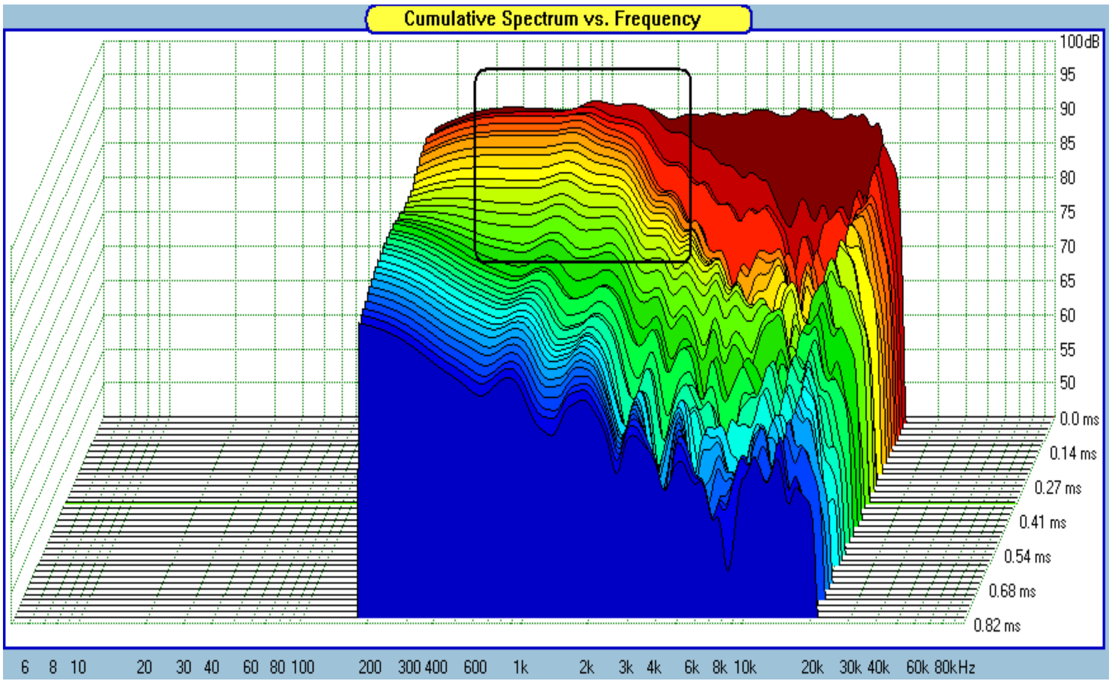


Figure 16 Waveguide 2-way loudspeaker.

The next loudspeaker shown above is actually somewhat bizarre. In the highlighted area ( black rectangle) the first immediate CSD waterfalls seem to have quite high amplitude. I am not sure what is responsible for this problem. The loudspeaker is a waveguide type, and seems to produce a lot of output around 600-2kHz range. The ridge around 17kHz is also quite high, and the result is a border-line case.

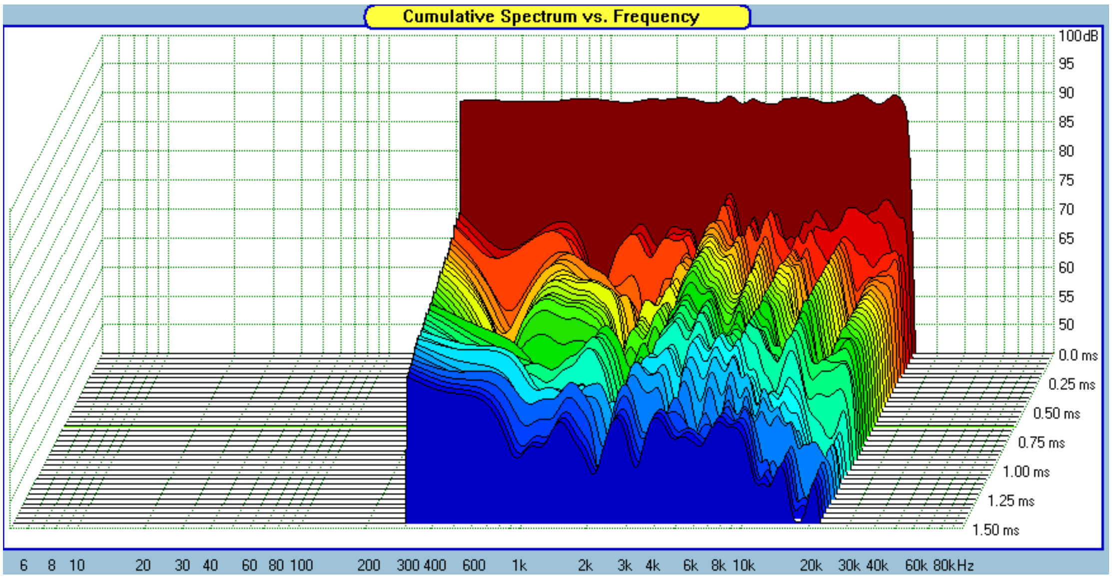


Figure 17. The two-way loudspeaker (12dB/oct LR @ 2kHz ) with DSP.

The two-way loudspeaker (12dB/oct LR @ 2kHz ) system shown above would easily pass even the (2.5/f) transient fidelity test right up to the 20kHz limit. The (2.5/f) requirement calls for -8.64dB drop immediately at 0.125ms. As you can see on the CSD above, the waterfall drops by around -20dB immediately. This loudspeaker system was measured in-room, using windowed MLS technique. This outstanding result is due to DSP equalization applied to amplitude and phase, and would even satisfy S. Linkwitz's idea of transient fidelity.

For this loudspeaker, I can only repeat findings provided by M. Corrington:

"…Extensive measurements show that for a high-quality audio system the sound-pressure curve must be smooth and properly shaped, and that the transient distortion should be down at least 18dB throughout the range. One can then be fairly certain that the system will pass very careful listening tests…."

## Conclusions

Intuitively, better loudspeaker symmetry (phase, level and frequency response matching), should improve perceived localisation of the phantom source in terms of better source accuracy and stability, and less smearing or blur. All this will manifest itself before the loudspeaker-room interaction comes into play.

Many years ago, Boston Audio Society (BAS) has developed interesting observations about time-corrected loudspeakers.

Source: http://www.bostonaudiosociety.org/bas_speaker.htm
http://www.bostonaudiosociety.org/pdf/bass/

"….If the stereo loudspeakers differ in their time-shift behaviour by more than about thirty millionths of a second (or a finer tolerance, perhaps, for critical listeners), the stereo image will be perceptibly smeared. **The two speakers must "speak" together at all frequencies if the subtlest details in the stereo field are to be preserved.**

This, quite simply, may be the principal advantage to be gained from "linear-phase" or "time-corrected" loudspeakers. The manufacturers who are striving to reduce the time dispersion of loudspeakers to zero may also be ensuring that there will be no significant differences in signal propagation timing between the two speakers in a stereo pair. The delicate timing information in a stereo recording is thus accurately retained and is transmitted to the listener unaltered…"

This was quite a while ago.

More recently (in 2002), the AESTD1001.1.01-10 drove the stake in the ground, and pegged the 10usec as the maximum allowed timing difference between stereo loudspeakers across the entire audio band.

Examples shown above shed some light on the contributing sources of loudspeaker's disparity in the time domain. Some of them contribute less and some contribute more, but all conspire to built a time domain error between the loudspeakers that can easily exceed the allowed 10usec (or corresponding phase difference).

If these factors are left unchecked or unaccounted for, the acoustic image generated by the pair of loudspeakers may be distorted, blurred and more difficult to localize.

Transient fidelity test, if performed using sine-wave bursts is a cumbersome test, but highly recommended. The substitute CSD method will allow you to at least evaluate the “resonating overhangs”, and pinpoint design deficiencies as well. The CSD method shows all decaying resonances at all frequencies at once, and is developed from the Impulse Response, so it’s a natural extension of loudspeaker characterization set of parameters, recommended by the AESTD1001.1.01-10

Thank you for reading.

Bohdan