Some Attributes of Linear-Phase Loudspeakers

Compiled by Bohdan Raczynski, January 2014

Introduction

Source: http://www.kfs.oeaw.ac.at/content/blogcategory/0/378/

"…Early psychoacoustic research suggested that the human auditory system is insensitive to differences in the relative phases of spectral components of a multi-component sound. However, research from the last two decennia provides evidence that listeners can detect phase differences between the stimulus components that interact within a single auditory filter. The most impressive demonstration of phase sensitivity is given by the masker-phase effect, i.e. the more than 20-dB variation in masking effect caused by a harmonic complex when varying the phase relations between its components. This masking paradigm is widely used to obtain a psychoacoustical measure of the phase response of the cochlea.…"

I must admit, I did not know about the above research and it's results. I have been researching the internet for about a year before I came across the above, simple information. There is a lot more too. Now, I realised, there is a volume of research results that clearly indicates, that rather than asking "is phase distortion audible?" we should now be asking question "how does the phase distortion manifest itself?".

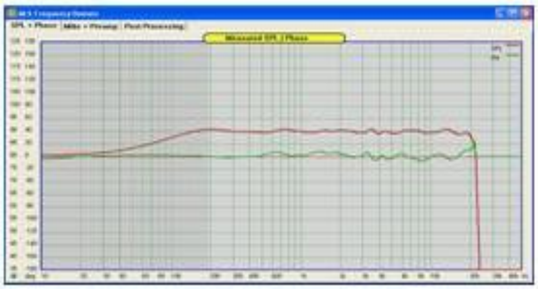
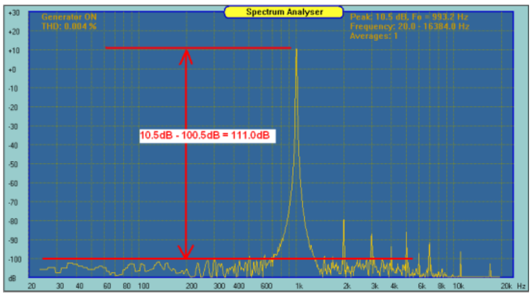
Naively, and without any prior experience in how should I actually do it, I conducted my own listening tests by comparing the sound from traditional, minimum-phase loudspeakers to the sound of linear-phase loudspeakers. I am talking here about acoustically linear-phase loudspeaker. During my short, initial listening tests on linear-phase loudspeakers, I was surprised by how indifferent the linear-phase mode was to my ear.

Was I doing the right thing then?. This result definitely required further investigation on much more diverse listening material.

Listening Habits

Traditionally, when I listened to the quality of the sound reproduced by my audio playback equipment, I focus on tonal balance (frequency response), dynamics of the sound (SNR), residual noise floor (inaudible), distortion (inaudible).

Interestingly, all of the above characteristics can be assessed and visualized in frequency domain. It was simply the easiest way to listen to the sound and evaluate what I was hearing, but I now realize, that I was only considering the steady-state analysis in the frequency domain – see pictures below.

Frequency response, distortion, dynamics and noise floor – all in frequency domain.

I was doing the same type of analysis over, and over again for years, and grew accustomed to this ritual. It was easy to compare with measured results, so it felt comfortable, that I can correlate my measurements with what I can easily hear (or can not hear).

Recently, things have changed for me. I came across a simple paper, http://www.audiophilerecordingstrust.org.uk/articles/speaker_science.pdf which inspired me to take a more comprehensive look at my listening tests. Having read the paper, I re-examined information from other internet re-sources, and as a result I came to the conclusion, that my listening tests were only a starting point of what I should have listened to when examining linear-phase loudspeakers.

To put it simply – I needed to significantly extended the evaluation of time-domain characteristics of the loudspeaker in my listening habits.

In the brief conclusions of my short, initial listening tests presented in http://www.bodziosoftware.com.au/Home_Theatre_Conclusions.pdf I have pointed out one perceptible difference – I felt closer to the stage/musicians. This was more of an accidental and unexpected impression, to which I did not pay much attention. But this indeed relates to time-domain characteristics of a loudspeaker, rather than frequency domain.

Yes, it appears, that I have been covering only half of what I should have been paying attention to. And the paper mentioned above made it startlingly clear to me.

New Listening Habits

The remaining part of this paper is my crude attempt to summarise audible attributes of linear-phase loudspeakers. This is what you need to listen for when evaluating linear-phase loudspeakers. I do not pretend, that the list is complete, but it's a start. It clearly points to the time-domain characteristics of the loudspeaker, and this is something, which may of us (till recently, including myself) are not accustomed to. I simply did not know what to listen for.

Below, I present the "nominated attribute(s)", showing the source, followed by a short description from the source.

1. Tighter bass

2. Wider and deeper sound stage (quite dramatic)

Source: http://redspade-audio.blogspot.com.au/2012/03/bathurst-2011-audio-event-of-year.html

DEQX demo

"…A highlight this year was a demo of the capabilities of DEQX. This came about from discussions of my active crossover listening comparisons, in which a small group could not hear any improvement with DEQX. Terry argued that we had dumbed down the DEQX and prevented it from showing what it can do. This is certainly true, we wanted to test sound quality only and in that regard found no reason to spend the extra compared to cheaper options. However, Terry set up a demo in which two profiles were created on DEQX. One was limited to the processing power of MiniDSP and DCX. The other allowed DEQX to strut its stuff. In particular, it was allowed to correct for phase and group delay. We then blind tested this with instant switching, not knowing what was being heard. I was the first to sit in the chair and do the demo and quite soon I didn't need to be told which was which, because the difference was obvious.

Changes noticed with DEQX:

much tighter bass
wider and deeper sound stage (quite dramatic)

Both had a basic level of time alignment with digital delays. Both were matched in level and in response closely. These differences were related to the group delay correction. **Without it, the sound was flat and almost lifeless in comparison**.

I then watched as others sat through the demo, each person noticing the same differences, differing only in the amount of time taken before declaring what they heard….."

Personally, I can testify to the tighter bass audible during linear-phase mode. I operate large, 18"/vented enclosure subwoofers, tuned to 20Hz. Playing impulsive sounds, in minimum-phase mode, the subs overshot and then add and prolong the ringing - past steep, impulse-like signals. This unwanted flabbiness is unfortunately audible in minimum-phase mode on low-frequency impulsive signals.
http://www.bodziosoftware.com.au/LP_MP_Subwoofer_Tests.pdf

However, in linear-phase mode, the punch is still deep, but tight, without the "aftersounds".

3. Realism

Source: http://www.audiophilerecordingstrust.org.uk/articles/speaker_science.pdf
(This is a must-read article in it's entirety)

".....Another area in which loudspeakers are disreputable is in the neglect of the time domain. The traditional view is that all that matters is to be able to reproduce continuous sine waves over the range of human hearing.

A very small amount of research and thought will reveal that this is a misguided view. Frequency response is important, but not so important that the attainment of an ideal response should be to the detriment of **realism**. One tires of hearing that "phase doesn't matter" in audio or "the ear is phase deaf". These are outmoded views which were reached long ago in flawed experiments and which are at variance with the results of recent psychoacoustic research.

The ear works in two distinct ways, which it moves between in order to obtain the best outcome from the fundamental limits due to the Heisenberg inequality. The Heisenberg inequality states that as frequency resolution goes up, time resolution goes down and vice versa. Real sounds are not continuous, but contain starting transients. During such transients, the ear works in the time domain. **Before the listener is conscious of a sound, the time domain analysis has compared the time of arrival of the transient at the two ears and established the direction.** Following the production of a transient pressure step by a real sound source, the sound pressure must equalise back to ambient.

The rate at which this happens is a function of the physical size of the source. The ear, again acting in the time domain, can measure the relaxation time and assess the size of the source. Thus before any sound is perceived, the mental model has been told of the location and size of a sound source.

In fact this was the first use of hearing, as a means of perceiving a threat in order to survive. Frequency analysis in hearing, consistent with the evolution of speech and music came much later. After the analysis of the initial transient, the ear switches over to working in the frequency domain in order to analyses timbre. In this mode, the mode that will be used on steady state signals, phase is not very important. However, the recognition of the **initial transient and the relaxation time are critical for realism**. Anything in a sound reproduction system which corrupts the initial transient is detrimental.

Whilst audio electronics can accurately handle transients, the traditional loudspeaker destroys both the transient and the relaxation time measurement. Lack of attention to the time domain in crossover networks leads to loudspeakers which reproduce a single input step as a series of steps, one for each drive unit at different times..."

4. Depth

5. Resolution

6. Separation of ambience

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

Boston Audio Society has an interesting view on time-corrected loudspeakers.

"….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…"

They also point to some of the advantages of such loudspeakers:

1. Depth.
This may surprise some listeners when they first hear it, since many speakers (and records) elicit only a general left-to-right spread. But "stereo", as originally conceived, implied a three-dimensional sound in which voices or instruments could be localized at different apparent distances from the listener as well as at various lateral positions. Listeners to time-aligned speakers consistently report hearing a stereo image with unusual depth.

2. Resolution.
The stereo image is reproduced precisely, each voice or instrument having its proper place and width. In complex sound sources such as symphony orchestra, individual instruments can be resolved with unexpected clarity. In the old cliche, "I hear details I never knew were in the recording. " Some listeners have incorrectly attributed the improved resolution of detail to more accurate transient response, but the better definition of details is simply the result of the reduction of blending in the stereo image.

3.Separation of ambience.
With loudspeakers whose stereo image is slightly blended because of time-smear, any hall ambience or reverberation in the recording tends to become slightly mixed with the instrumental sounds, causing coloration of those sounds. Consequently, with such speakers closely-microphoned recordings tend to sound better because of their distinctly defined sound. But with time-corrected loudspeakers, the ambience is resolved as a separate sound, and larger amounts of hall ambience in recordings can be enjoyed……"

7. Inter-channel accuracy of sound reproduction.

Source: http://www.cirrus.com/en/pubs/whitePaper/DS668WP1.pdf

"…….5. Audibility of Phase Distortion

One of the confusing issues regarding the audibility of phase is that the discussion is generally considered to be a single topic when in reality should be discussed as two distinct situations. The audibility of phase distortion must be evaluated as follows:

1) Inter-channel phase distortion. Characterized as differences in phase response between two or more channels.

2) Intra-channel phase distortion. Characterized by non-linear phase response within a channel with the stipulation that the phase response is matched between all channels within the system (i.e. inter-channel phase distortion is equal to 0 msec)

6. Inter-Channel Phase Distortion

We use the amplitude and phase relationship between the sounds received by our ears to localize the source of the sound. Modern audio systems use this attribute to create what is known as imaging, or the perception that an instrument or vocal is coming from a location that is different than the actual speaker location. The audible effects of inter-channel phase distortion can be easily demonstrated by simply reversing the speaker connections on one channel of an otherwise properly configured stereo system. The loss of imaging is immediately noticeable even to those without a trained ear. Granted this test is rather dramatic and 180 degrees of inter-channel phase distortion is not indicative of standard operation but it does demonstrate the potential effects. As a result of this test, you would be hard pressed to find someone that would argue that 180 degrees of inter-channel phase distortion is acceptable, but where between the two extremes is the threshold of audibility? **Tom Holman reports [10] that in his laboratory environment at the University of Southern California that is dominated by direct sound, a channel-to-channel time offset equal to one sample period at 48 kHz is audible. This equates to 20 µsec of inter-channel phase distortion across the entire audio band. Holman [10] also mentions, "one just noticeable difference in image shift between left and right ear inputs is 10 µsec**".

7. Intra-Channel Phase Distortion

Recall that we use the differences in signal amplitude and phase to localize or determine the source of sound and relatively small amounts of inter-channel phase distortion can be audible. But how does our hearing react when each channel in a multi-channel system is subjected to non-linear phase response but the phase response is matched between all channels? Douglas Preis [11] did an extensive survey of existing literature and Tom Holman's [10] experiences and research through his work at USC gives us an interesting insight into this phenomenon. Both report that the threshold of audibility is frequency dependent, which correlates with all other audibility thresholds. In laboratory environments when using test tones and headphones, research has shown that the human ear is sensitive to intra-channel phase differences of 0.25 msec [8] or +/-0.5 msec [9] in the mid-range with the threshold increasing at higher and lower frequencies. Preis states "the tolerances shown.... are not directly applicable to speech or music signals irradiated by loudspeakers in a reverberant environment. Most likely, the perceptual thresholds for these conditions would be at more than twice those shown". Essentially, the data suggests that for high quality music or speech reproduction in a reverberant environment intra-channel phase distortion of 1 msec is inaudible to a trained listener. Notice that this threshold is a relatively conservative statement and is still two orders of magnitude greater than that for inter-channel phase distortion!....."

8. Precedence effect or "law of the first wavefront"

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

"…..The **precedence effect** or **law of the first wavefront** is a binaural psychoacoustic effect. When a sound is followed by another sound separated by a sufficiently short time delay (below the listener's echo threshold), listeners perceive a single fused auditory image; its perceived spatial location is **dominated by the location of the first-arriving sound** (the first wave front). The lagging sound also affects the perceived location. However, its effect is suppressed by the first-arriving sound…..

The precedence effect appears, if the subsequent wave fronts arrive between 2 ms and about 50 ms later than the first wave front.

The precedence effect is important for the hearing in enclosed rooms. With the help of this effect it remains possible to determine the direction of a sound source (e.g. the direction of a speaker) even in the presence of wall reflections…."

9. Importance of Phase in Transients

Source: http://sound.media.mit.edu/Papers/kdm-phdthesis.pdf

Page 44

"….Since Helmholtz, there has been a figurative tug-of-war between proponents of his "spectral theory" of musical sound and researchers who recognized the importance of sound's temporal properties. *Analysis-by-synthesis* research, by trying to discover methods for synthesizing realistic sounds, has revealed several critical limitations of purely spectral theories. Clark demonstrated that recordings played in reverse—which have the same magnitude spectra as their normal counterparts—make sound-source identification very difficult. Synthesis based on Fourier spectra, with no account of phase, does not produce realistic sounds, in part because the onset properties of the sound are not captured (Clark et al., 1963). Although most musical instruments produce spectra that are nearly harmonic—that is, the frequencies of their components (measured in small time windows) are accurately modeled by integer multiples of a fundamental—deviations from strict harmonicity are critical to the sounds produced by some instruments. For example, components of piano tones below middle-C (261

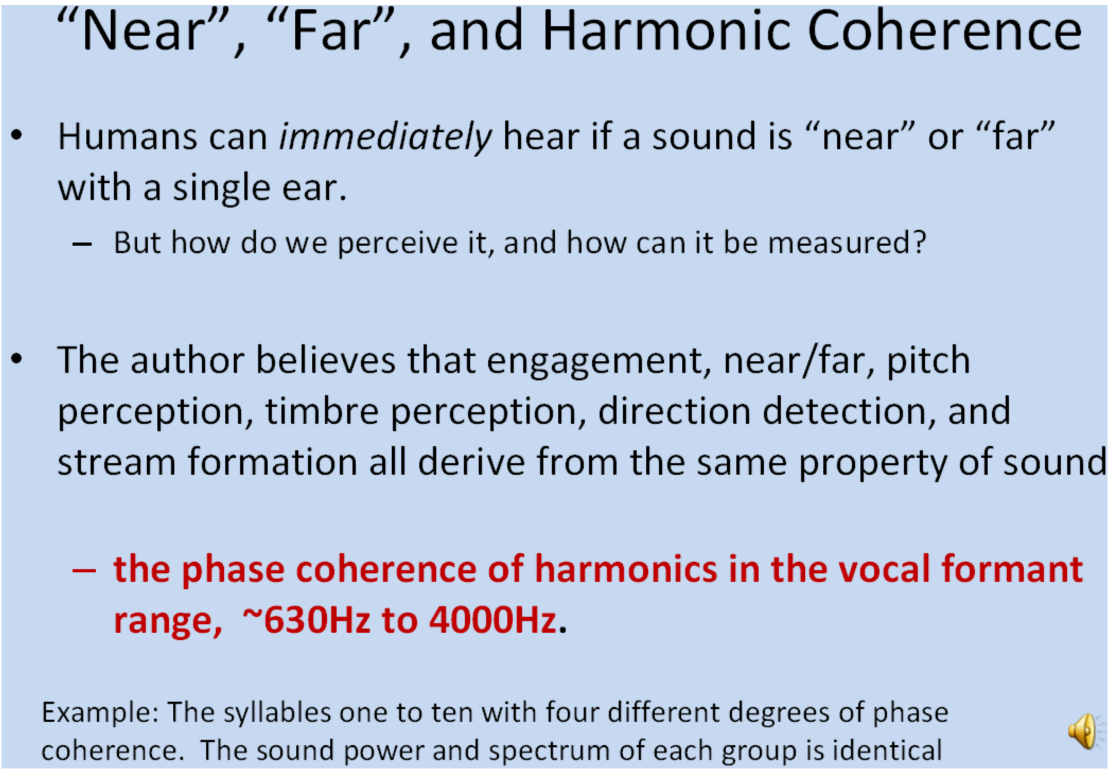
Hz) must be inharmonic to sound piano-like (Fletcher et al., 1962). In fact, all freely vibrating strings (e.g., plucked, struck, or released from bowing) and bells produce inharmonic spectra, and inharmonicity is important to the attack of many instrument sounds (Freedman, 1967; Grey & Moorer, 1977). Without erratic frequency behavior during a note's attack, synthesized pianos sound as if they have hammers made of putty (Moorer & Grey, 1977).

So Helmholtz's theory is correct as far as it goes: the relative phases of the components of a purely periodic sound matter little to perception. However, as soon as musical tone varies over time — for example, by turning on or off — temporal properties become relevant. In the real world, there are no purely periodic sounds, and an instrument's magnitude spectrum is but one of its facets….."

10. Pitch, Timbre, and Source Separation

Source: David Greisner http://www.davidgriesinger.com

http://www.davidgriesinger.com/Acoustics_Today/Pitch,%20Timbre,%20Source%20Separation_talk_web_sound_3.pptx



11. Confirmation of two-stage processing by the ear, as discussed in (3).

Source: http://www.hauptmikrofon.de/theile/ON_THE_LOCALISATION_english.pdf

4.3.1 The "law of the first localisation stimulus"

"….**For a conventional stereo-up, a phantom source shifts from ϕ = 0° to ϕ = 30° if the time difference between two broadband loudspeaker signals is increased from zero to about 600 µs**. The association model could explain this phenomenon (time- as well as level-based stereophony) by means of psychoacoustic principles of the gestalt association stage. The localisation stimulus arriving at the gestalt association stage first has a greater weight compared to the second stimulus (the equivalent for level based stereophony would be the localisation stimulus with the higher level). Despite their identity and relative time delay, the localisation stimuli can be discriminated, since each of them is present in the binaural correlation pattern in a complete and discriminable form (see Section 4.1).

Yet, a further increase in the inter-channel time difference leads to an exceedance of the maximal time delay τmax. For stationary broadband signals (continuous noise), this causes a disruption of the localisation stimulus selection, which manifests itself in the form of a reduced suppression of the comb filter effect, for example. In this particular sound field constellation, the law of the first wavefront cannot be observed in accordance with the association model. Analysable wavefronts that would allow for a localisation stimulus selection of the impinging sound components do not exist.

In contrast, for non-stationary impulsive signals (clicks, speech, impulsive tones) an increase in the inter-channel time difference has a different effect. In the association model, evaluation of the amplitude envelope ensures that the primary and the delayed sound (reflection) can be discriminated as localisation stimuli. According to a hypothetical function of the gestalt association stage, the primary localisation stimulus determines the auditory event. It does this even more so the larger the time difference between the arriving localisation stimuli gets. Only when a time difference of about 10 … 30 ms is exceeded will the subsequent localisation stimulus gain in perceptual weight. Beyond the echo threshold (for a definition see BLAUERT 1974), it will be perceived as a separate auditory event.

It appears that the "law of the first wavefront" can be interpreted as the **"law of the first localisation stimulus"**….."

"…..6. Summary
According to the association model presented in the preceding chapters, the functioning of the auditory system with respect to spatial hearing is due to two different processing mechanisms. Each of these two processing mechanisms manifests itself in the form of an associatively guided pattern selection.

A current stimulus stemming from a sufficiently broadband sound source gives rise to a location association in the first and to a gestalt association in the second, higher-level processing stage because of auditory experience. Although the two stages work independently of each other, they always determine the properties of

one or multiple simultaneous auditory events in a conjoint manner. The rigorous differentiation of these two stimulus evaluation stages corresponds entirely to the two elementary areas of auditory experience. The received ear signals can be attributed to the two sound source characteristics of "location" and "signal", which are independent of each other but always occur in a pair-wise fashion. Therefore, the presented association model is in agreement with many phenomena related to localisation in the superimposed sound field……"

12. Confirmation of a need to process timing information:

Source: http://arxiv.org/pdf/1208.4611v2.pdf

Gave the following summary:

"..The time-frequency uncertainty principle states that the product of the temporal and frequency extents of a signal cannot be smaller than 1/(4PI). We study human ability to simultaneously judge the frequency and the timing of a sound. Our subjects often exceeded the uncertainty limit, sometimes by more than tenfold, mostly through remarkable timing acuity. **Our results establish a lower bound for the nonlinearity and complexity of the algorithms employed by our brains in parsing transient sounds, rule out simple "linear filter" models of early auditory processing, and highlight timing acuity as a central feature in auditory object processing**…."

And further:

"…In many applications such as speech recognition or audio compression (e.g. MP3 [18]), the first computational stage consists of generating from the source sound sonogram snippets, which become the input to latter stages. **Our data suggest this is not a faithful description of early steps in auditory transduction and processing, which appear to preserve much more accurate information about the timing and phase of sound components [12, 19, 20] than about their intensity**…."

And finally:

"…Early last century a number of auditory phenomena, such as residue pitch and missing fundamentals, started to indicate that the traditional view of the hearing process as a form of spectral analysis had to be revised. In 1951, Licklider [25] set the foundation for the temporal theories of pitch perception, in which the detailed pattern of action potentials in the auditory nerve is used [26, 28], as opposed to spectral or place theories, in which the overall amplitude of the activity pattern is evaluated without detailed access to phase information. The groundbreaking work of Ronken [22] and Moore [23] found violations of uncertainty-like products and argued for them to be evidence in favour of temporal models. However this line of work was hampered fourfold, by lack of the formal foundation in time-frequency distributions we have today, by concentrating on frequency discrimination alone, by technical difficulties in the generation of the stimuli, and not the least by lack of understanding of cochlear dynamics, since the active cochlear processes had not yet been discovered.

Perhaps because of these reasons this groundbreaking work did not percolate into the community at large, and as a result most sound analysis and processing tools today continue to use models based on spectral theories. We believe it is time to revisit this issue….."

13. Transient and localization

Some very interesting information on transients and localization comes from the development work of Joseph Manger. The whole paper is recommended for reading.

Source: http://www.manger-audio.co.uk/PDFs/acoustical_reality.pdf

2/ Perception and hearing explained by new research

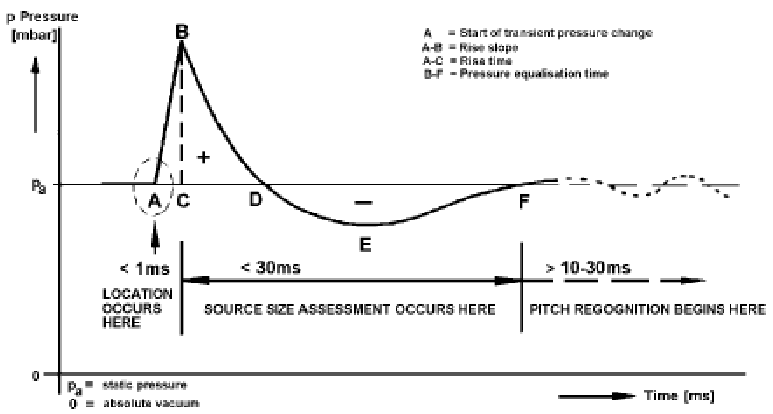
The human hearing mechanism does not just detect the existence of sound, it also estimates the direction of the source as well as analysing the content of the sound to determine the most likely cause. In musical sounds, the pitch will also be determined. Josef Manger has been studying these mechanisms for over 20 years. He has found that each mechanism takes a different time to operate following an initial transient. The location and nature of the sound source are completely discerned before the pitch is recognised.

Pitch and timbral recognition is described by the well-established place theory, described in Part 1, in which different parts of the basilar membrane resonate according to the frequencies in the sound. However, various authorities, such as Keidel, Spreng, Klinke and Zenner, have suggested that there is another, faster acting, mechanism which works in the time domain.

The theory could not be tested with conventional loudspeakers. Confirmation of the theory was not possible until Josef Manger used his newly developed transducer as the sound source.

Fig.3 illustrates this principle of transient analysis and shows an idealised transient pressure waveform following an acoustic event. There are three important points made in the figure:

Fig. 3



1/ A complete cycle is quite unnecessary for the recognition of the sound source. Only the initial transient pressure change A-B is required. The time of arrival of the transient at the two ears will be different and will locate cause, i.e. the source laterally within around a millisecond.

2/ Following the event which generated the transient, the air pressure equalises itself along the line B-F. The period of time between B and F varies and allows the listener to establish the likely size of the sound source.

3/ Only after the recognition of the source from the transient is the pitch recognised according to the place theory of the basilar membrane from the part of the wave-form beyond F.

The information in the initial transient pressure waveform goes beyond locating the source. Fig. 4 illustrates how the size of a sound source affects the pressure equalisation time. Pressure waveforms from a hand gun, a rifle and a cannon are shown. It will be seen that the larger the source, the longer the pressure equalisation time.

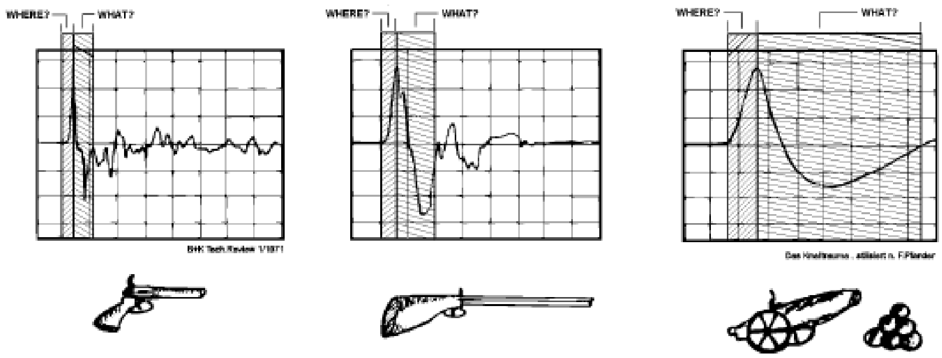
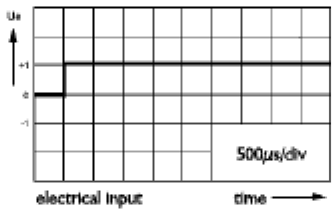
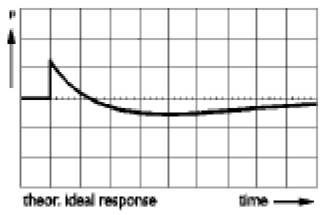


Fig. 4

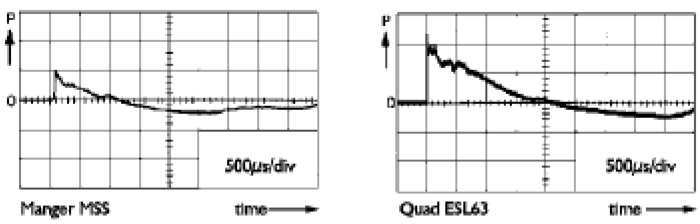
How distorted transients can be – Manger illustrates it on the following pictures:



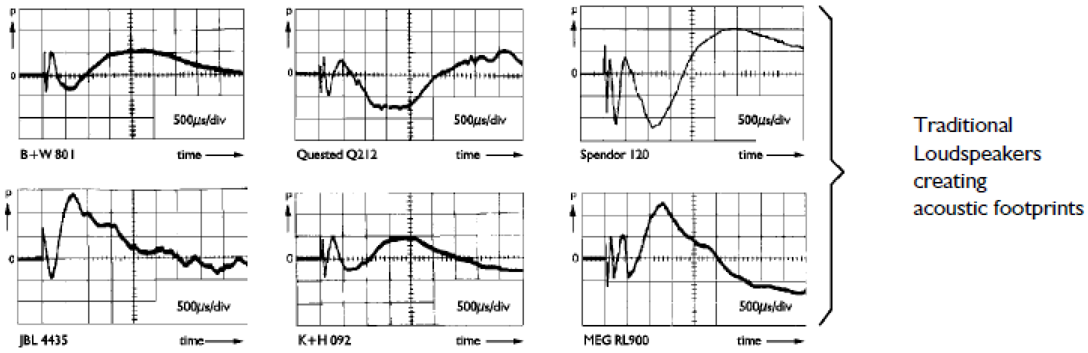
electrical input



theoretical ideal presuure
step response in air



closed to the ideal



14. AES Technical Document on phase accuracy and transient fidelity.

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.

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

Some comments are presented in
http://www.bodziosoftware.com.au/AES_Document_Comments.pdf

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

Parameters	*Units/Conditions*	*Value*
Amplitude/frequency response	40 Hz–16 kHz 0°	Tolerance 4 dB
	±10° Horizontal ±30°	Deviation to 0°, 3 dB 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 >100 Hz	–30 dB (=3%) –40 dB (=1%)
Transient fidelity 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] (preferably $2.5/f$)
Time delay Difference between stereo loudspeakers	∂t	≤10 µs
System dynamic range Maximum operating level (measurement acc. to IEC 60268, § 17.2, referred to 1 m distance)	$L_{\text{eff max}}$	>112 dB (at IEC 60268 program simulation noise or special condition)
Noise level	L_{noise}	≤10 dBA

15. Audibility of transients

I have come across an interesting paper in JAES ,Vol.38, No.11,1990 November, "On the Correlation between the Subjective Evaluation of Sound and the Objective Evaluation of Acoustic Parameters for a Selected Source".

The authors performed subjective and objective analysis of several woofers using impulsive tones, and concluded:

"…**A detailed analysis of the results of the subjective evaluation of loudspeakers showed that the subjective evaluation of the obtained sounds was decisively influenced by the work of the loudspeaker in a transient state. It appeared that the longer the duration of final transient and the smaller the value of coefficient D, the greater the sharpness of the sounds emitted by the loudspeaker**…."

16. Transients and Localization

The following paper, clearly indicates, that transients are critical in localization process.

Source: http://www.pa.msu.edu/acoustics/rooms1.pdf

I have located an interesting piece of information in the paper "Localization of sound in rooms", from JASocAm. 74 (5) Nov 1983. The paper is by WM Hartman from Michigan State University, Dept. of Physics, and provides the following summary:

"…This paper is concerned with the localization of sources of sounds by human listeners in rooms. It presents the results of source-identification experiments designed to determine whether the ability to localize sound in a room depends upon the room acoustics, and how it depends upon the nature of the source signal.

The experiments indicate that the **localization of impulsive sounds, with strong attack transients, is independent of the room reverberation time**, though it may depend upon the room geometry.

For **sounds without attack transients, localization improves monotonically with the spectral density of the source**.

Localization of continuous broadband noise does depend upon room reverberation time….."

More papers by Hartmann and Rakerd.

"Localization of sound in rooms, II: The effects of a single reflecting surface"

http://www.pa.msu.edu/acoustics/rooms2.pdf

"….Our results indicate the following: (1) **A sound must include transients if the precedence effect is to operate as an aid to its localization in rooms.** (2) Even if transients are present the precedence effect does not eliminate all influences of room reflections. (3) Due to the interference of reflections large interaural intensity differences may occur in a room and these have a considerable influence on localization; this is true even at low frequencies for which IID cues do not exist in a free field. (4) Listeners appear to have certain expectations about the reliability and plausibility of various directional cues and perceptually weight the cues accordingly; we suggest that this may explain, in part, the large variation in time-intensity trading ratios reported in the literature and also the differing reports regarding the importance of onsets for localization. (5) In this study we find that onset cues are of some importance to localization even in free field.

"Localization of sound in rooms: III: Onset and duration effects"

http://www.pa.msu.edu/acoustics/rooms3.pdf

Conclusions

"…(1) **A rapid onset facilitates localization in a free field by a measurable but small amount, about 0.5deg. It facilitates localization in rooms by substantially larger amount because the onset allows the precedence effect to operate and without the precedence effect localization is poor due to misdirection cues in steady-state sound field.**

(2) **The precedence effect is maximally effective when the signal onset is instantaneous. Its effectiveness begins to diminish as the onset duration is increased….."**

17. More on localization and transients

Source: http://www.pa.msu.edu/acoustics/rakhar2.pdf

In a paper by Brad Rakerd and William M. Hartmann "Localization of noise in a reverberant environment" (Michigan State University), they conclude:

"…(1) Localization of noise is enhanced by an attack transient. An attack transient appears to be particularly helpful when the direct-reverberant ratio is low. Attack transients give an advantage over slow onsets when the reflections are not much delayed *re* the direct sound. By contrast, attack transients are of only marginal value when noise is presented by headphones or tones are presented in an anechoic room (Tobias and Schubert, 1959; Rakerd and Hartmann, 1986).

(2) Onsets are a great leveler among individuals. Whereas the ability to localize steady steady-state sounds varies greatly among listeners, the ability to localize sounds with an onset transient shows best to worst differences less than 1.5 degrees among our seven listeners…."

18. Even more on localization and transients

Source: AES library. Preprint 2745. 86th Convention.
"Localization of sound in a room with reflecting walls" W.M. Wagenaars

"….3. CONCLUSIONS

In this study localization of sound in a room with reflecting walls was tested. Eleven stimuli were used, differing in spectral and temporal information. For such a room the following can be concluded:

- Signal bandwidth is an important cue for localization. The broader the frequency spectrum of a sound, the better localization performance.

- Offsets seem to be an equally important cue for localization as onsets. Localization performance are similar for signals with an abrupt onset, offset, or both.

- Localization performance for steady state sinusoids is frequency-dependent. For simply gated sinusoids performance is not dependent of frequency.

Although many of the errors made were distance errors, subjects are able to localize distance quite well. Furthermore subjects usually select the correct side, even for the hard to localize steady state sinusoids….."

19. Sound Quality and Transient Response.

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

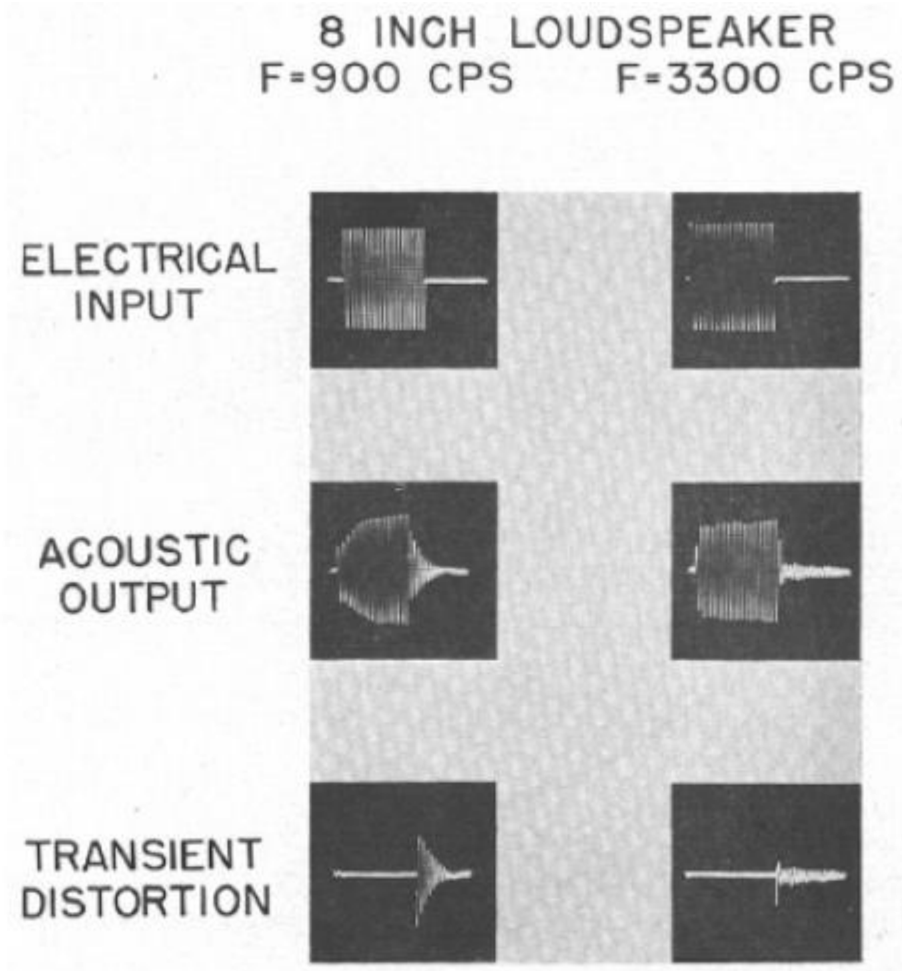


FIG. 3. Transient distortion for 8-in. loudspeaker.

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…."

20. Confirmation of two-stage processing by the ear, as discussed in (3) and (11).

Yet another interesting paper. It puts the early reflections in somewhat different perspective.

"The Significance of Early High-Frequency Reflections from Loudspeakers in Listening Rooms", Preprint 4094, David Moulton, David Moulton Professional Services, Groton, MA

"…Any reverberant space yields comb-filtering effects, and virtually all listening to music via loudspeakers is done in such spaces. Therefore, logically speaking, all listening is done under compromised conditions, where a primary attribute of accurate sound reproduction (fiat amplitude response) is negated. Yet we must acknowledge that music playback systems seem to work well: listeners enjoy listening, they readily and accurately identify sounds (and will testify to their realism), and some listeners are able to detect truly microscopic differences between alternate components in the playback system.

This anomaly raises the question: how can individuals listen effectively to loudspeakers in reverberant spaces and why don't the ubiquitous comb-filtering interference effects always pose problems for the listener?

I suggest that the answer lies in the nature of our auditory localization capability, which makes use of interference effects such as comb-filtering as a function of performing the sound source localization task.

That task is performed at a pre-conscious neurological stage and most early reflections are localization information that is not presented to the conscious mind. Further, we do not consciously perceive the amplitude response characteristics of comb-filtering effects that occur in reverberant spaces as a result of early reflections, even though such effects are clearly measurable…."

The above statement confirms earlier findings of Gunter Theile, Watkinson and Manger about the ear processing the incoming audio stimulus in two stages: The received ear signals can be attributed to the two sound source characteristics of "location" and "signal", which are independent of each other but always occur in a pair-wise fashion.

21. General Conclusions From Papers Presented Above

First of all – the room itself.

Accordingly to Bernd Theiss, Malcolm O. J. Hawksford in AES Preprint 4462:

"…Early reflections < 2.5 ms.

Early reflections occurring less than 2.5 ms after the original sound sensation are known to shift the image towards their direction and to blur the image.

Early reflections < 5 ms.

Early reflections occurring more than 2.5 ms but less than 5 ms after the original sound sensation are known to blur the image, although they keep the direction of the image constant….".

So if your goal is to deliver the sharpest image, or most accurate localization, you would be well advised to take care of transient origination (loudspeakers) and also provide some acoustical treatment to the walls/room.

There are basically three areas where linear-phase loudspeakers differ from minimum-phase loudspeakers.

1. Linear-phase speakers provide more accurate spatial information, rather than timbral. Tonal balance is the same for both loudspeaker types. This is where the tests are falling apart, because listeners are looking for tonal differences, rather than subtle spatial clues – sharper image, better located soloists, stage depth. It's subtle, but it is there.

2. Identical phase response for all loudspeaker in the system. The phase response in correctly equalized multi-channel linear-phase system is 0deg in every loudspeaker. Therefore it immediately satisfies AESTD1001.1.01-10 for phase accuracy and transient fidelity to perfection. The measurements of linear-phase loudspeaker are presented on my website, and comments on AESTD1001.1.01-10 are presented in - http://www.bodziosoftware.com.au/AES_Document_Comments.pdf

3. Tighter bass. Even Dr. Floyd Toole quoted other researchers (Craven and Gerzon) on this subject on page 420. The most obvious difference is the tighter bass. I have conducted extensive tests on this subject - http://www.bodziosoftware.com.au/LP_MP_Subwoofer_Tests.pdf

22. Square Wave Loudspeaker Testing

Another interesting paper from 94'th AES Convention. I would recommend reading the entire paper.

Source: "Directions for Qualified Loudspeaker Evaluations", AES Preprint 3603, Peter M. Pfleiderer, 1993.

The paper concludes with the following summary:

”…**Summary**

An almost unbelievable state of perfection has been reached for electronic components within the electroacoustical reproduction chain due to competent applications of measurement technology. With loudspeakers, on the other hand, competent measurement methods are currently not even in practical use. Obviously, test methods are required which are capable of uncovering major changes to signal waveforms originating from linear and acoustical errors.

Measurements with square wave signals should be included as standard testing procedures in order to be able to detect errors with sound quality and spatial imaging in all HiFi components, but especially in loudspeaker systems. Many technical and acoustical faults can namely not be registered with SPL or frequency measurements, although they have induced significant irregularities into the relevant audio signal waveform.

This is the reason why loudspeakers of proven square wave response capability are an important prerequisite for the natural reproduction of sound. Moreover, it is only possible to detect acoustic faults with this type of technically faultless reference loudspeaker. It should be clearly noted that all other current components in the electroacoustical reproduction chain already transmit square wave signals correctly.

Correct square wave reproduction with loudspeakers has the same importance as was the case for correct square wave reproduction with amplifiers in the 1960's. Both constitute fundamental advances and establish important conditions for high quality reproduction of music. Nothing can propagate the concept of high fidelity more than these types of advances….”

23. JAES ,Vol.61, No.11, 2013 November, "**"Sensitivity of Human Hearing to Changes in Phase Spectrum", MIKKO-VILLE LAITINEN, AES Student Member, SASCHA DISCH, AND VILLE PULKKI, AES Fellow.**

The paper is 18-pages long, and has interesting conclusion too. Here is only a short excerpt:

"….Human ability to perceive differences in sounds due to the modification of the phase spectrum was studied in this article. Formal listening tests were arranged and synthetic harmonic complex signals were used as test signals. The results of the tests confirm that humans are not "phase deaf," the perceived difference due to randomization of the phase spectrum can be larger than the difference due to randomization of the magnitude spectrum with a standard deviation of 4 dB…."

24. More listening tests conducted by other people.

Here is a quote two well-known audio workers and publishers:

http://www.stereophile.com/reference/706deep/index.html

Michael Gerzon **—**

"…The subjective effect of phase compensation of the bass from loudspeakers is very marked, giving a much tighter and more 'punchy' quality, with greater transparency, and interestingly a subjective extension of bass response of at least half an octave. The improvement is audible even on loudspeakers with a very high cut-off frequency, such as Quad electrostatic designs. . . . The benefits of bass phase equalisation are considered, **by those who have heard it**, to be a substantial improvement over what was hitherto possible with analog technology, and digital equalisation provides a way of improving bass performance without going to ridiculously large giant space-consuming power-hungry monster speakers, and is certainly a much cheaper route…"

Keith Howard describing listening test with linear phase subs.

"…As I pressed Play, I didn't know what to expect: a mild improvement or a revelation. To use a curry-house analogy, I was half-prepared for korma rather than phal. But when I compared the unprocessed and phase-corrected tracks, it took only a few seconds of the latter to persuade me that here was a significant improvement. Just as Michael Gerzon described, the phase-corrected sound was both weightier and punchier, and distinctly more coherent. It simply sounded more like a bass guitar, to the extent that I almost hummed along—not something that I would normally expect to do when listening to a bass-guitar accompaniment shorn of all else (sorry, John), and certainly not an urge I felt with the unprocessed track. Still, time-consuming as it was, this experiment involved only a single music excerpt and was conducted in mono, and so hardly offers a comprehensive insight into the full benefit of bass phase correction. But what I heard convinces me that this is an area in which the application of DSP can make a significant contribution to fidelity…."

Personally, as I mentioned before, I have tested linear-phase subwoofers quite extensively and published the results on my website. I used 2-5ms pulse, bi-polar pulse, square wave, and LFE from movies. The linear-phase bass remains seismic, deep and powerful, but is also tight and has punch to it now.

In short – this bass is accurate.

25. On Importance of Impulses, Transients and Waveforms.

Earthworks' president David E Blackmer:

http://www.earthworksaudio.com/wp-content/uploads/2012/07/The-world-beyond-20kHz.pdf

"…Human hearing is generally, I believe, misunderstood to be primarily a frequency analysis system. The prevalent model of human hearing presumes that auditory perception is based on the brain's interpretation of the outputs of a frequency analysis system which is essentially a wide dynamic range comb filter, wherein the intensity of each frequency component is transmitted to the brain…"

Further:

"…The inner hair cells clearly relate to the frequency analysis system described above. Only about 3,000 of the 15,000 hair cells on the basilar membrane are involved in transducing frequency information using the outputs of this traveling wave filter. The outer hair cells clearly do something else, but what?

There are about 12,000 'outer' hair cells arranged in three or four rows. There are four times as many outer hair cells as inner hair cells(!) However, only about 20% of the total available nerve paths connect them to the brain. The outer hair cells are interconnected by nerve fibers in a distributed network. This array seems to act as a waveform analyzer, a low-frequency transducer, and as a command center for the super fast muscle fibers (actin) which amplify and sharpen the traveling waves which pass along the basilar membrane thereby producing the comb filter. It also has the ability to extract information and transmit it to the analysis centers in the olivary complex, and then on to the cortex of the brain where conscious awareness of sonic patterns takes place. The information from the outer hair cells, which seems to be more related to waveform than frequency, is certainly correlated with the frequency domain and other information in the brain to produce the auditory sense.

Our auditory analysis system is extraordinarily sensitive to boundaries (any significant initial or final event or point of change). One result of this boundary detection process is the much greater awareness of the initial sound in a complex series of sounds such as a reverberant sound field. This initial sound component is responsible for most of our sense of content, meaning, and frequency balance in a complex signal. The human auditory system is evidently sensitive to impulse information imbedded in the tones. My suspicion is that this sense is behind what is commonly referred to as 'air' in the high-end literature. It probably also relates to what we think of as 'texture' and 'timbre' - that which gives each sound it's distinctive individual character. Whatever we call it, I suggest that impulse information is an important part of how humans hear.,,,"

26. Ability of the brain to process time and frequency information in parallel.

THIS IS YOUR BRAIN ON Music, The Science of a Human Obsession, Daniel J. Levitin

Published by Penguin Group (USA) Inc.375 Hudson Street, New York, New York 10014, U.S.A.
MSR ISBN 0-7865-8404-1
AEB ISBN 0-7865-8405-X

How does the brain figure out, from this disorganized mixture of molecules beating against a membrane, what is out there in the world? In particular, how does it do this with music?

It does this through a process of feature extraction, followed by another process of feature integration. The brain extracts basic, low-level features from the music, using specialized neural networks that decompose the signal into information about pitch, timbre, spatial location, loudness, reverberant environment, tone durations, and the onset times for different notes (and for different components of tones).

These operations are carried out in parallel by neural circuits that compute these values and that can operate somewhat independently of one another— that is, the pitch circuit doesn’t need to wait for the duration circuit to be done in order to perform its calculations. This sort of processing— where only the information contained in the stimulus is considered by the neural circuits—is called bottom-up processing. In the world and in the brain, these attributes of the music are separable. We can change one without changing the other, just as we can change shape in visual objects without changing their colour.

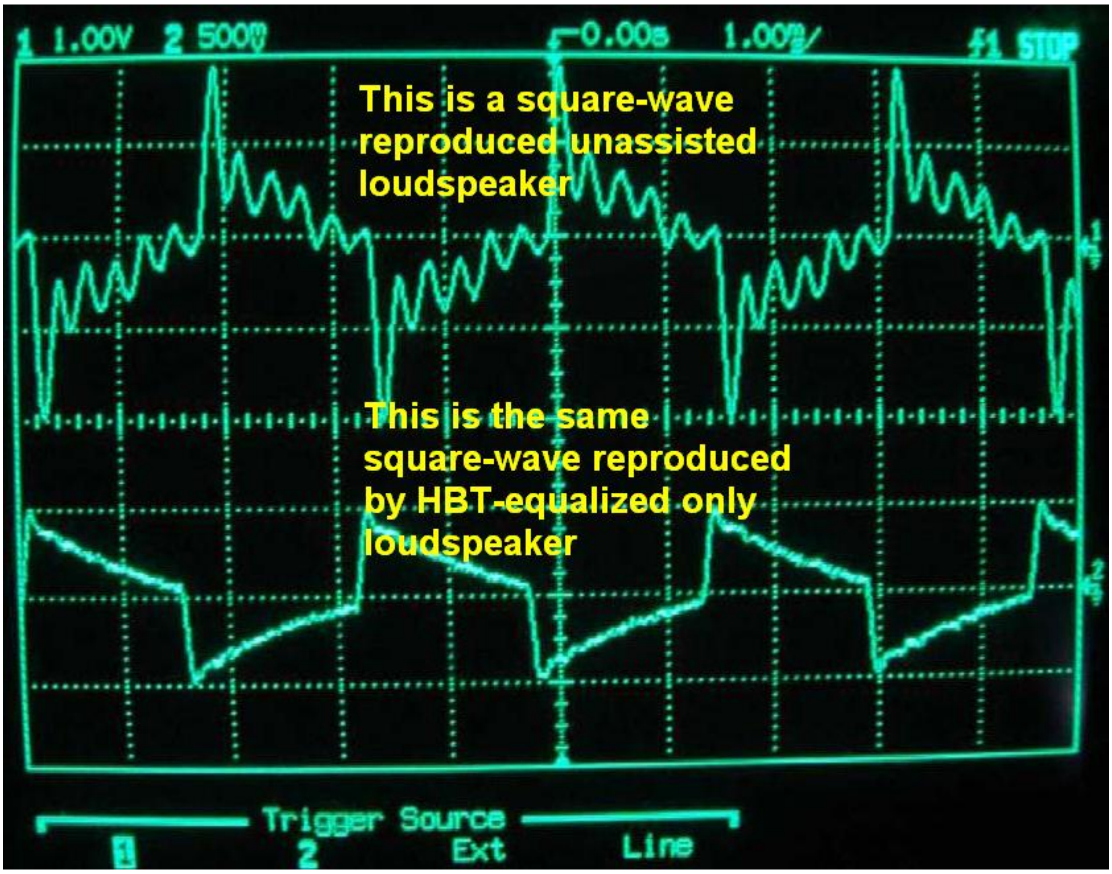
Time Domain Instrument Testing

Real-life loudspeaker example

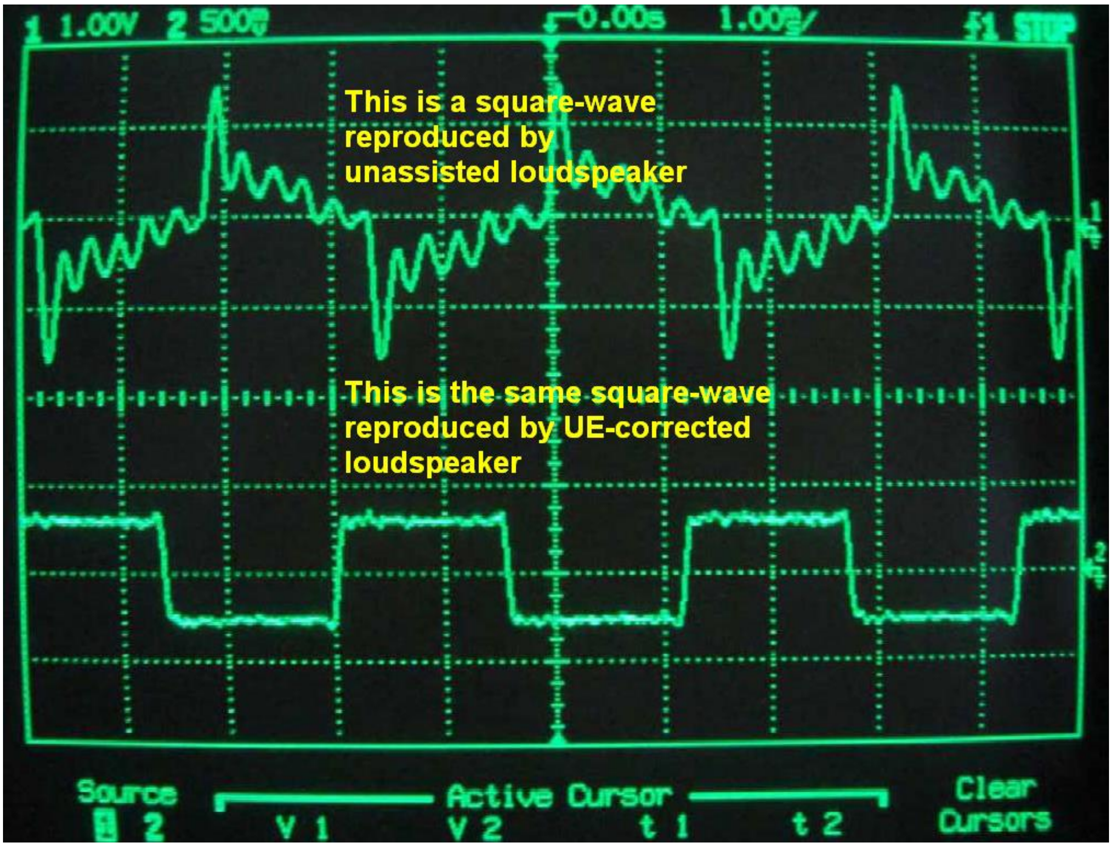
The system under test discussed here consists of a filter and a loudspeaker in an enclosure. These two components that will introduce time delay are the filter and the combination of driver and the enclosure itself. To illustrate the above, a 12” guitar loudspeaker in a vented box was measured and it’s minimum-phase responses were obtained with a help of an MLS measurement technique – see below. It is immediately observable, that the loudspeaker has rather irregular frequency response. Since the loudspeaker is essentially a minimum-phase device, the corresponding phase response is also highly irregular, and definitely not flat.

Let’s establish the frequency response of interest, which is the frequency range where the SPL will be equalized to flat response. In my example it will be: 90Hz – 5500Hz.

A 300Hz square wave reproduced by this loudspeaker is highly distorted. Strong ringing is due to 10dB sharp SPL peak located at 3.5kHz. You can see, that there are about 11periods of ringing waveform in one period of 300Hz square wave.



This is a square-wave reproduced unassisted loudspeaker
This is the same square-wave reproduced by HBT-equalized only loudspeaker
1 1.00V
2
0.00s
STOP
Trigger Source
2
Ext
Line

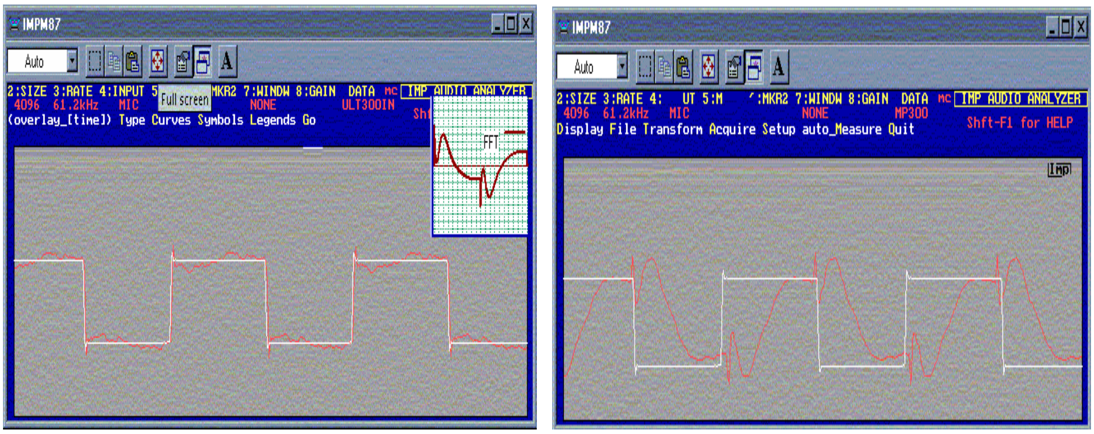


This is a square-wave reproduced by unassisted loudspeaker
This is the same square-wave reproduced by UE-corrected loudspeaker
1 1.00V
2
0.00s
STOP
Source
2
Active Cursor
V 1
V 2
t 1
t 2
Clear Cursors

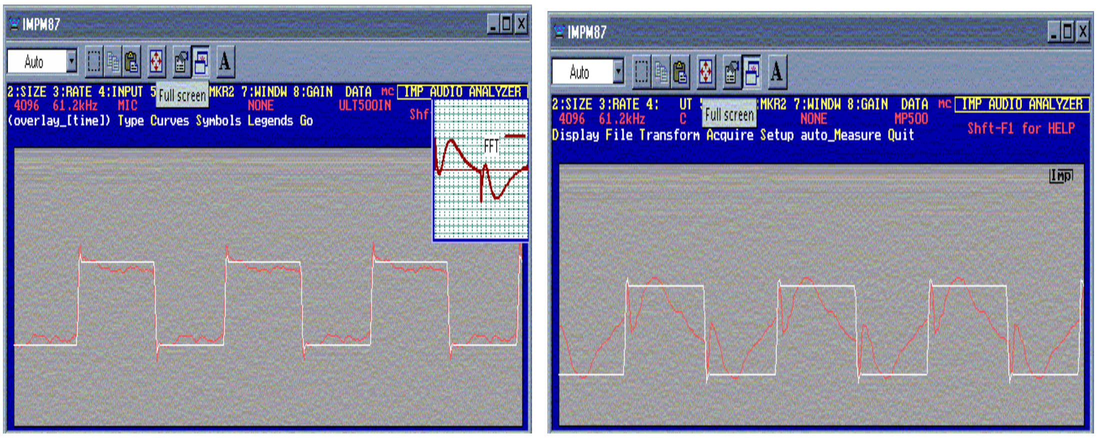
Instrument test results obtained from linear-phase loudspeakers reveal their true superiority in time domain. The following test results were obtained by John Kreskovsky of Music and Design (http://www.musicanddesign.com)

As John points out: "….The measurements were not taken in an anechoic environment and are of the continuous time type, recorded over numerous cycles, windowing over a reflection free period can not be performed. Thus, there is some contamination by room reflections resulting is some degradation in the observed response.

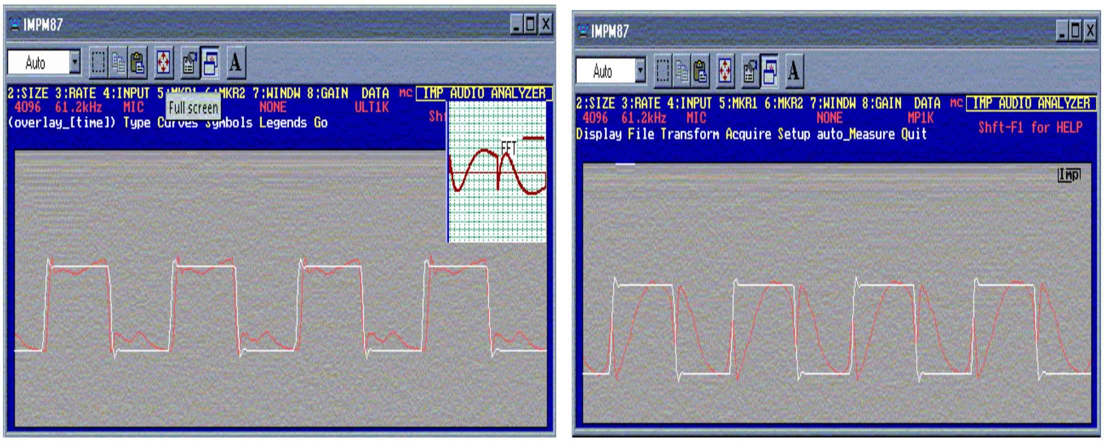
The first figure shows the 300 Hz response. This is close to the low frequency cut off of the system where the phase rotation and group delay due to the 200 Hz high pass cut off would normally result in loss of flat top behaviour and the 2k Hz crossover would cause distortion of the initial rise. This is shown in the insert at the upper right of the plot for the linearized system and confirmed by the lower plot which if for the standard LR4 system. The white trace is the input, orange the acoustic output from the speaker system.



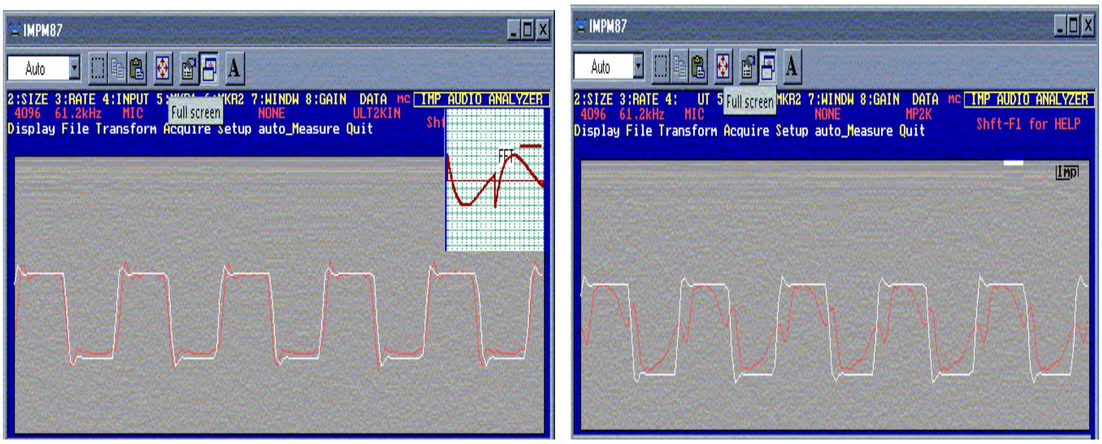
300 Hz square response of linearized system, left, and standard LR4 crossover, right.



500 Hz square response of linearized system, left, and standard LR4 crossover, right.



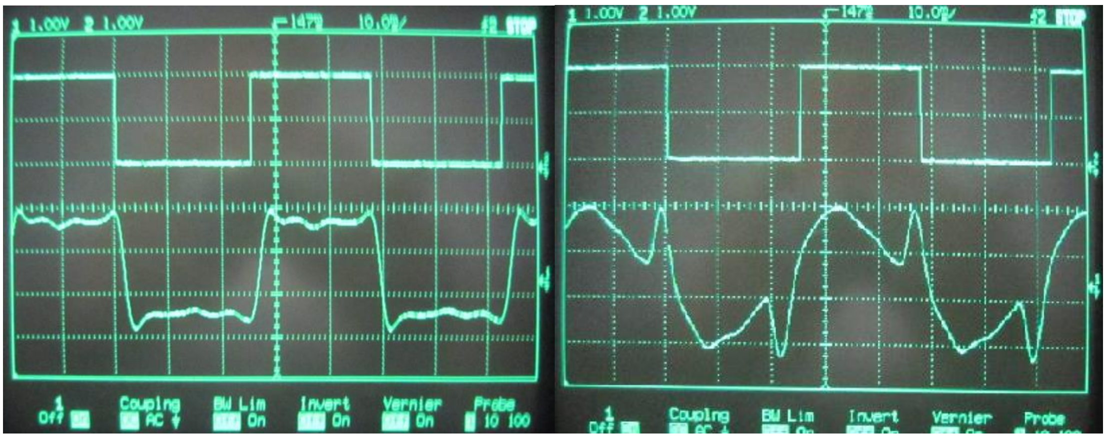
1kHz square response of linearized system, left, and standard LR4 crossover, right.



2kHz square response of linearized system, left, and standard LR4 crossover, right.

….” End of quote.

My own measurements on 18” McCauley subwoofers further confirm time domain superiority of linear-phase loudspeakers.



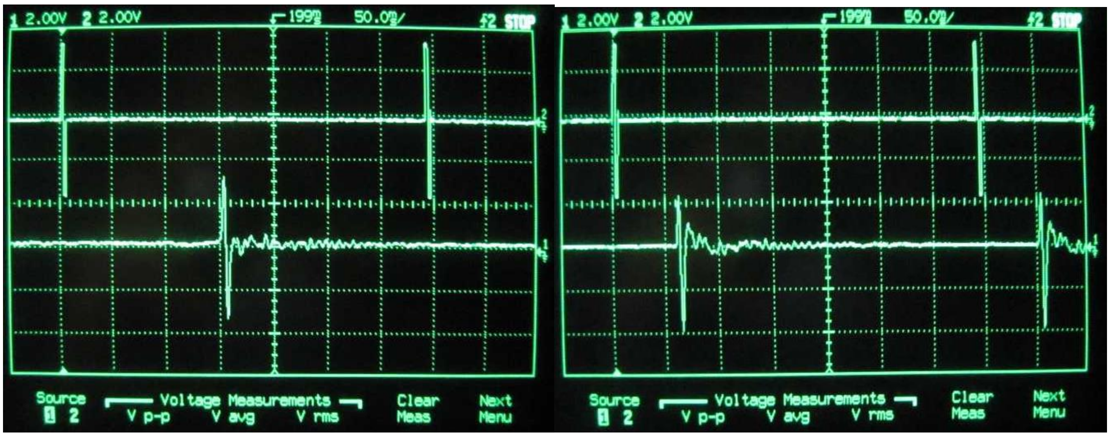
20Hz square wave: Linear-Phase Mode and Minimum-Phase Mode

Shown above, the time-domain comparison measurement results speak for themselves. It needs to be remembered, that we are dealing here with a very heavyconed, 18" driver, low-pass filtered, in a vented (resonating) enclosure, and yet, the time domain performance is near-perfect accurate. It's pretty amazing to see a vented loudspeaker, holding the acoustic pressure nearly constant for 25ms.

Next, I used 2ms-wide pulses separated by 350ms space as the source signal. On the 2ms pulse, the minimum-phase version delivered a more of a "thump" instead of a pop or a click. This is perhaps not surprising, as the post-ringing of the pulse extended to130ms and far exceeded the 30ms "memory effect" of the auditory system. Here, the driver, filter and vented enclosure added it's own, combined signature. It is also observable, that the minimum-phase version of the subwoofer has converted the clearly asymmetrical pulse into a much more symmetrical bi-polar pulse with post-ringing. This is clearly visible on the screen shots below.

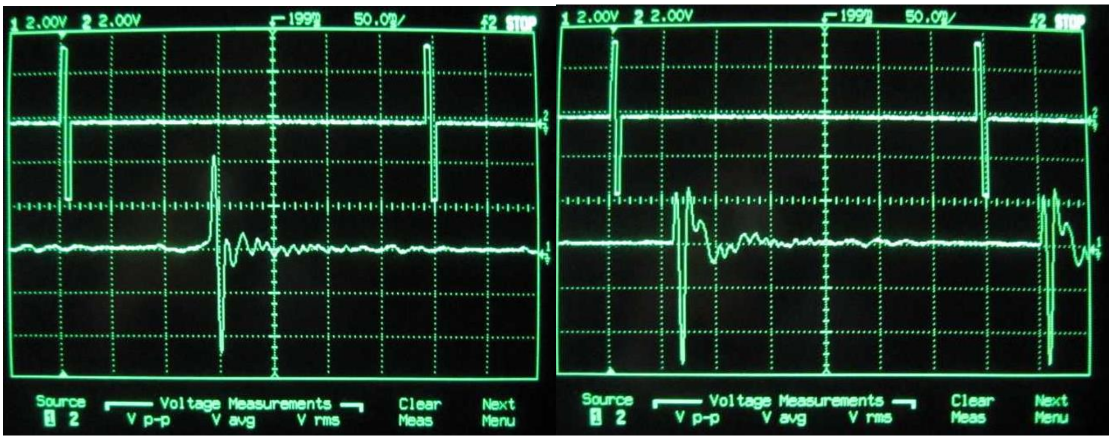
5ms Impulse in Linear-Phase Mode and Minimum-Phase Mode

When a 2ms bi-polar pulse was used for excitation, the minimum-phase version has done the opposite, and converted the symmetrical bi-polar pulse into a pulse with clear asymmetrical tendency. The ringing past the pulse is due to a more distant microphone placement, so now, the mike picks some of the room reflections.



2ms Bi-polar pulse in Linear-Phase Mode and Minimum-Phase Mode

When a 10ms bi-polar pulse was used for excitation, the minimum-phase version has even more asymmetrical tendency.

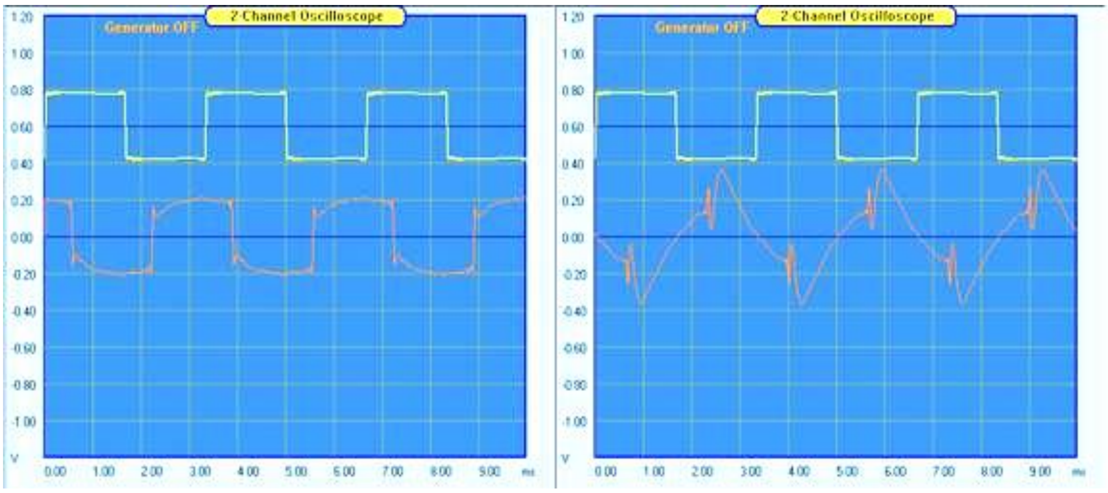


10ms Bi-polar pulse in Linear-Phase Mode and Minimum-Phase Mode

Finally some more square wave measurements from UE User's Manual.

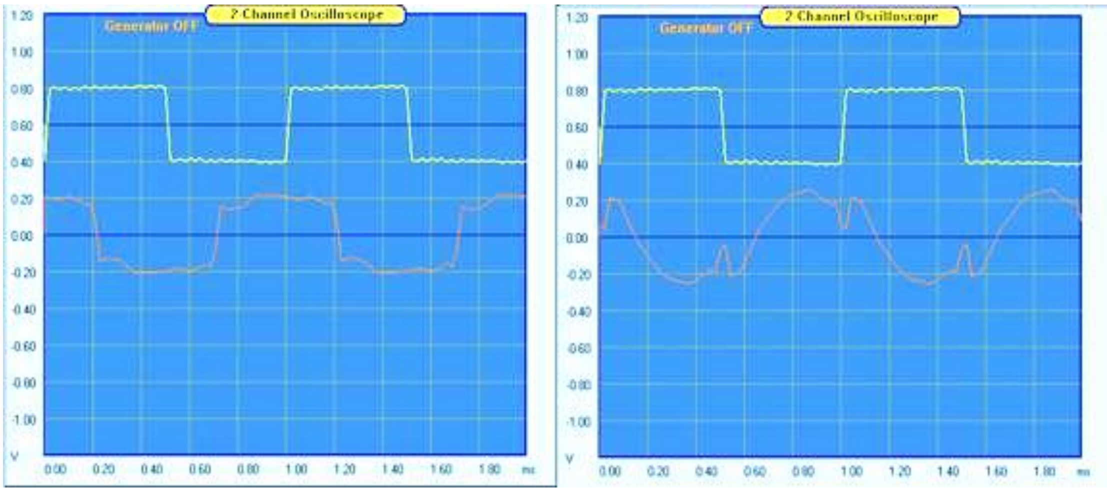
The linear phase result is on the left and the nonlinear phase result on the right. It should be noted that there is some distortion in the wave forms that that must be attributed to room reflections. Square wave testing is a steady state test and without a true anechoic chamber the effects of room reflections can not be eliminated.

Never the less, for the 300 Hz case shown in the first figure, the linear phase system shows the sharp rise and fairly flat top expected. The nonlinear phase case shows early tweeter response followed by the woofer response and the sloped top is an artefact of the nonlinear phase. The response also significantly overshoots the correct level. This latter effect is seldom discussed when comparisons of linear and nonlinear phase systems are made. Even though the amplitude of reach frequency component is correctly reproduced in the nonlinear phase system, the lack of linear phase means that the different frequency components do not sum correctly since that are delayed by different amounts. The overshoot is a result of time distortion.



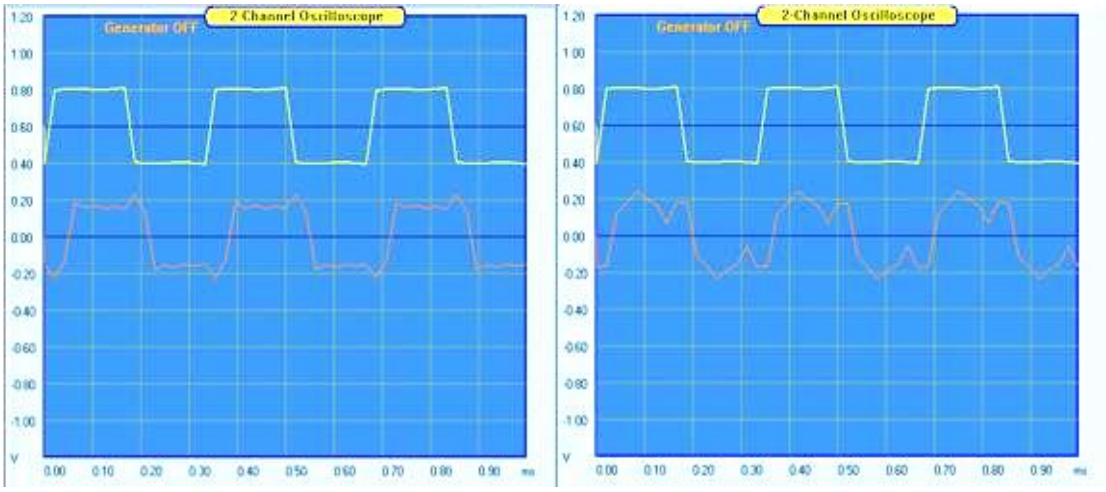
300 Hz square wave response, Linear phase, left; Nonlinear phase, right

The next figure shows the same comparison for a 1kHz square wave. Again, some distortion is observed due to room reflections. However, the linear phase case again shows the expected sharp rise and relatively flat top. The nonlinear phase system more clearly shows the time lag between the woofer and tweeter response.



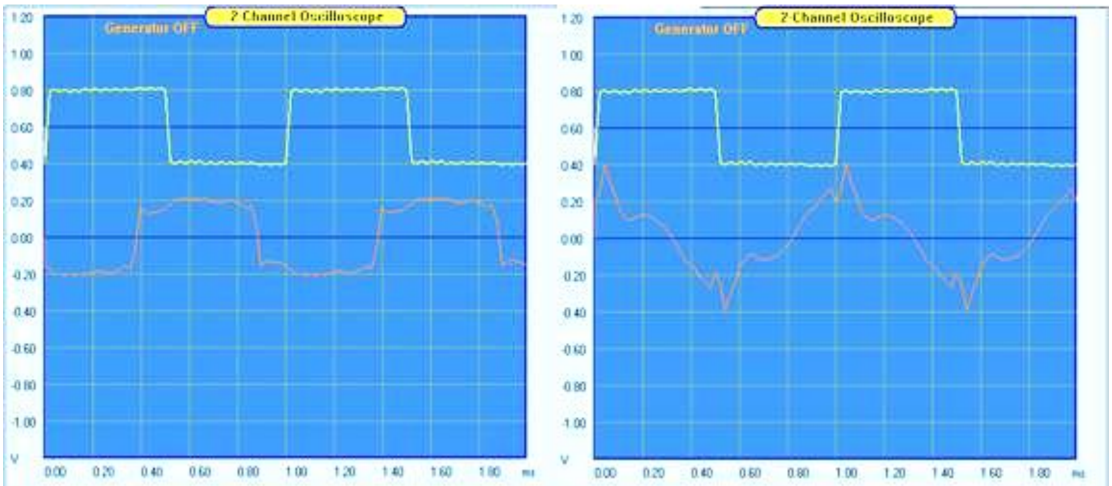
1kHz square wave response, Linear phase, left; Nonlinear phase, right.

The next figure shows the result for a 3k Hz square wave. The differences between linear and nonlinear phase, while clearly evident, are less significant because the fundamental is above the crossover point and there is little contribution from the woofer due to the 4th order low pass response. With the system designed another interesting feature of the linear phase system can be examined, the effect of crossover slope.



3kHz square wave response, Linear phase, left; Nonlinear phase, right.

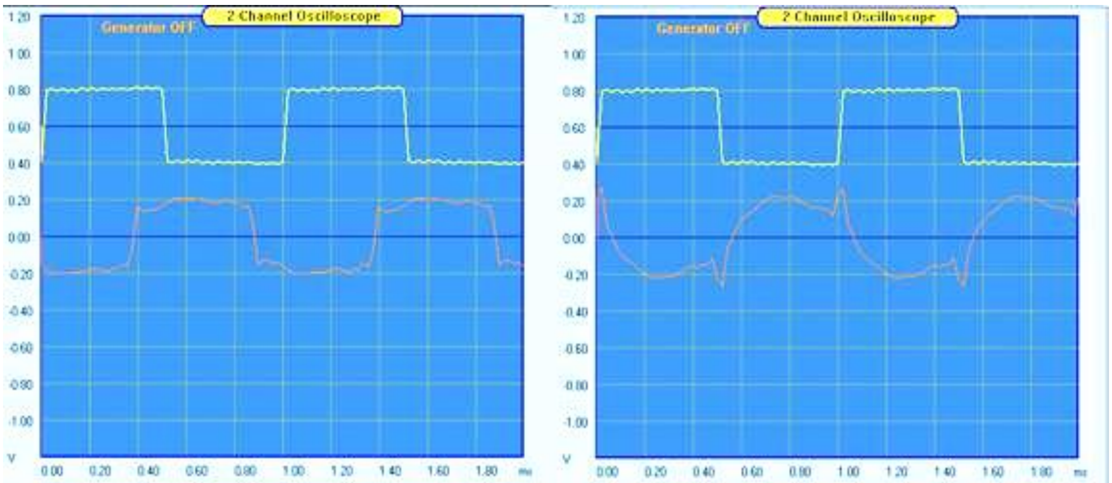
The next figure shows the 1kHz response of the linear phase and nonlinear phase system when the slope of the woofer to tweeter crossover is increased to 8th order, 48dB/octave. With the *Ultimate Equalizer* this is easily accomplished by selecting the new 48dB/octave slopes and clicking *Show complete system* to calculate and load the new filters.



1kHz response of linear and nonlinear phase system with 8th order crossover.

This result should be compared to that of figure where the crossover was 4th order. Changing the order has no effect on the linear phase system at the design point. The nonlinear phase system response is significantly different solely due to the change in crossover order.

Finally, the last figure shows the effect of reducing the crossover to 2nd order. The response of the nonlinear phase system looks somewhat better now. However, for flat response the tweeter must be connected with inverted polarity in the nonlinear phase system and the initial tweeter pulse is therefore in the wrong direction. It should be noted that many audio enthusiasts feel the 2nd order crossover sound better than those of higher order. This may be a result of the improved wave form observed here and could be an indication of the potential of linear crossover and speakers of any order since they will all preserve wave form relative to the design point.



1kHz response of linear and nonlinear phase system with 2nd order crossover.

Conclusions

At the time of this writing, linear-phase loudspeakers are still a new "kid on the block". Past attempts in creating them resulted in offerings that were simply too expensive for wide-spread use. The most accurate implementation of linear-phase loudspeaker requires a full set of individual driver measurements, coupled with a DSP approach, in addition to an active amplification system. This really makes the linear-phase system highly customized device – a world of difference in comparison to the current approach of loudspeaker industry.

However, this particular feature makes the linear-phase system an ideal DIY device. In our world, everything is custom-built, with an aim to typically outperform comparable commercial designs. Linear-phase loudspeakers offer everything that minimum-phase loudspeakers can offer, and then reward you with often vastly superior performance in time domain, as explained in the pages above.

It appears, that my poor and outdated listening/evaluating habits, coupled with lack of standard listening methodology for time/space-domain assessment of loudspeakers conspired to cloud my ability to really critically listen to the full set of my loudspeakers during some of my evaluation tests. Secondly, not every musical material will reveal all time-domain characteristics to the same degree. For instance, tight, well-defined bass, will manifest itself on gunshots and explosions in DVD movies, but will not stand out during low-frequency, seismic earthquake effects on LFE channel. In more critical tests, I did pick the "tighter bass" characteristic, as it was too obvious to miss on the large, 18" subs. Also, I pointed out earlier the effect of feeling closer to the orchestra, as if I could better discriminate their sitting arrangement. Both of these effects have really nothing to do with frequency domain – they are both more of the time/space domain phenomena.

It is clear, that designing loudspeakers using frequency-domain characteristics as the main (or only) criteria leads to stagnated, oversimplified, and ultimately inaccurate system. If I continued to design loudspeakers that never reveal time-domain or spatial-domain subtleties, I would never even know of the existence of such subtleties, therefore, I would never be motivated to change – thus allowing the vicious cycle to continue. It is evident, that the ear examines the incoming audio stimulus in two-stage process: (1) location – here the transient of the stimulus is examined, and (2) signal – here the spectral properties of the stimulus are examined. The two processes always work in-tandem. It is therefore essential, that the loudspeaker provides undistorted waveforms to the auditory system to enable correct processing of both stages.

So, here I am. Struggling to come out of the "frequency-domain box" and into the new world of time/frequency/space-domain characteristics of contemporary loudspeakers. But even at these early stages of adopting a new technology, I find it already very rewarding. This is because it's evident that a new, **accurate and realistic acoustic transduction technology** is being achieved in much more accessible commercial way.

Thank you for reading.
Bohdan

Appendix A

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

Lateral information (left, ahead, right)

For determining the lateral input direction (left, front, right) the auditory system analyzes the following ear signal information:

- Interaural time differences
  Sound from the right side reaches the right ear earlier than the left ear. The auditory system evaluates interaural time differences from
  - Phase delays at low frequencies
  - group delays at high frequencies
- Interaural level differences
  Sound from the right side has a higher level at the right ear than at the left ear, because the head shadows the left ear. These level differences are highly frequency dependent and they increase with increasing frequency.

For frequencies below 800 Hz, mainly interaural *time* differences are evaluated (phase delays), for frequencies above 1600 Hz mainly interaural *level* differences are evaluated. Between 800 Hz and 1600 Hz there is a transition zone, where both mechanisms play a role.

Localization accuracy is 1 degree for sources in front of the listener and 15 degrees for sources to the sides. Humans can discern interaural time differences of 10 microseconds or less.[5][6]