Perception of Temporal Response and Resolution in Time Domain

Workshop #3, Berlin, 2017

Temporal Coherence

Introduction of panel members

- David Griesinger
- Mike Turner
- Menno van der Veen
- Hans van Maanen

Contents of Workshop

- Background
- Anecdotal evidence
- Perceptional test of tweeters / temporal responses
- The “mid range” by David Griesinger
- The “low range” by Mike Turner
- Consequences for microphones and loudspeakers
- Plans for follow-up workshop
- Discussion

Background

- Sound reproduction still a big issue when compared with (non-electronic) live music
- Differences will be illustrated by anecdotes
- Many aspects defy explanation with common theories
- Theories incorrect and / or incomplete?
- Which aspects are overlooked / ignored / underestimated?
- Perceptual tests

Anecdote 1: In the HiFi shop

- In the late 1970's, I wanted a better set of loudspeakers
- In those days, transmission lines were rather popular
- In the shop, a whole set of transmission lines *<u>from the same manufacturer</u>*, but with different sizes, were lined up
- I listened to the largest and the one just a bit smaller type
- Their frequency responses were virtually identical, except for the lowest frequencies
- Surprisingly, these sounded completely different, not only at the lowest frequencies
- The shop owner agreed completely with my findings, but had no explanation
- Flabbergasted, I left the shop to think about this

Anecdote 1: In the HiFi shop

- In the end, my conclusion was that, accepting that the frequency responses could not explain the perceptual differences, it *had* to be the *temporal properties*
- This was my first step to have a better look at the temporal response of audio systems

Anecdote 2: The pick-up cartridge

- In the late(r) 1970’s, I used long-play grammophone records as my primary, high quality source for music reproduction at home
- In those days, a hefty discussion was going on about moving magnet (MM) and moving coil (MC) cartridges
- MM *measured* better, MC *sounded* better (according to the reviewers). How come?
- Measurements revealed that the MM’s used a resonance of the magnet on the stylus to extend the frequency response to 20 kHz, MC’s did not, these went up to > 50 kHz without mechanical support of resonances
- The resonance of the MM’s degrades their ability to resolve signals of higher frequencies in time domain

Anecdote 2: The pick-up cartridge "as is"

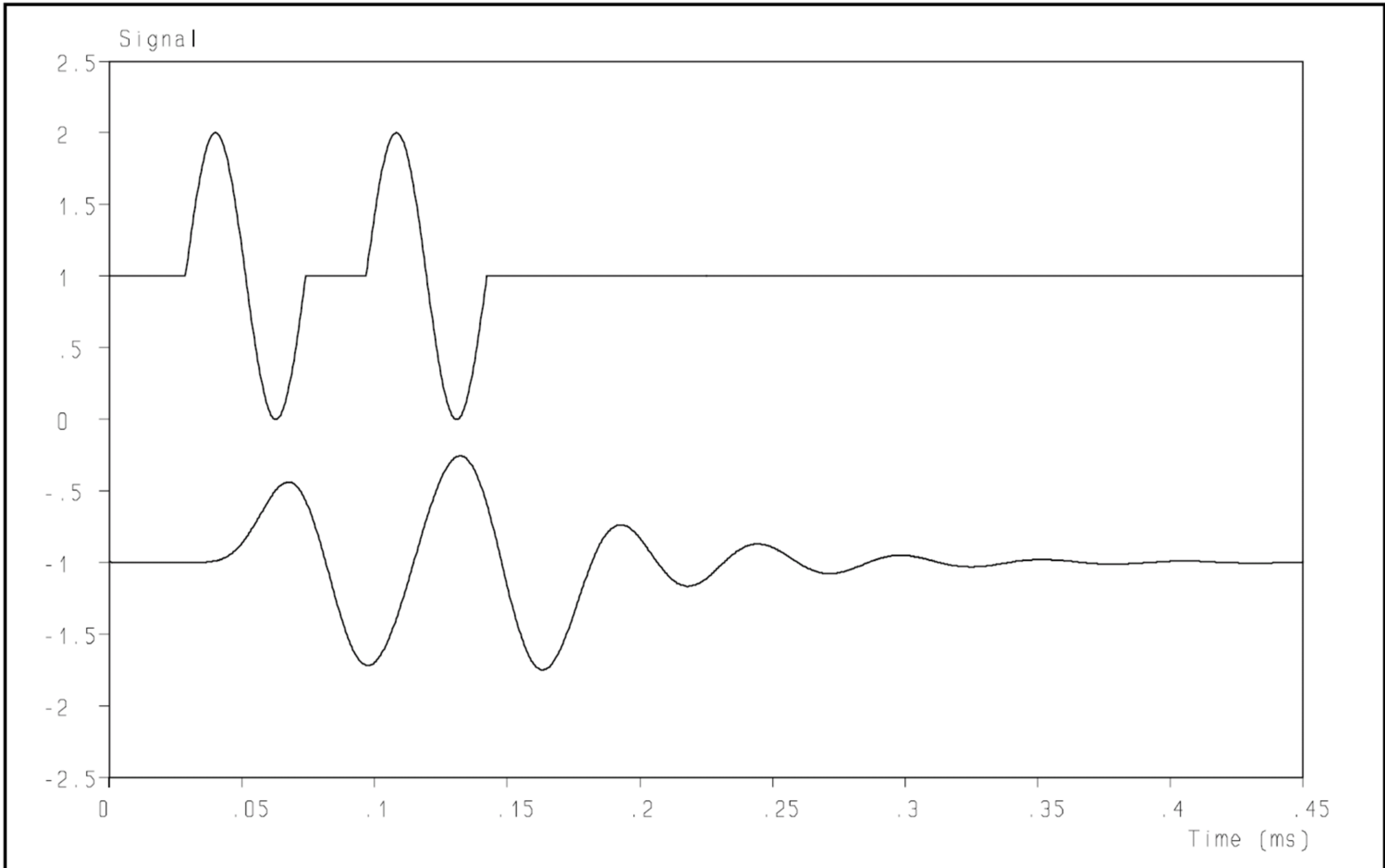
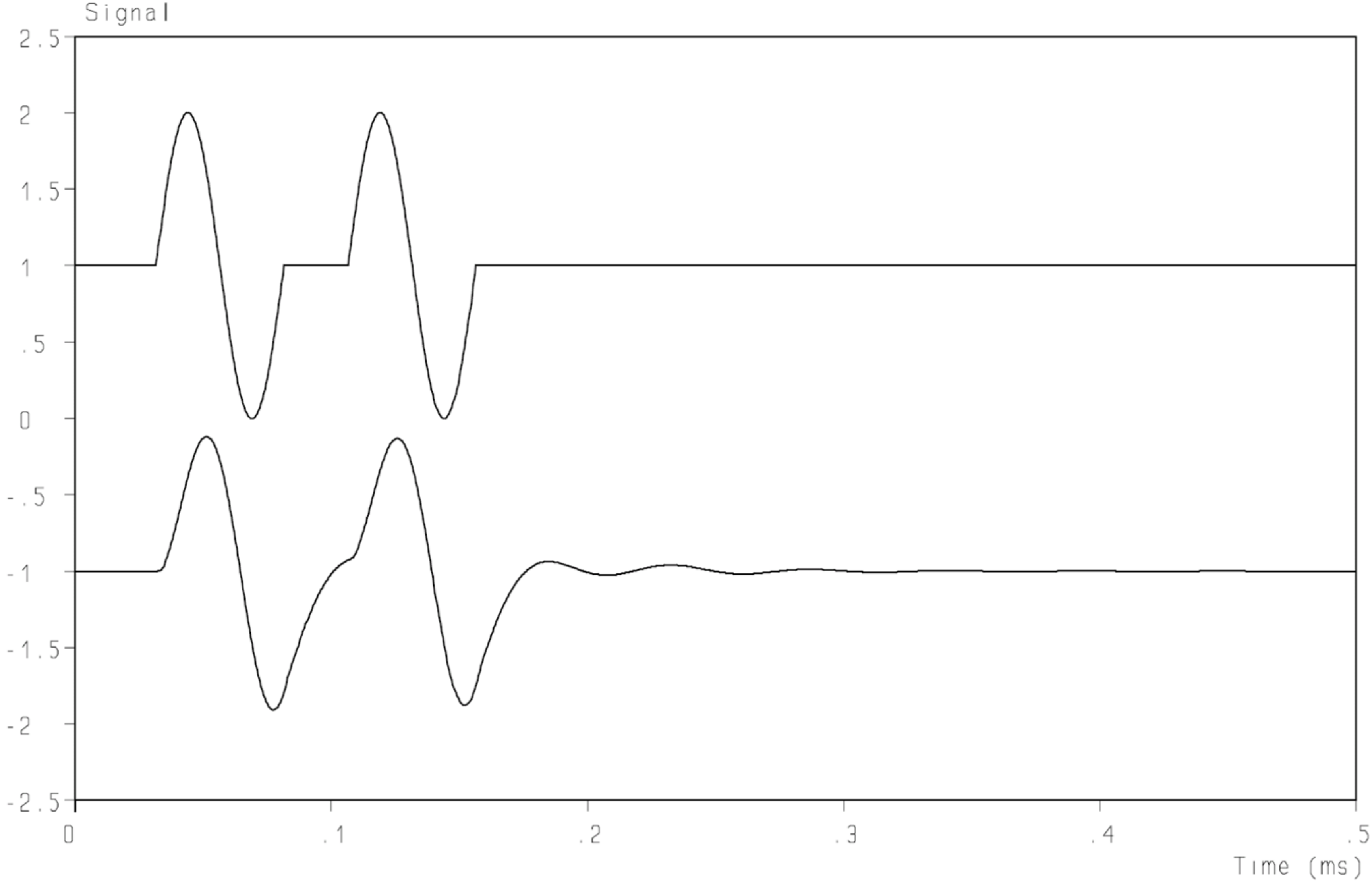


Figure 2: Response of a moving magnet cartridge to a complex signal.

Anecdote 2: The pick-up cartridge

- Electronic correction brought a major improvement in the *perceived* quality of the same cartridge, even though its frequency response < 20 kHz was still the same. How come?
- The better resolution in time domain hints at the answer:

Anecdote 2: The pick-up cartridge after correction



Anecdote 2: The pick-up cartridge

- The most logical explanation was the improvement in its temporal response and the resolution (detail) in time domain

Both anecdotes indicate that the requirement “20 Hz – 20 kHz” is *insufficient* to explain the perceived differences between the loudspeakers and cartridges

Anecdote 3: The CD reconstruction filter

- With the coming of the CD (44.1 kHz / 16 bit), also steep reconstruction filtering was introduced
- Although all filters are flat up to 20 kHz, there are perceived differences
- The behaviour > 20 kHz is usually different, but those frequencies cannot be heard by humans, certainly not seniors, like myself, isn’t it?

Anecdote 4: Own experiences

- Frequent visitor of the Concertgebouw in Amsterdam for live concerts of classical symphony orchestra (mostly)
- Cymbals and triangle sound distinctively different (read: clearly better, more resolved) compared to reproduced versions even though the higher frequencies should have been damped by the air humidity and distance (literature: 30 m ≈ 17 dB damping at 20 kHz)
- Best reproduction of metal percussion instruments I ever heard was by an ionophone
- Able to hear 15 kHz filter although hearing is limited to 11 kHz ☹

Anecdote 4: Own experiences

- After the experiences of the 1970's, my development of audio equipment has always taken the consequences in time domain into account
- *Everything* I have done to improve the temporal response had a positive impact on the *perceived* quality
- This includes the low frequency response (electronic correction for the woofer response), the mid frequency response (correct temporal response of the cross-over filters) and the high frequencies (MM cartridge correction, tweeter impulse response)

Anecdote 5: Book of G.A. Briggs

- In the book by G.A. Briggs (“Loudspeakers”), the founder of the Wharfedale factory, he describes a test in which two senior listeners (limited to 10 and 11 kHz) were unambiguously able to tell whether a 12.5 kHz LP filter was active or not (my copy is from 1963)

Question: what are the spectral distribution and the temporal properties of instruments like cymbals, triangle, piano? Or, is there life above 20 kHz? Boyk published measurements, some of which will be shown in the next slides

Measurements by Boyk: Cymbals

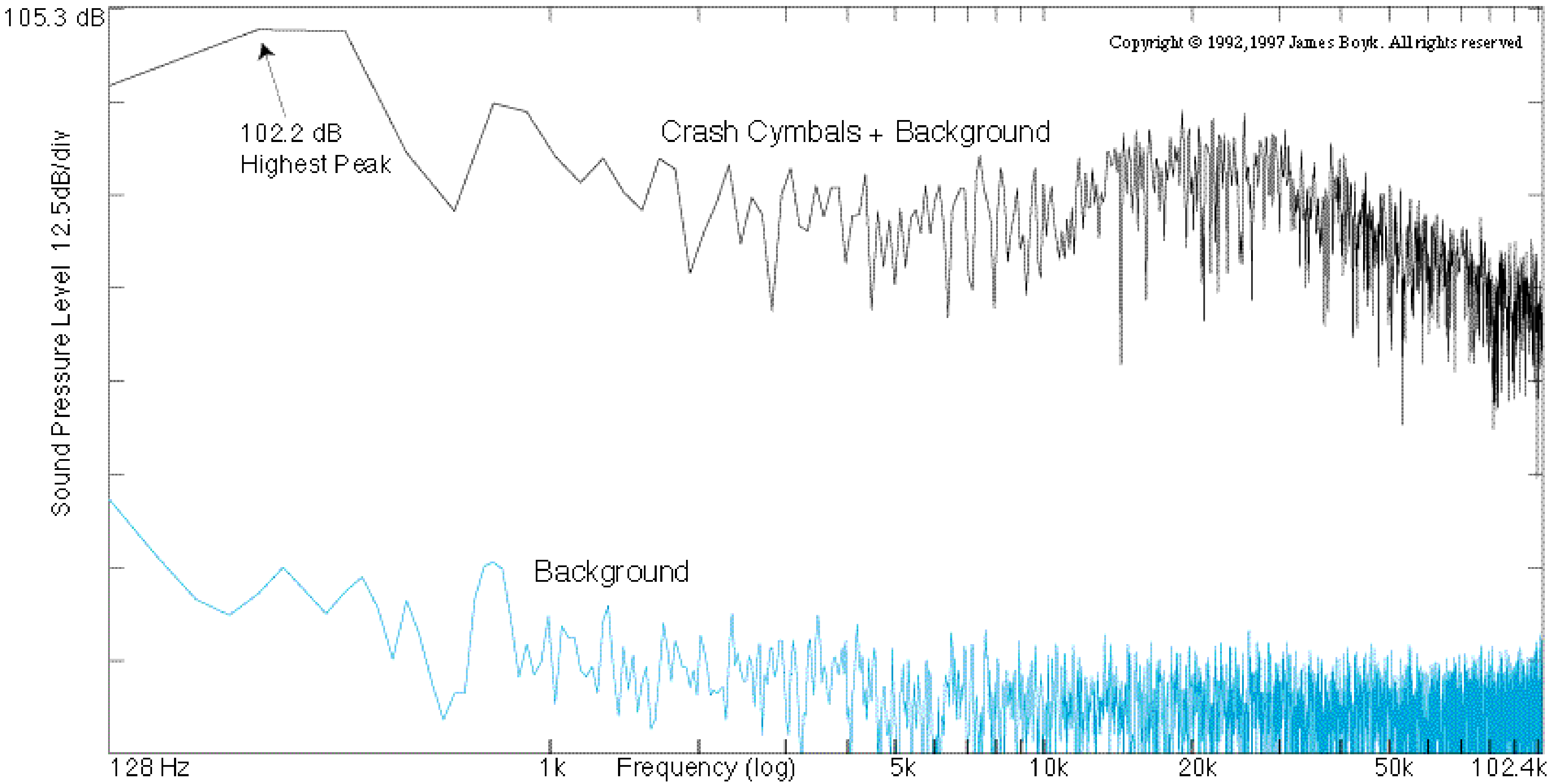


Figure 13(a) (Amplitude *vs* frequency). **Crash Cymbals**; 108.3 dB at B&K 4135 microphone with grid off, approximately one and a half feet away. Upper trace: Cymbals + background, corrected to 100 kHz. Lower trace: Background alone.

Measurements by Boyk: Claves

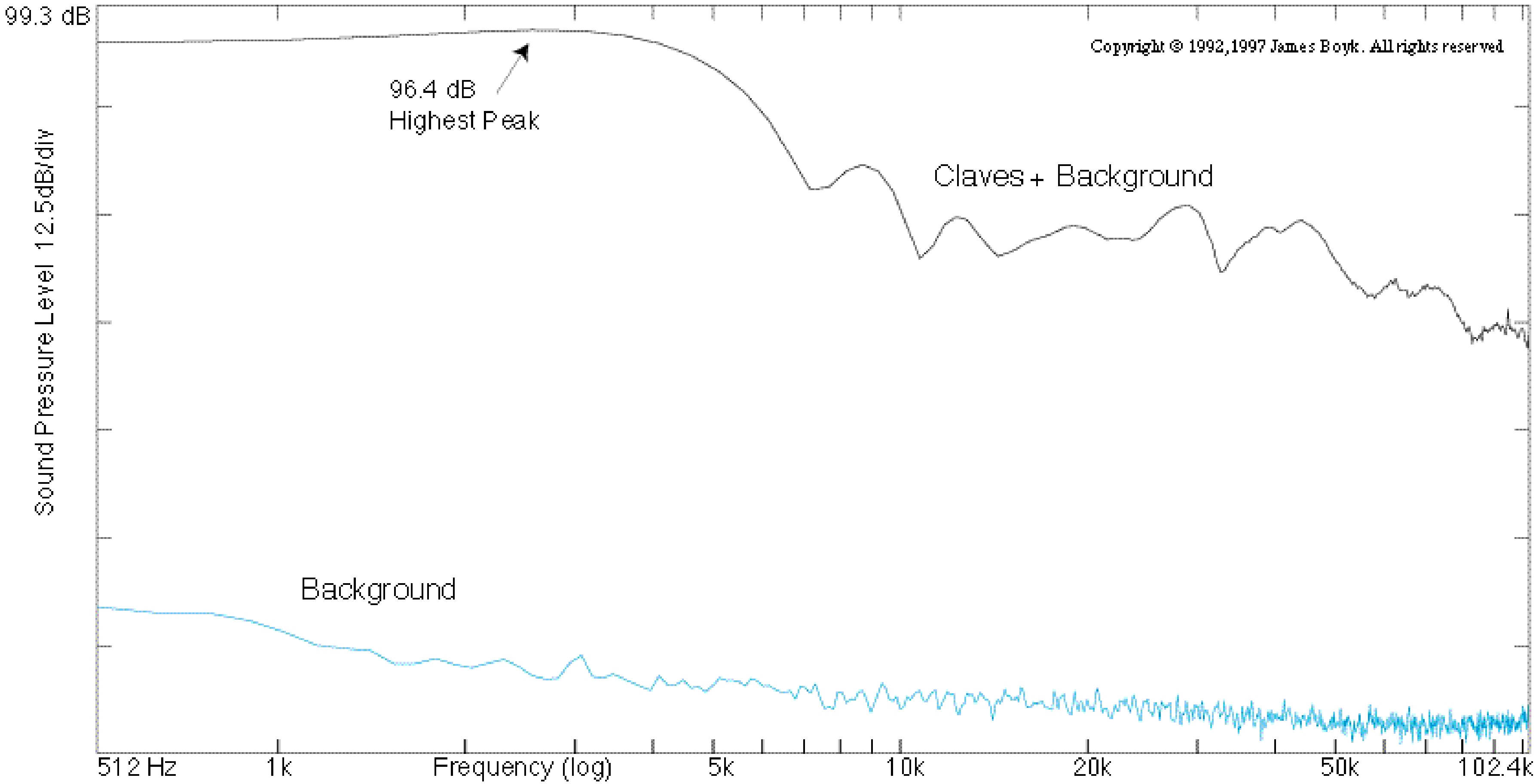


Figure 11(a) (Amplitude *vs.* frequency). **Claves**; 104 dB at B&K 4135 microphone with grid off, approximately 18 inches away. Upper trace: Claves + background, corrected to 100 kHz. Lower trace: Background alone.

Measurements by Boyk: Attack claves

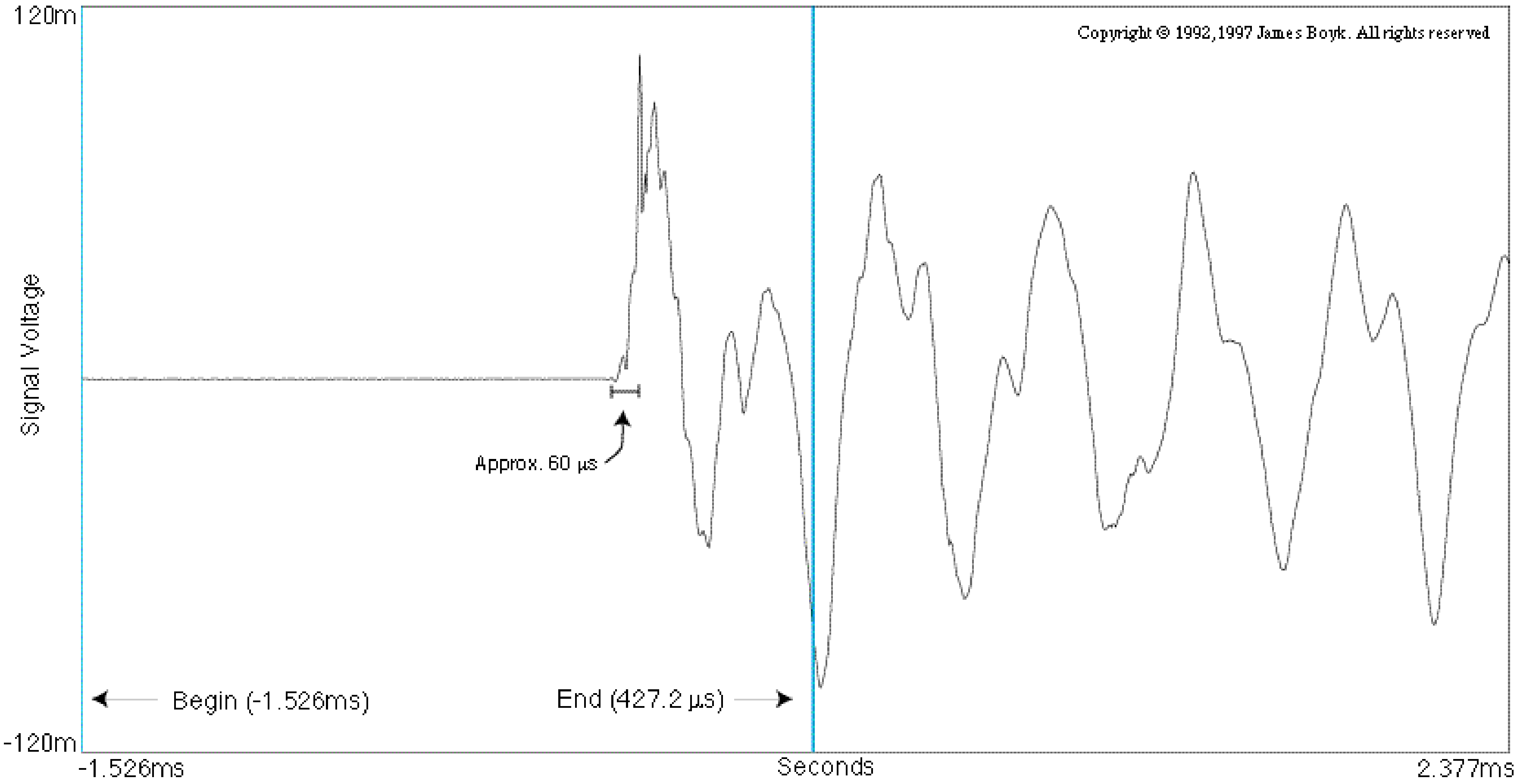


Figure 11(b) (Amplitude *vs.* time). **Claves**. The time sample analyzed to make Figure 11(a). The segment between the "Begin" and "End" marks is the 2-millisecond record from which the spectrum was created by Fourier analysis.

Measurements by Boyk: Attack piano

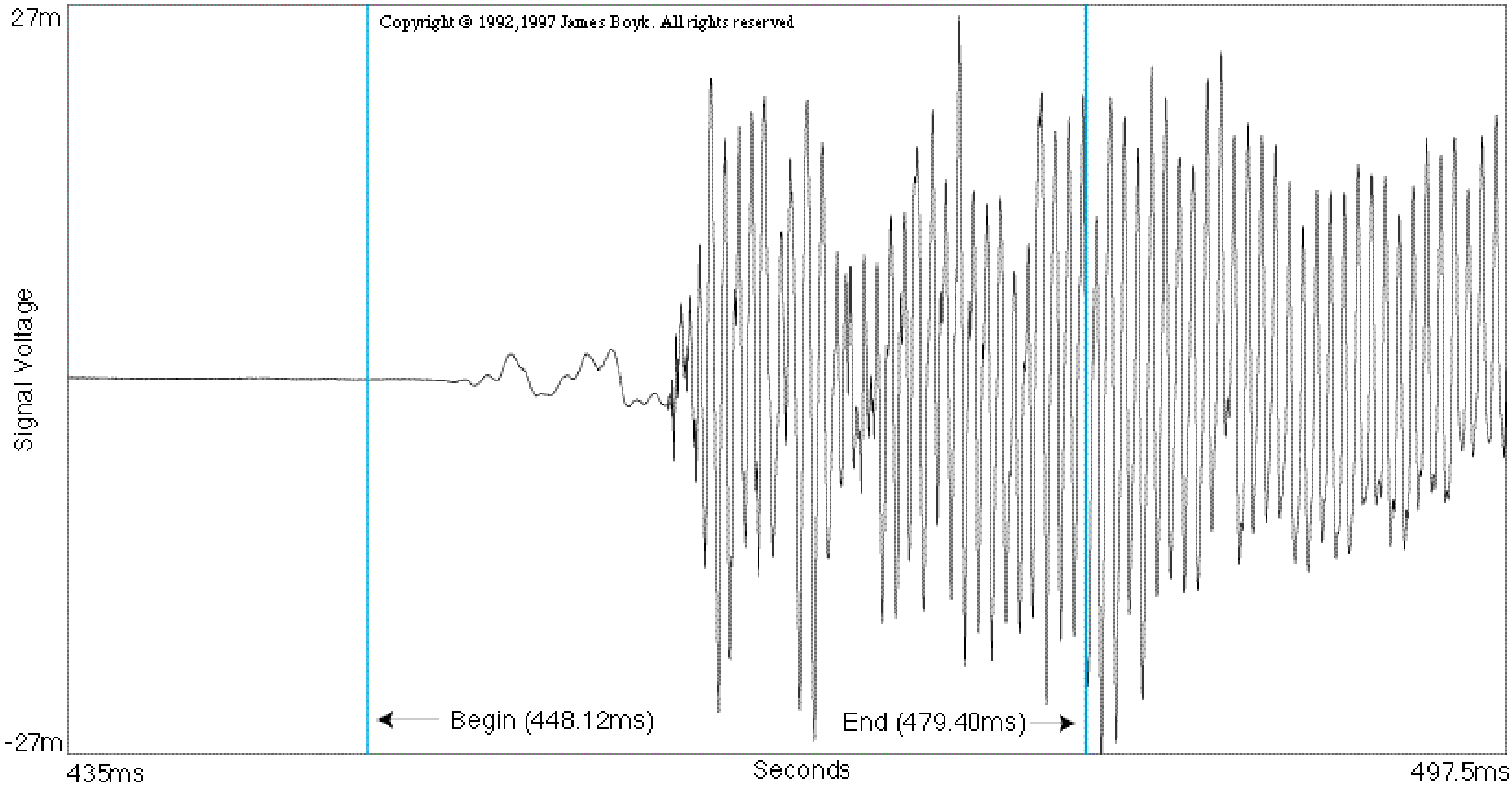


Figure 16(c) (Amplitude *vs* time). **Piano**. The sample analyzed to make Figure 16(a). The marked segment shows the 31.25-millisecond record from which the spectrum was created. Note that while the length of the record was 31.25 ms, it starts before the beginning of the sound and thus ends before the first reflection from the nearest wall. (The floor is taken as part of the instrument; see text.)

Measurements by Boyk: Attack triangle

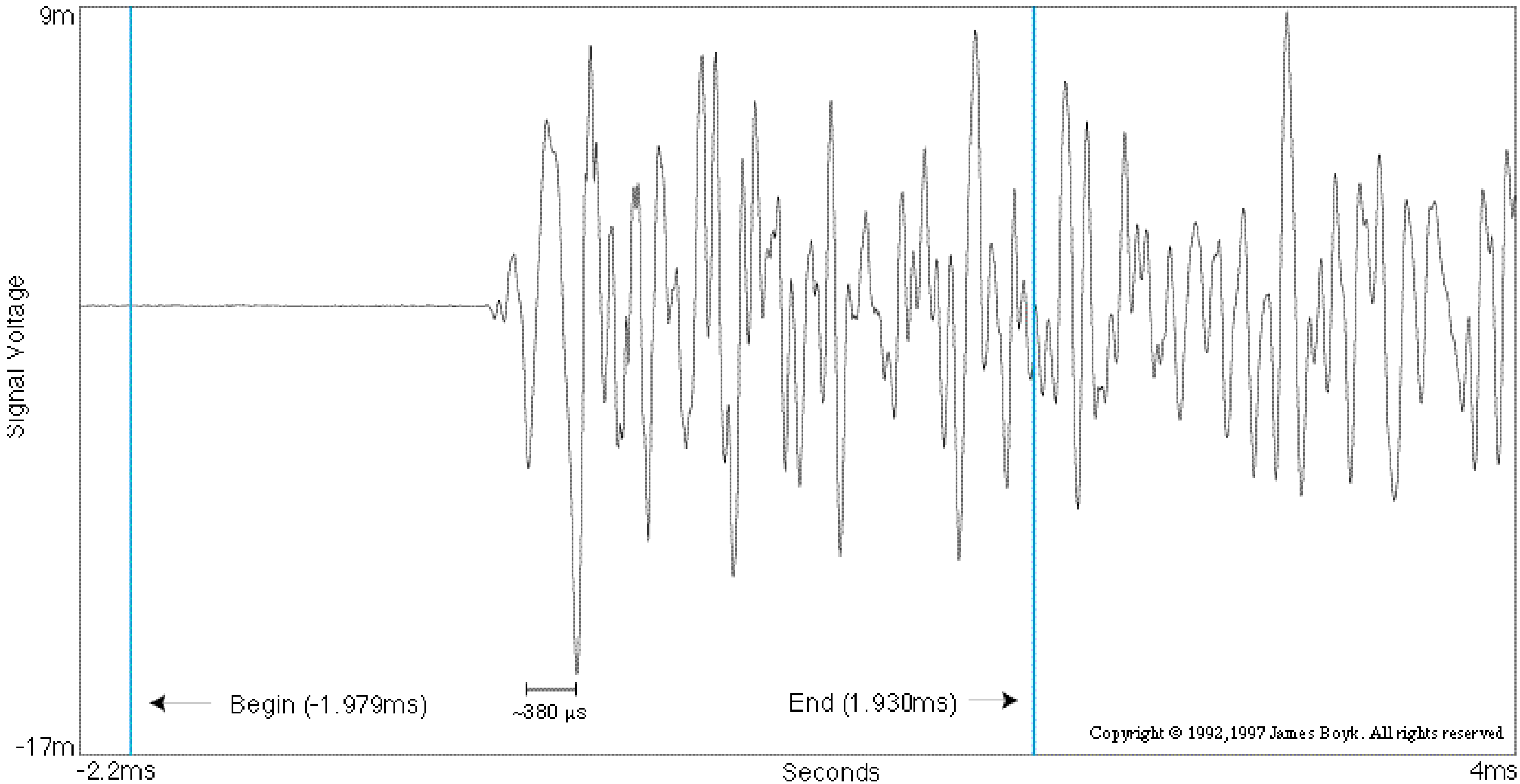


Figure 14(b) (Amplitude *vs.* time). **Triangle**. The time sample analyzed to make Figure 14(a). The segment between the "Begin" and "End" marks is the 3.9-millisecond record from which the spectrum was created.

Learnings from measurements Boyk

- Several instruments have a strong contribution above 20 kHz
- Several instruments have a strong attack, rapid change of signal at start, with very clear high-frequency content

Learnings from literature

- Attack is essential part of the specific sound of the instrument
- Instruments with a strong attack are the toughest to reproduce in a “natural sounding” way
- Specific instruments: Turkish drum, percussion, (grand) piano, cymbals, triangles
- But also human voices

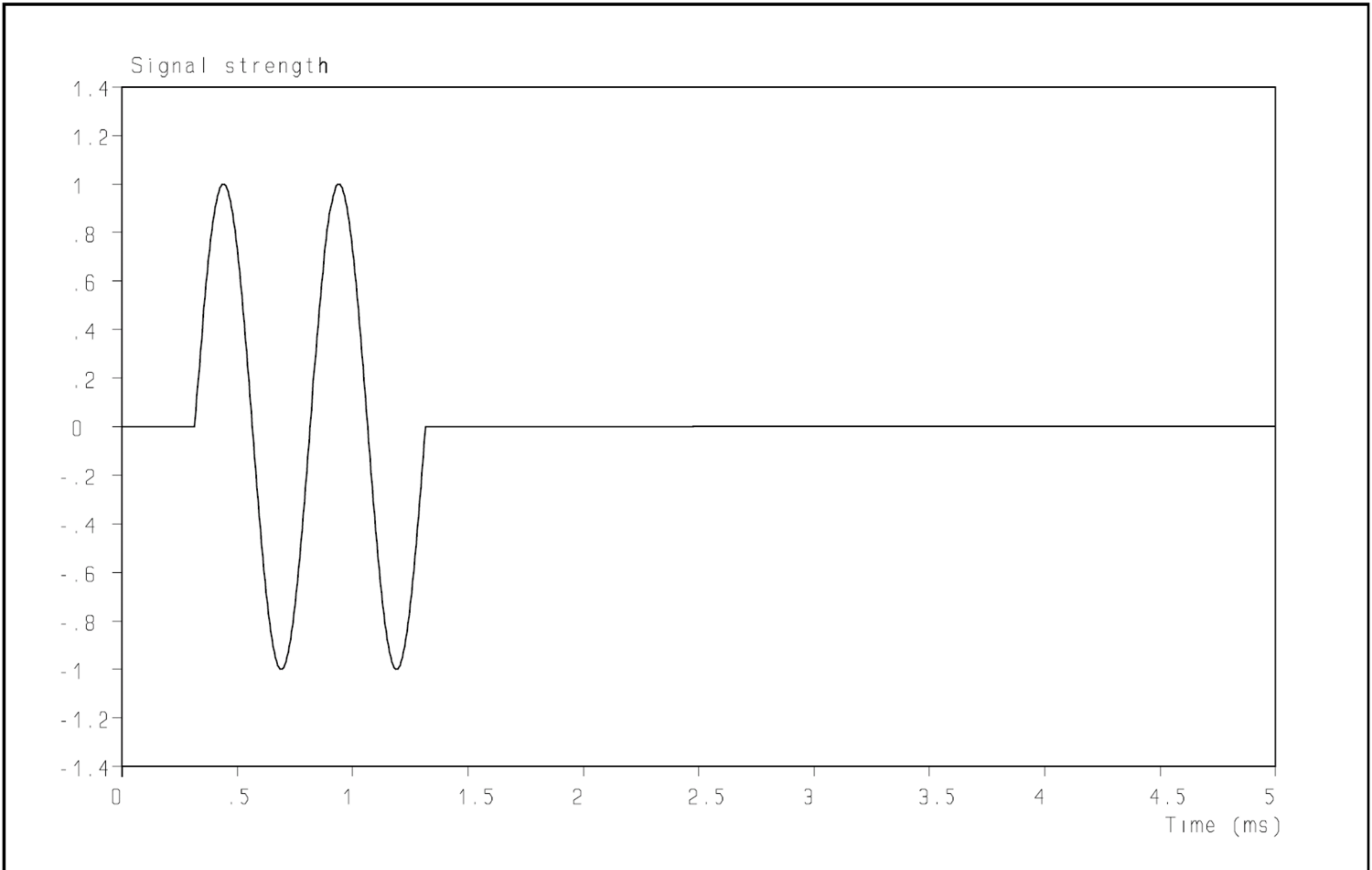
Theory

- The Fourier theory is one of the fundamental basics on which the whole sound reproduction building rests
- It says that any signal can be separated in an infinite series of (co)sine waves of increasing frequency
- It is known that humans cannot hear continuous sine waves above 20 kHz and the upper limit decreases with age
  (I know!)
- Tests have shown that human hearing is insensitive to the phase of *continuous* sine wave sound signals
- The common conclusion is that reproduction of sound from 20 Hz – 20 kHz <u>*with only the correct amplitude*</u> is completely sufficient for sound reproduction, indistinguishable from the original, but quite in conflict with the above mentioned anecdotal findings and with what I hear

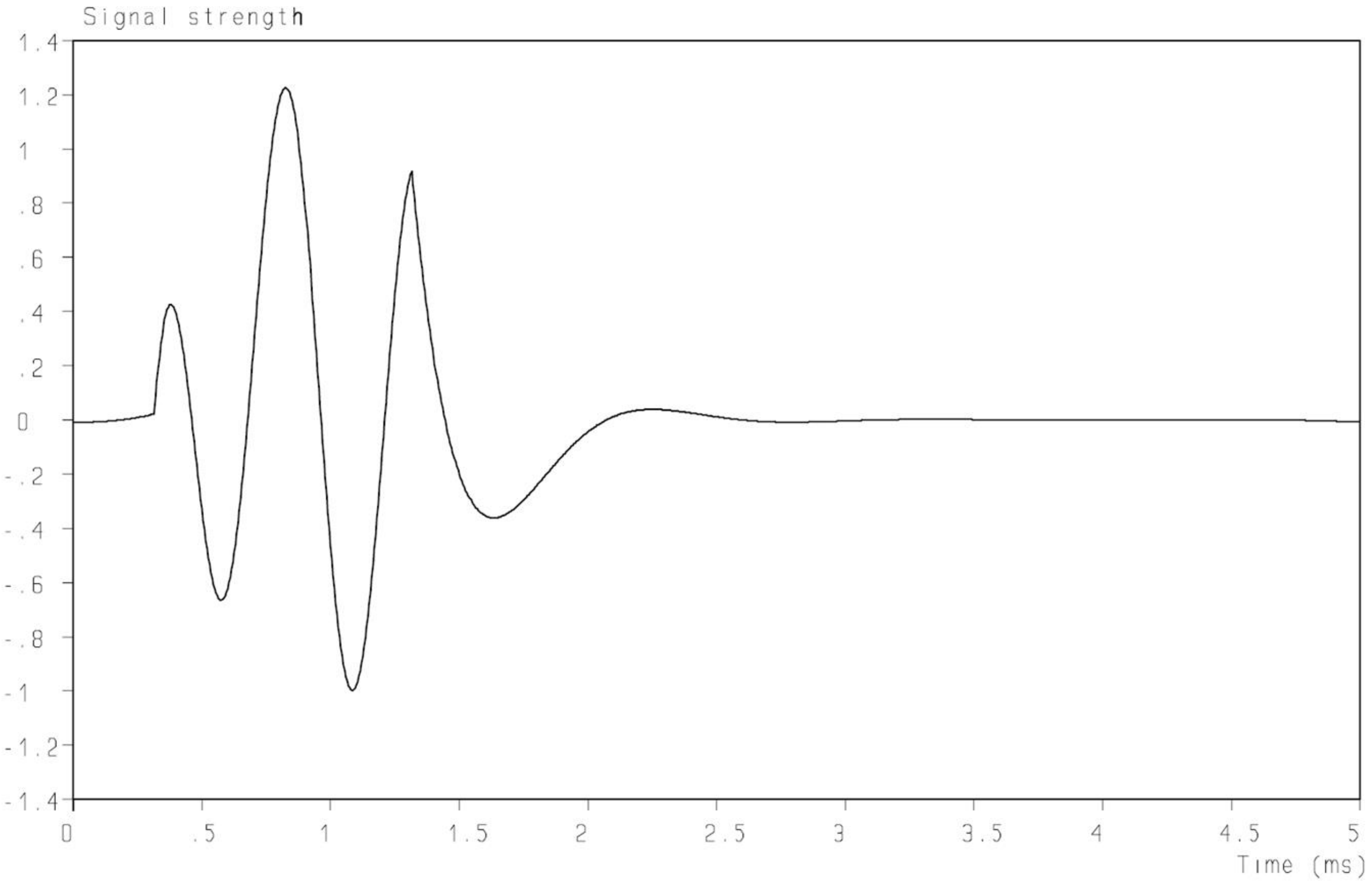
Theory

- Theory learns that to reconstruct the original signal from the Fourier components also requires the correct use of the phase
- Ignoring the phase response means that the reproduced signal can, in time domain, be different from the original, *even if the amplitudes are identical*

Theory: original signal



Theory: same modulus, different phase



Theory

- As is shown, ignoring the phase leads to a change in the temporal properties of the signal, which is clearly seen from its *envelope*
- This has consequences for e.g. the attack of percussion instruments and the grand piano

So is the change of the signal *in time domain* really inaudible?

Theory

- The anecdotes indicate that the temporal properties are of importance for the perceived quality of reproduced sound
- Tests of Kunchur indicate temporal resolution of human hearing of 5 – 6 μs (which is rather surprising with 20 kHz upper limit of hearing)
- The Fourier theory has several conditions, like a.o.:
  - the system should be *linear*
  - the system should be *time-invariant*
- Human hearing is *neither*

So is the Fourier theory directly applicable to human hearing?

Hypothesis

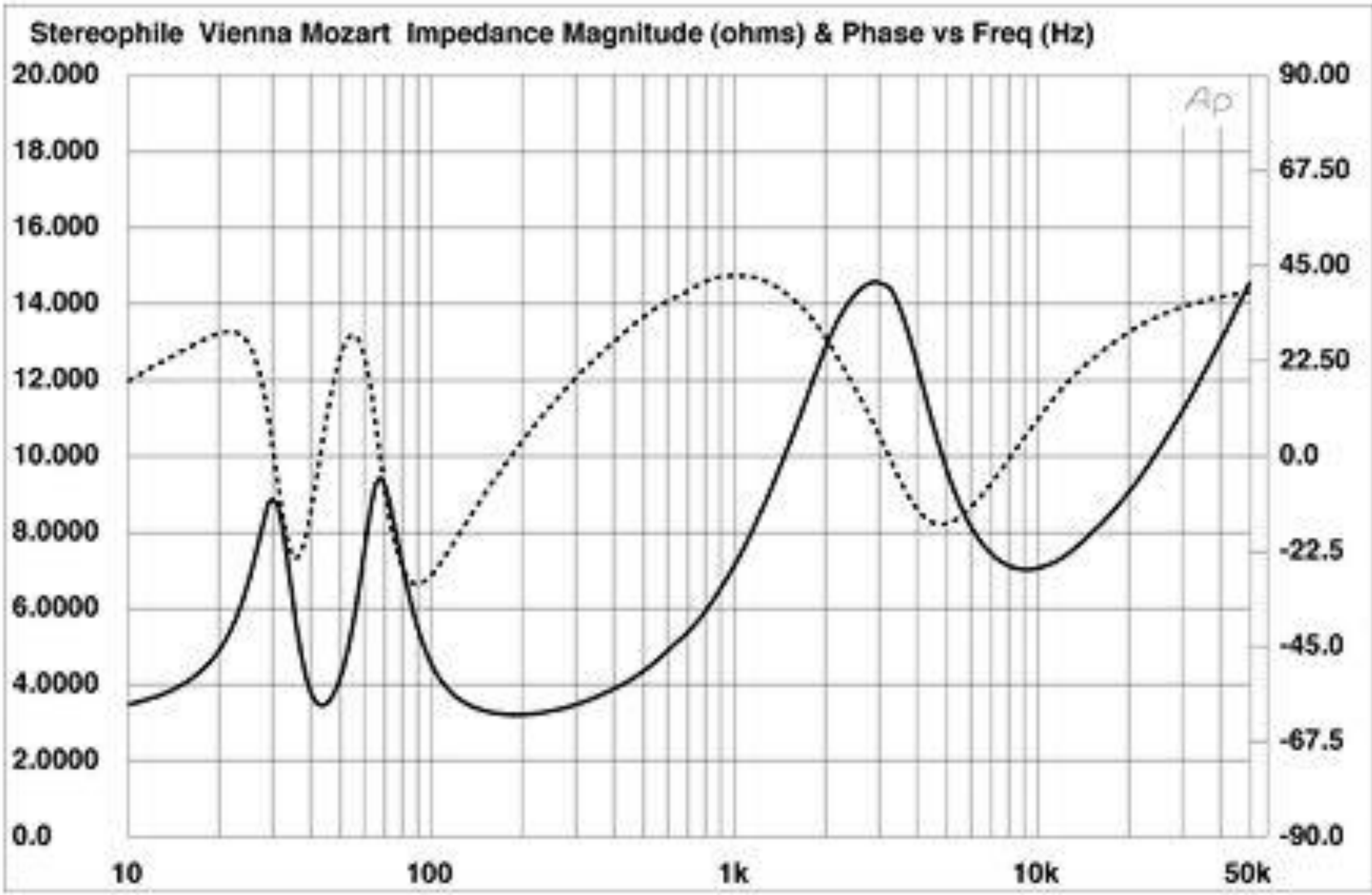
- Human hearing is sensitive for temporal properties of sound due to its non-linear properties (e.g. envelope detection)
- This shows most clearly in impulsive sounds like Turkish drum (low frequencies), attack of grand piano (mid range) and metallic percussion (high frequencies)
- Also the human voice could be effected

Current situation

- For the low frequency range very often base-reflex systems are used as these extend the response for continuous sine waves from a moderately sized cabinet.
- High-end audiophiles find that this design leads to “woolly” reproduction of the low frequencies
- As a base-reflex uses two resonances (port and woofer), it is likely to have a problem with the time response as resonances need time to start and time to decay

Current situation

- Many thanks to Stereophile:



Current situation

- For the low frequency range very often base-reflex systems are used as these extend the response for continuous sine waves from a moderately sized cabinet.
- High-end audiophiles find that this design leads to “woolly” reproduction of the low frequencies
- As a base-reflex uses two resonances (port and woofer), it is likely to have a problem with the time response as resonances need time to start and time to decay
- An acoustic box or a baffle do not suffer from timing problems, but both have a low output at lower frequencies
- Some designs use electronic compensation for AB or baffle to correct the reduced output

Current situation

Design with a baffle for the woofer and electronic correction for the low frequency roll-off

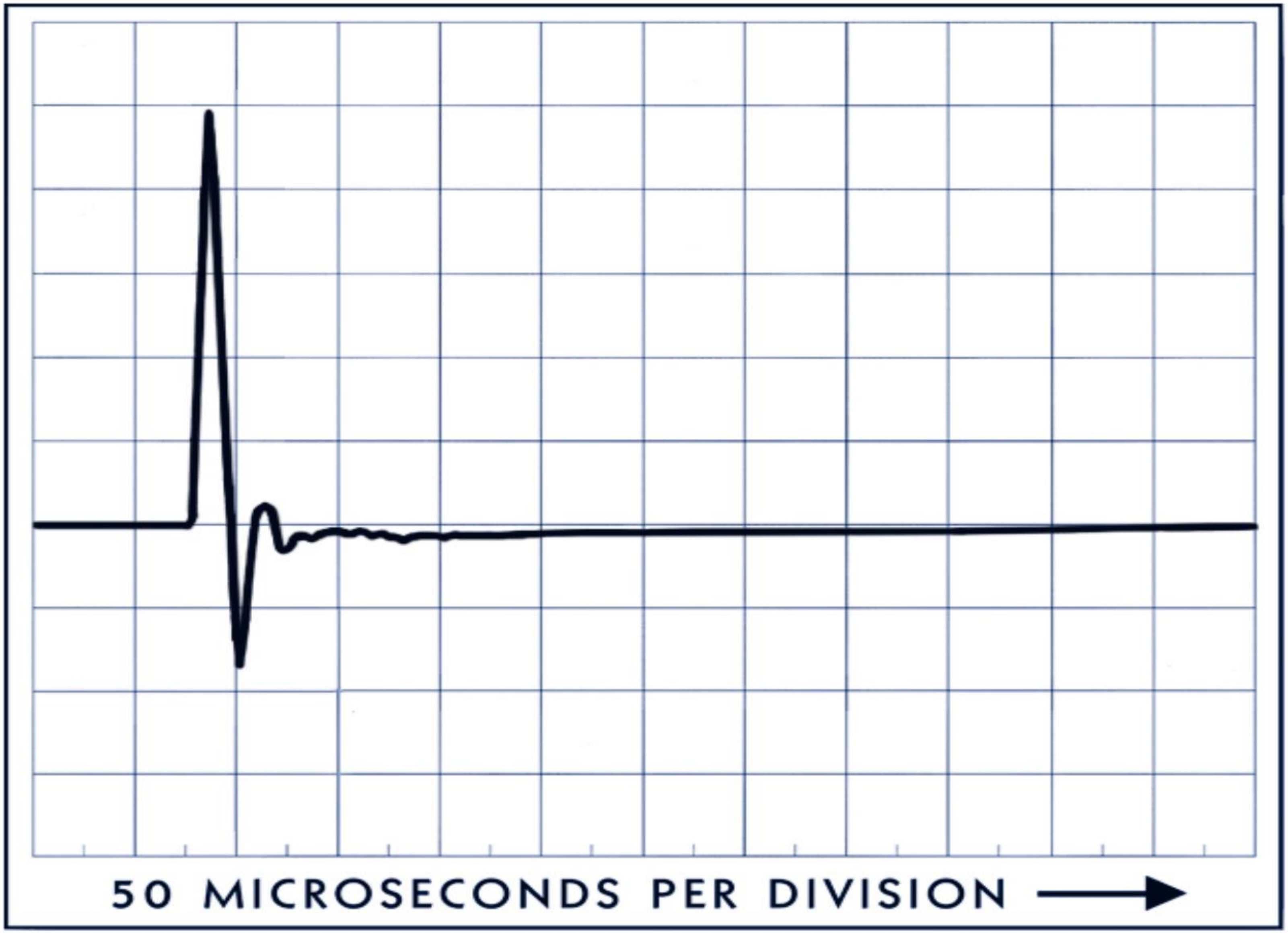
Current situation

- The midrange is effected by the cross-over filters between woofer & squawker and squawker & tweeter
- Cross-over filters, steeper than first order, influence the temporal response of the system
- Attack of grand piano, but also e.g. kettle drums, suffer
- Electrostatic loudspeakers and Magneplanars use a single unit for the midrange from relatively low up to relatively high frequencies and are rated as very good in the reproduction of the mid-ranges, including human voices
- Two-way systems can have less problems, but are not free of it (a loudspeaker unit in a housing is also a filter!)

Current situation

- The tweeter impulse response is essential for the temporal response of the loudspeaker as a whole
- Rarely, if ever, specified by the manufacturer
- But the whole chain from musician to listener is of influence
- Microphones can also suffer from a –diplomatically put- less than optimal impulse response
- The use of resonances to fill the gap to 20 kHz is just as detrimental to the perceived quality as it is with MM cartidges
- Some microphone manufacturers emphasize this and do specify the impulse response of their products

Current situation: impulse response microphone



Current situation

- Based on the linear Fourier theory, there should be no audible differences between microphones which go up to 20 kHz
- In practice, these are obvious
- The anti-aliasing filter, the (limited) bandwidth of the transmission / recording path, the reconstruction filter, etc. all impose limits on the temporal resolution of reproduced sound
- The result is loss of detail, which explains at least a part of the perceived differences between “live” sound and reproduced sound
- Improvement should be obtained when the temporal response is improved

Getting supporting evidence

- In my view, there is sufficient “anecdotal” evidence to conclude that the temporal properties of audio systems are critical for realistic sound reproduction
- Yet, there is still a lot of debate (compare with the discussions on high-resolution formats) which could benefit from additional supporting evidence
- A number of people tried to organise this for this workshop
- The original idea was to do this for the low and the high frequency ranges
- Unfortunately, Mike Turner was too busy and tied up with work, to do perceptional tests, but he has some interesting developments to report

Getting supporting evidence

- Some preliminary results on the low side can be presented:
- A base-reflex housing from Hepta Design Audio could simply be modified into an acoustic box
- The response of the AB could be corrected by an electronic circuit to obtain the same frequency response (at least the -3dB frequency), but without the additional time delay, introduced by the resonances (regard it as a “motional feed forward” approach)

Getting supporting evidence

- The bass from the corrected AB was rated as “better controlled”, “less wooly” and “faster” in comparison to the base-reflex
- As there is little discussion about the temporal resolution of human hearing in these ranges, this can easily be understood and accepted
- Further tests are planned for a follow-up workshop, Mike will tell more about this in his contribution

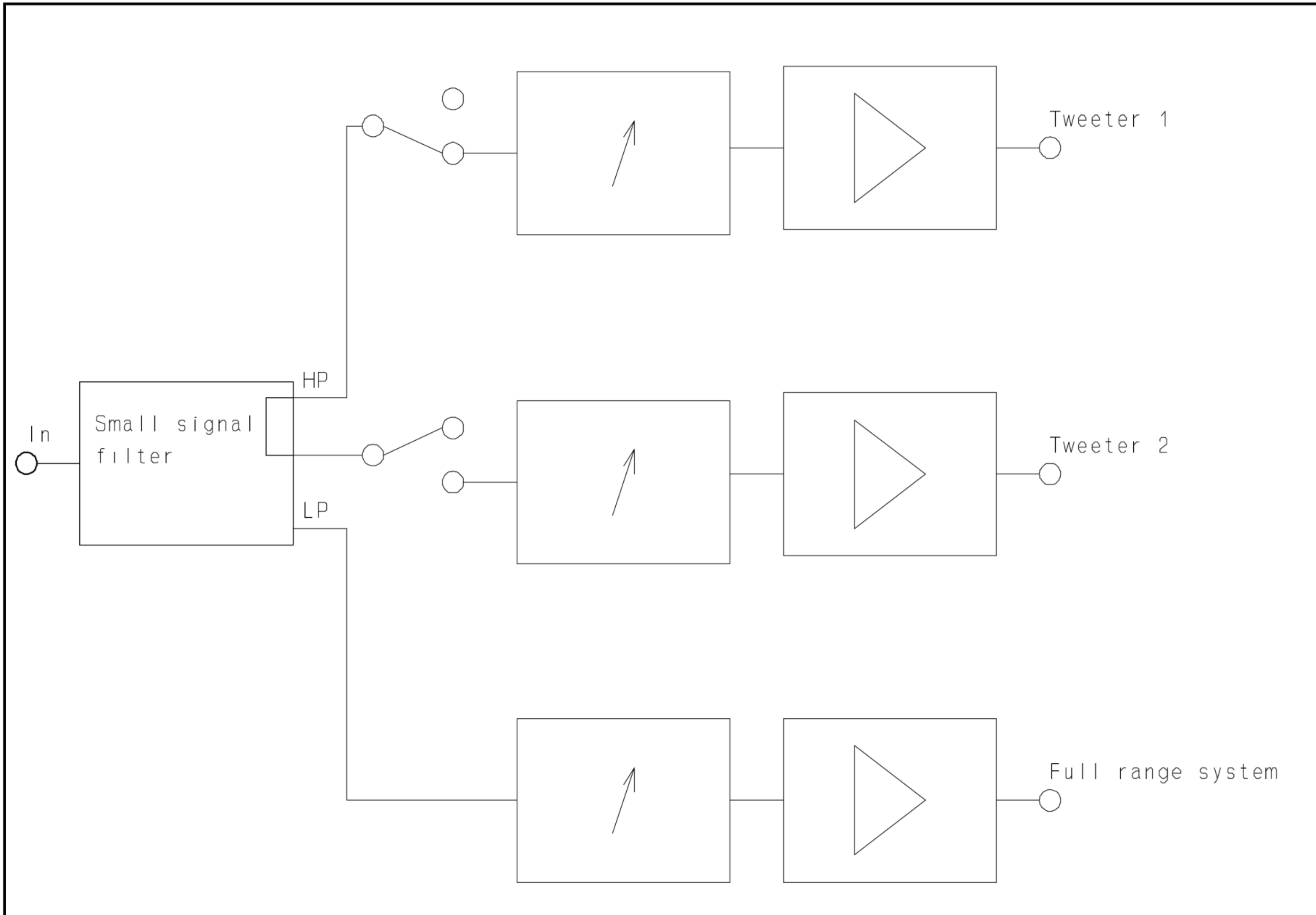
Getting supporting evidence

- For the high frequency side, we took the following actions
  - three different tweeters were measured by Geoff Hill, using a B&K 4135 at 192kHz sampling rate via a 100kHz amplifier (many thanks, highly appreciated)
  - the tweeters were mounted in small baffles
  - in pairs, these were compared for their perceived quality
  - the low and midranges are covered by a full range system (Hepta Design Audio, Emmarantus)
  - each unit has its own control and power amplifier
  - sound balancing is done within 1 dB

Getting supporting evidence

- **N.B.** The choice of the tweeters was based on historical grounds and is, of course, only a very small sample of the available types of tweeters. Therefore, we will not mention names or brands
- **N.B.** The aim of these tests was solely to see whether a correlation could be found between the *perceived* quality of tweeters and their measured temporal properties. It is hoped that the findings can help developers of loudspeaker units and microphones in their quest for better products.

Test set-up



In
Small signal filter
HP
LP
Tweeter 1
Tweeter 2
Full range system

Getting supporting evidence

- The cross-over filter was created with passive components only, operated at a low power level
- No electronics in the filtering, which could be of influence
- No passive filter in series with the tweeters which could influence the tweeter response, tweeters were used in the same way as their impulse responses were measured
- The cross-over frequency was chosen sufficiently high to accomodate all three tweeter types
- The control and power amplifiers were from “Temporal Coherence” and were identical for all three branches

The test set-up at Hepta

The test set-up at Hepta

A more detailed view of the loudspeaker / tweeter configuration

The listening team

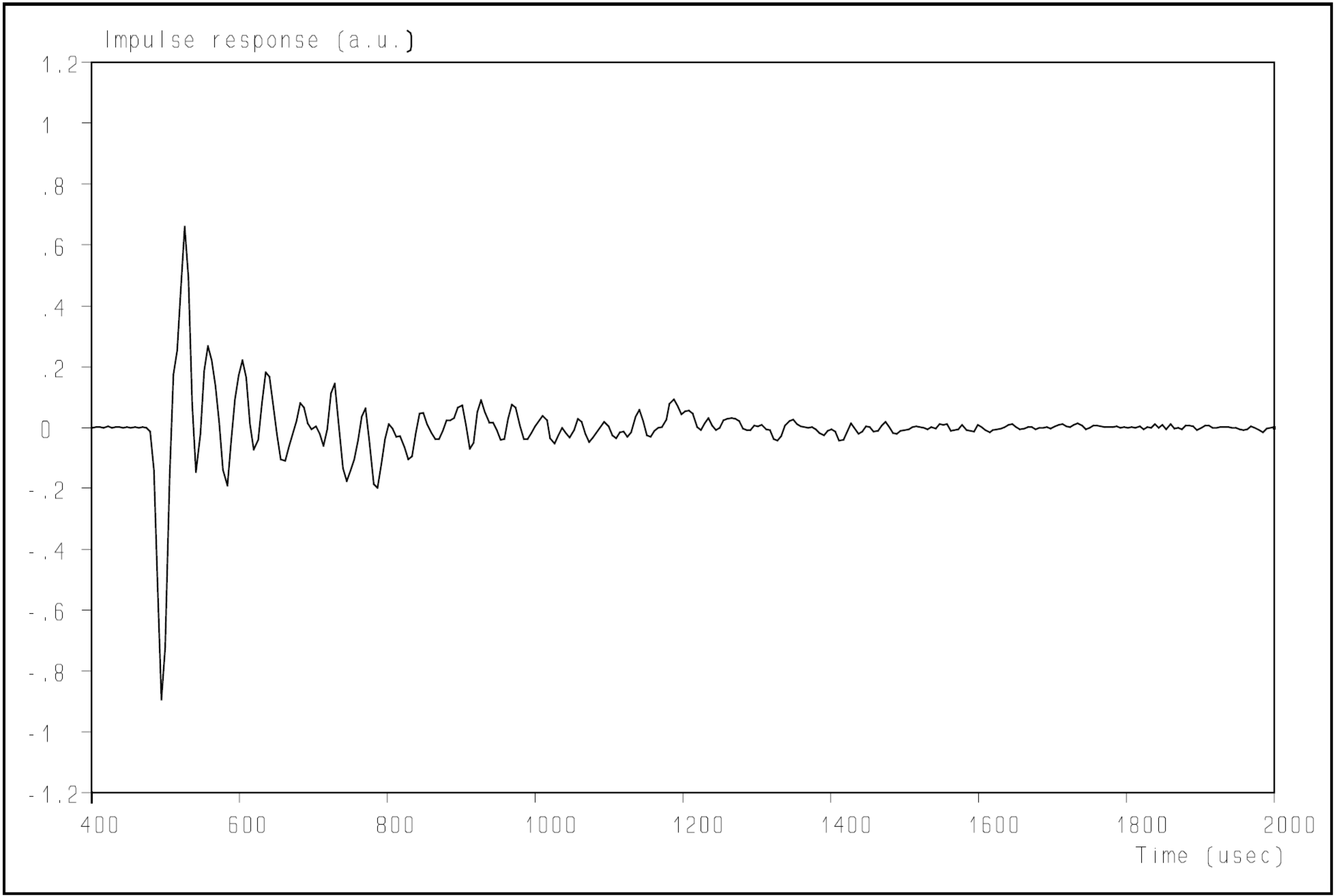
- A listening team of 8 people was formed
- All were rather senior
- Most had extensive experience with listening to reproduced music
- Upper frequency limit was 10 – 11 kHz for most
- Signal source was an SACD player (Denon 2010 AE)
- Only SACD recordings were used for testing
- Only “mechanical” instruments were used
- Only recorded music was used, no artifical sounds or signals
- The team members did not have any information about the impulse responses of the different tweeters

The listening team

The listening team

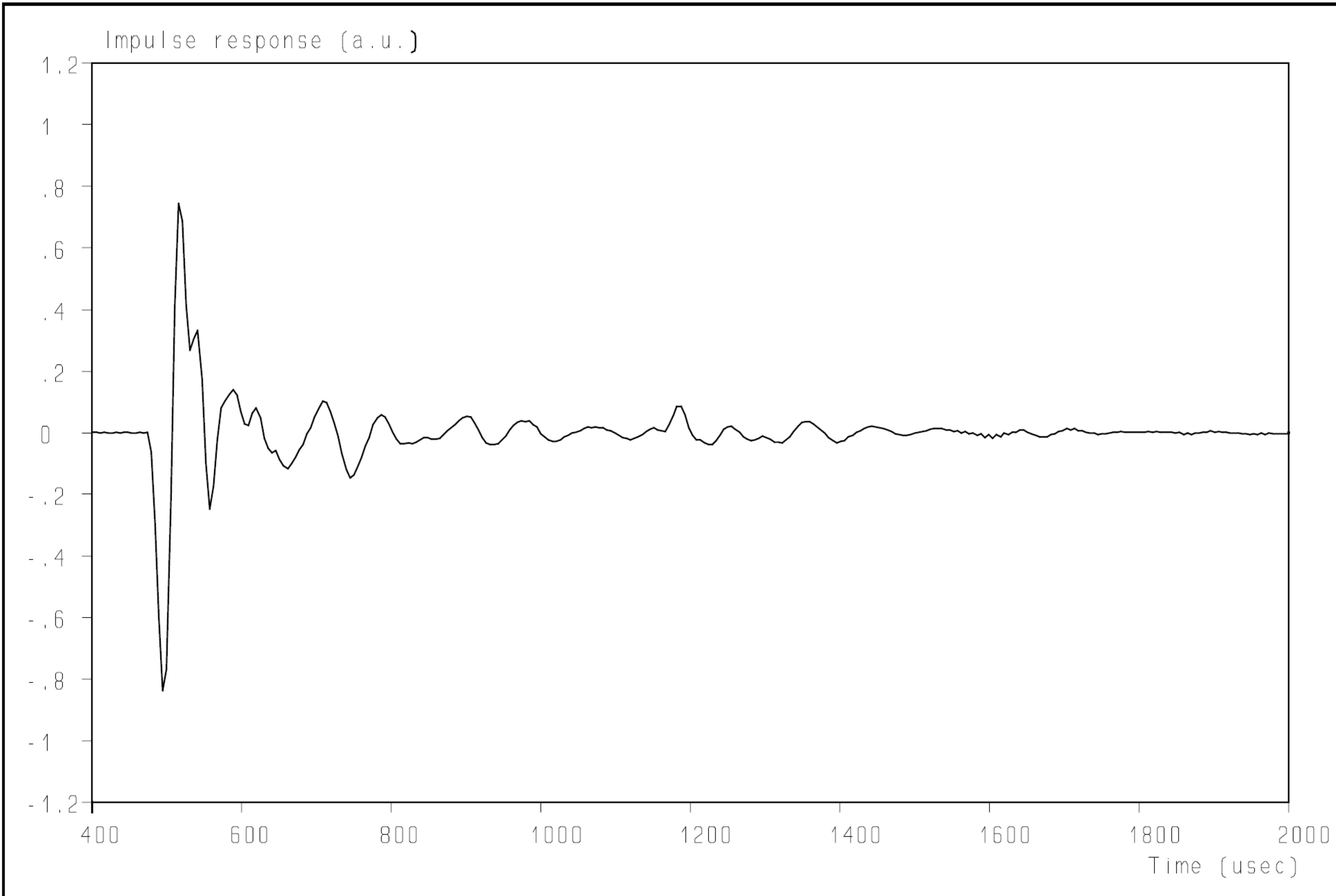
- It was not the intention to obtain quantitative results, the idea was to see whether a *correlation* between temporal properties and *perceived* quality could be discerned

Tweeter A



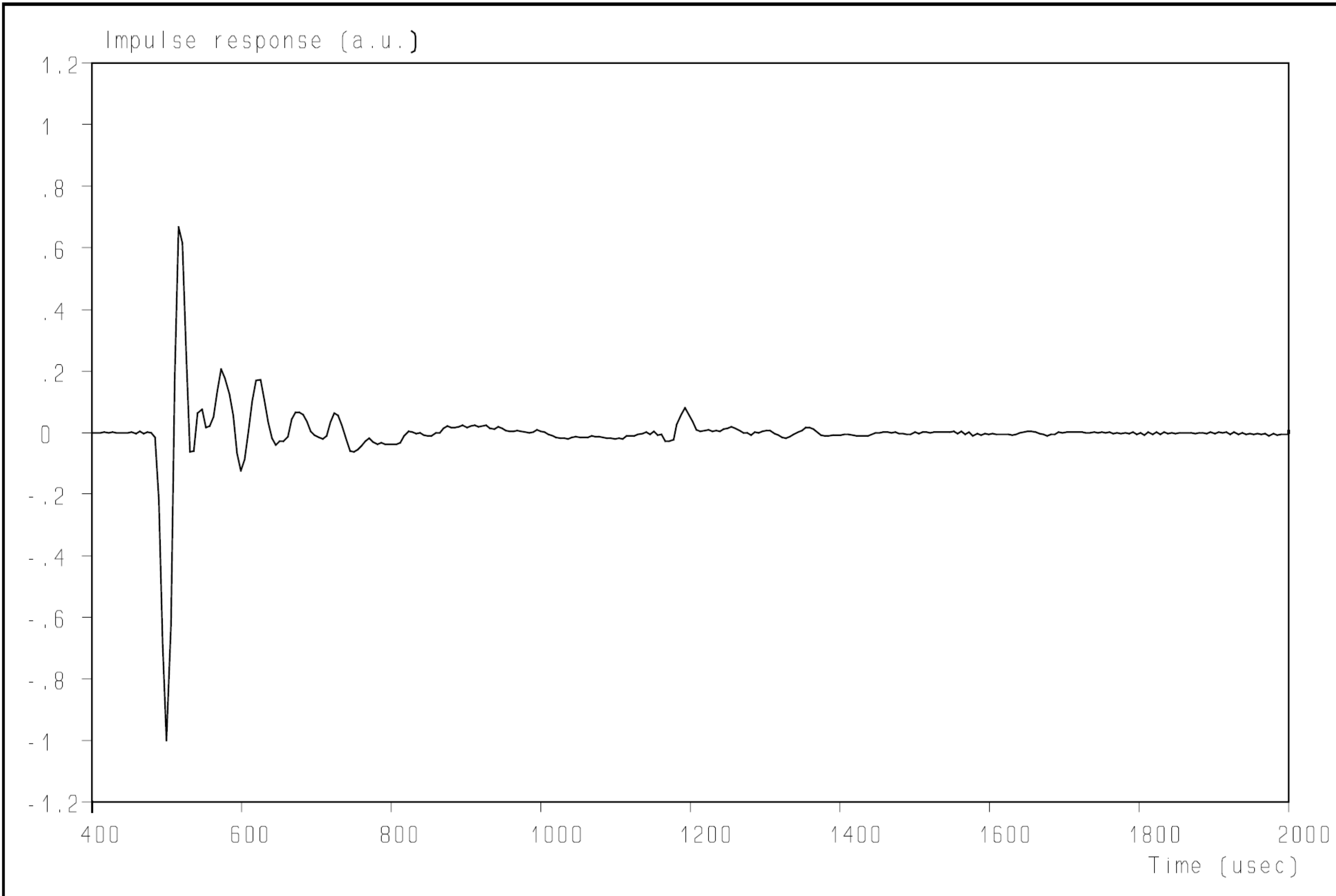
Impulse response (a.u.)
1.2
1
.8
.6
.4
.2
0
-.2
-.4
-.6
-.8
-1
-1.2
400
600
800
1000
1200
1400
1600
1800
2000
Time (usec)

Tweeter B



Impulse response (a.u.)
1.2
1
.8
.6
.4
.2
0
-.2
-.4
-.6
-.8
-1
-1.2
400
600
800
1000
1200
1400
1600
1800
2000
Time (usec)

Tweeter C



Impulse response (a.u.)
1.2
1
.8
.6
.4
.2
0
-.2
-.4
-.6
-.8
-1
-1.2
400
600
800
1000
1200
1400
1600
1800
2000
Time (usec)

Results

Three pairs were compared: A ⇔ B

A ⇔ C

B ⇔ C

The panel members could select their preference for each pair. The preferences were combined to 1st, 2nd and 3rd place rankings. The scores are presented by the times each tweeter got the specific ranking

Results

Tweeter	# of scores	1st	2nd	3rd
Tweeter A		0	3	5
Tweeter B		1	4	3
Tweeter C		7	1	0

It should be noted that the choice for the second place was found hard by the listening team, so the preference for B over A is only marginal

Results

Additional remarks

- The choice for C was almost unanimous, the only member who chose B is not used to listening to SACD's
- The reasons to choose C were
  - detail of reproduction, clarity, open sound
  - musical, transparent, neutral, least tiring
  - better attack of cymbals and better definition

Results

Additional remarks

- Although I did not participate in the listening team, my preference, made up before I got the impulse response results, was C, B, A
- Actually, the choice to apply C in our systems was made after listening to it in another system
- Independently from us, Hepta Design Audio came to the same conclusion and applies this tweeter in their best system (the Superior Orator)
- The start (onset) of the impulse response of A was better than the start of B

Results

These results indicate that there is a correlation between the perceived quality vs. impulse response & onset of the impulse response:

Tweeter impulse response

Tweeter A: 3

Tweeter B: 2

Tweeter C: 1

Onset of impulse response

Tweeter A: 2

Tweeter B: 3

Tweeter C: 1

Results

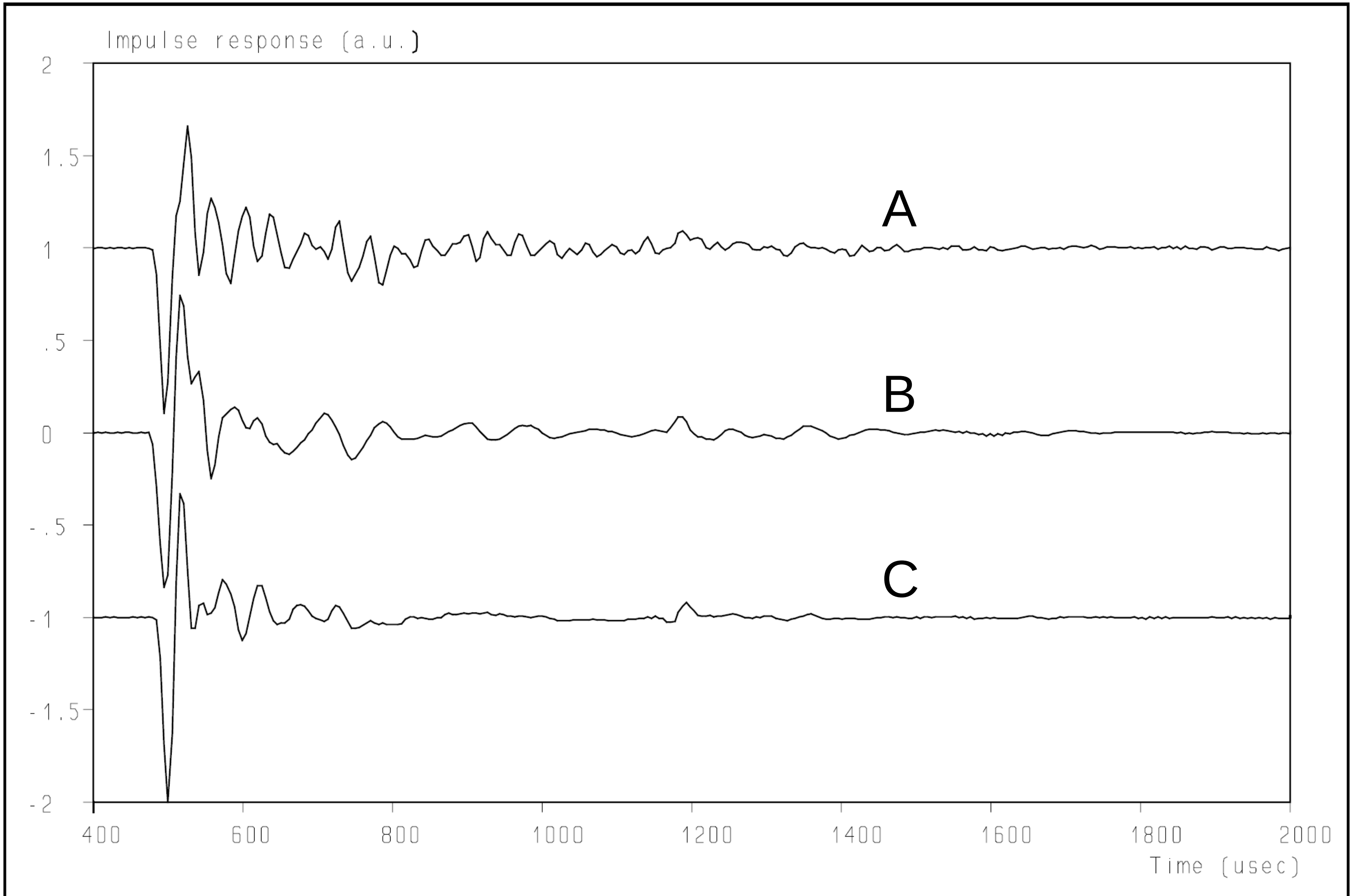


Figure 1: Impulse responses of three tweeters.

Presentations on mid and low frequencies

- Mike Turner will describe some of his work on the low frequency side
- David Griesingen will present his experiences with the focus on the mid-ranges

Conclusions

- The corrected acoustic box enclosure is the prime choice when it comes to well defined and controlled reproduction of low frequencies
- As the temporal resolution of the human hearing in these regions is undisputed, this is easy to understand

Conclusions

- The anecdotal results point at the importance of the temporal response of audio systems for high quality sound reproduction
- Taken the non-linearity and the other properties of human hearing into account, this can be understood
- Similar findings have been reported by others (both loudspeaker and microphone manufacturers)
- The tweeter with the best impulse response (and also the most extended frequency response) is clearly chosen as the best sounding (perceived quality)

Conclusions

- The onset of the impulse response also seems of influence on the perceived quality, which might explain the ambiguity between tweeter A and B
- The findings of the anecdotal experiences and the supporting test are in agreement with the findings of e.g. Kunchur and others and can explain the perceived difference e.g. between “live” and reproduced cymbals
- The results are even more convincing as the panel members were limited to < 11 kHz !

Consequences

- The simple requirement for audio systems that its frequency response should range from 20 Hz – 20 kHz is insufficient for high quality sound reproduction
- The response in time domain is at least as important, this holds for all frequencies, but notorious difficult parts are the low frequencies, the high frequencies and the mid range (for dynamic loudspeakers)
- The temporal resolution of audio systems need to be upgraded to at least 5 μs in order to become par with human hearing

Consequences

- The latter requirement means that the frequency response should be extended to at least 200 kHz with a moderate roll-off above this frequency
- This holds for microphones, recording equipment, transmission channels, amplifiers and loudspeakers and thus a major change in approach
- Resonances to “improve” the *frequency response* of any piece of equipment should be banned

Improvements

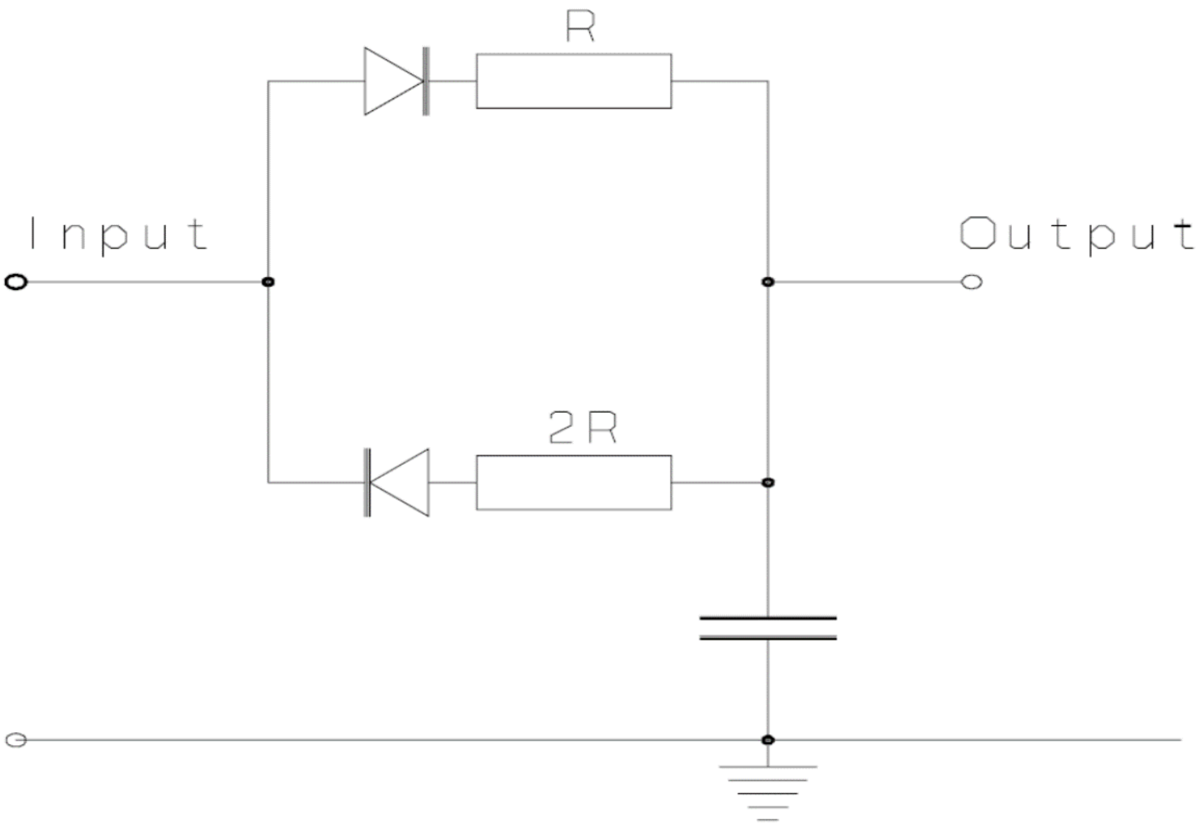
- The manufacturers of loudspeakers and microphones should pay more attention to the temporal properties of their products. This will be helpful for designers of recording and reproduction equipment to optimise their products for perceived quality
- Manufacturers of microphones, loudspeakers and headphones should specify the impulse response of their products as this would be very helpful for the selection of components and the combination

Improvements

- Developers of audio equipment should take the overall temporal properties into account to obtain the best possible temporal resolution (which e.g. can be quantified by the “temporal decay”)
- During the education, more attention should be paid to the relation of spectral and temporal properties of systems
- The theory of non-linear systems and the consequences in time domain should be developed further

Improvements

Note that the current theory is unable to predict the temporal response of this circuit:



Plans for follow-up workshop

- Simulate electronically the low-frequency response of base-reflex, acoustic box (with and without compensation), baffle and others and compare these by listening tests for their perceived quality
- Simulate electronically different cross-over filters in the midrange and compare these by listening tests for their audibility
- Extend the tweeter comparison test with other tweeters (preferably, those based on a different concept)

Plans for follow-up workshop

- Compare microphones with different impulse responses for their perceived quality

Discussion

Questions?

Remarks?