TEMPORAL DECAY: A USEFUL TOOL FOR THE CHARACTERISATION OF RESOLUTION OF AUDIO SYSTEMS?

Drs. H.R.E. van Maanen

TEMPORAL COHERENCE

Any band-limited system has a "time-smear" in its impulse response. The width of this impulse response is dependent on the bandwidth of the system <u>and of the way its band is limited</u>. This spread results in masking of fine details and the related "Temporal Decay" can vary from 0.16 to 1.09 dB/µs. The concept will be discussed and some examples will illustrate the results.

SUMMARY.

In the discussion about the perceived quality of sound systems the temporal aspect is often neglected or its importance underestimated. In this paper we propose a semi-quantitative property of systems to compare these, taking the temporal behaviour into account. We have tried to find a simple, easily to find and to interpret parameter which by no means will be the final answer to the problems encountered in audio, but can help to improve the comparison of systems in a more objective way and could help to direct future developments.

This so-called "temporal decay" is an excerpt of the Dirac delta pulse response of the linearised system. As the delta pulse response contains more information than the amplitude characteristic between 20 Hz and 20 kHz, this could be an interesting additional aspect of audio systems.

Application of this concept to various different filters, some of which could be used as anti-aliasing and/or reconstruction filtering in digital audio systems, reveals differences between these and possibly the perceived quality as determined by a critical audience.

In previous papers (ref. 1 and 2) was shown that the perceived quality difference between moving magnet and moving coil cartridges is largely due to a difference in impulse response. These two types thus show a strong difference in temporal decay in agreement with the perceived quality.

1. INTRODUCTION.

The discussion on the perceived quality of audio systems often lacks objective criteria. This is partly due to the subjective experience of the ill-defined property "quality", covering many aspects, partly to the lack of understanding of all the properties that influence the perceived quality. The latter is not synonymous with the technical quality of a system to begin with.

Disregarding non-linear distortions, the frequency response between 20 Hz and 20 kHz of a system is very often taken as a major parameter determining the quality of a sound reproduction system. The basic idea behind this is the Fourier analysis of sounds, in which any sound wave, no matter how complicated, can be decomposed into an infinite series of sine and cosine waves of different frequencies, starting at zero and "ending" at infinity. The, never mentioned, assumption is that the frequency components above the hearing limit, usually taken at 20 kHz, do not influence the perceived sound in any way.

Although this seems a reasonable assumption at first, it is not as straightforward as one would think. Two aspects play an important role: the first is that Fourier analysis only holds for linear systems and if there is one transducer which is non-linear, it is the human ear. In non-linear systems frequencies not present in the original signal can be generated and/or other frequencies can acquire more power than in the original signal. This can easily be demonstrated using a 3 kHz sine wave with 5 periods on and 5 periods off. Although Fourier analysis tells that 300 Hz is only a weak component in this signal, it is the strongest one hears. As 300 Hz corresponds to the envelope of the signal it is not surprising using the non-linear properties of our ears. It can be concluded that frequencies above the hearing limit can indeed generate signals that are below the hearing limit which could thus influence the perceived sound and the quality experienced.

The second aspect is that the limitation of the bandwidth of an audio reproduction system has consequences in the *time domain*, which we will discuss in the next section. The relation between the spectral response and the temporal behaviour will be discussed, followed by some examples and discussion on the perceived quality.

2. TEMPORAL DECAY.

We will try to find a parameter, that is able to characterize in "shorthand" the temporal properties of an audio system, in a way that is both easy to find and to interpret. The Wigner Distribution (ref. 3 and 4) is a very powerful technique, but requires sophisticated computational systems to obtain and is not easy to interpret, especially for people with less feeling for mathematics. The Energy-Time Curve (ref. 5) is simpler, but still interpretation is not easy, because of its non-causality. Before we will define temporal decay, we will make some simplifications:

Any audio system (starting with the microphone picking up the original sound and ending with the loudspeaker reproducing the sound) can be regarded as a band-pass filter as shown in fig. 1. The first approximation we will make is to neglect the small wiggles in the response, so we will idealise the response to that of fig. 2. However, all audio systems are more or less non-linear and thus produce (amongst others) harmonic and intermodulation distortions. Although these are very important for the perceived quality of an audio system we will concentrate on "high-end" systems and neglect these non-linear effects in order to simplify the problem. To simplify the problem even further, we will neglect the high-pass filtering at the low end of the audio spectrum and thus model an audio system as a linear low-pass filter as shown in fig. 3.

The theory of Fourier analysis yields that the inverse Fourier transform of the complex valued transfer function of any filter, and thus also of our idealised audio system, equals the Dirac delta function response of the system in time domain. Note that the impulse response thus tells us more than the amplitude response of a system, because it contains information about the amplitude response at ALL frequencies (not only those between 20 Hz and 20 kHz) and about its phase response, albeit in an implicit way.

Looking at the impulse response of a typical low-pass filter as shown in fig. 4, we see that:

- although the input signal is never below zero, the output signal is negative during a significant part of the time and
- although the input signal has an infinite short duration, the output signal differs from zero during a finite period of time.

In other words, any audio system has the tendency to "smear out" the signal both in amplitude and in time. These effects could reduce subjective experiences like the "definition" and "transparency" of the perceived sound. *This smearing will always be a degradation of the original sound* and we will try to study its influence on the perceived sound.

We think that the "temporal decay" can be used to quantify the temporal smearing by a simple characteristic value. It has units of dB/µsec. and in words it is the inverse of the time (in µsec.) required for the *envelope* of the impulse response to decay by 1 dB. For most analog filters it can be found directly from the time derivative of the envelope function, for digital filters (which can show pre-ringing, see below) this is a bit more complicated. The idea is illustrated in fig. 5 and 6.

If the temporal decay can be found from the time derivative it can written in formula:

$$T.D. = -10^{-6} \cdot {}^{d}/_{dt}\left(20 \cdot {}^{10}\log(env(V(t))/V(0))\right) \text{ (dB/µsec.)} \qquad (1)$$

in which:
${}^{d}/_{dt}$ = temporal derivative of
env = envelope function of
V(t) = output voltage as function of time
V(0) = output voltage at reference time e.g. at maximum of the impulse response.

3. RELATION BETWEEN SPECTRAL TRANSFER FUNCTION AND TEMPORAL DECAY.

Because the (temporal) impulse response is the inverse Fourier transform of the complex valued transfer function, the properties of the transfer function and the impulse response are closely interrelated. In popular terms this interrelationship states that fast phenomena in time domain are wide in frequency domain and things that are slow in time domain are narrow in frequency domain. Because the Fourier transform is a one-to-one projection from time into frequency domain and vice-versa, the above popular statement can also be reversed: a narrow spectrum indicates slow phenomena, a wide spectrum fast phenomena.

This can be illustrated by looking at idealised audio systems with the same cut-off frequency but with different roll-off rates as shown in fig. 7. In fig. 8 the impulse responses of these systems are shown, with a zero phase shift for all frequencies. The differences are obvious, although below 20 kHz. no difference exists! These differences and the related temporal decays could be emphasized even more if they have a different phase response inside or outside of the audio band. *This holds for all types of band limitation*, being it analog or digital, linear or non-linear phase, although there are differences. We will illustrate this in the next section with a number of examples.

4. EXAMPLES.

Some of the examples listed in this section can be solved analytically, others only numerically. We will restrict ourselves to the results in both cases, as most of the analytical cases will be more or less familiar. But before we will discuss this, another parameter will be defined: the t_{90}. This will be the time, required to decay to -90 dB of the maximum value of the impulse response. As -90 dB is approximately the resolution of the CD, this time tells us how long a signal can influence its (temporal) environment.

4.1. First order low-pass filter.

The impulse response of such a filter is:

$$V(t) = V(0).\exp(-{}^{t}/_{\tau}) \qquad (2)$$

in which τ is the time constant of the system.

If we choose τ to be 7.96 µsec., the -3 dB low pass frequency equals 20 kHz and the impulse response obtained is shown in fig. 9. Taking this value of τ and substituting (2) in (1) we obtain

temporal decay = 1.09 dB/µsec.

The temporal decay of this filter is independent of time and the t_{90} is 82 µsec.

4.2. Brick wall filter.

The so-called "brick wall" filter is often regarded as the most ideal low-pass filter that can be imagined. Its transmission coefficient is 1 below the cut-off frequency and 0 above it and is shown in fig. 10. The phase shift is 0 degrees for all frequencies. Although it cannot be realised in practice, for reasons we will see shortly, often designers try to approach it as close as possible in many digital audio systems. Its impulse response is given by:

$$V(t) = V(0) \frac{\sin(2.\pi.f.t)}{t} \qquad (3)$$

in which:
f = cut-off frequency of brick wall filter. (Hz)
t = time (s)

This impulse response is shown in fig. 11. The impulse response decays only very slowly and it also shows "pre-ringing", which starts at -∞. Because of the causal effect of real filters (they cannot produce any output signal before there is an input signal) a brick wall filter cannot be built. Not surprisingly, close imitations as often used in CD players, show similar behaviour, except that the pre-ringing occurs only during a limited period (equal to the time delay of the filter). The envelope of this function equals for t > 1/f:

$$\mathrm{env}(V(t)) = {}^{V(0)}/_{t} \qquad (4)$$

Substituting this is (1) for a 20 kHz cut-off frequency yields:

$$\text{temporal decay} = 8.69 \bullet 10^{-6}\ 1/t\ \text{(dB/µsec.)} \qquad (5)$$

The initial value (around t=0) is:

temporal decay = 0.16 dB/µsec.

From (5) we see that the temporal decay *decreases* with *increasing* time. This can also be illustrated by the following figures: the t_{36} is 1000 µsec. and the t_{90} is 503 000 µsec! This example clearly illustrates the undesirable temporal behaviour of such steep low-pass filters. Yet the amplitude and phase characteristics are far superior to that of the first order low pass filter.

4.3. Elliptical filter, 80 dB/octave.

An example of the impulse response of a filter that can be built and has a steep roll-off of approximately 80 dB/octave is shown in fig. 12. The absence of pre-ringing, compared to fig. 11, is marked. Numerical calculation of the temporal decay shows similar behaviour as the brick wall filter, it decreases with increasing time. The initial value is:

temporal decay = 0.22 dB/µsec.

This value is higher than the one of the brick wall filter, in agreement with the Fourier inversion theorem. Because it is a real (existing) filter it only shows output signal after an input signal has arrived. In this case no pre-ringing occurs, but alteration of the phase characteristic can create this. This will be illustrated in the next example.

The t_{90} is approximately 500 µsec., which also is much less than the brick wall filter value.

4.4. Elliptical filter, 80 dB/octave phase linear.

In this example we study a theoretical filter, it cannot be built, but it is useful to illustrate some effects. We calculate the behaviour of a filter with exactly the same amplitude response as the one in the previous example, but we give it zero degrees phase shift for all frequencies. The thus obtained filter has an impulse response shown in fig. 13.

If we compare this with the response shown in fig. 12, we see that the ringing is now divided symmetrically around the delta pulse but shows a lower value of the temporal decay. The value found for the initial temporal decay is 0.19 dB/μsec.

Also t_{90} (750 μsec.) is larger than this number for the original filter (500 μsec.) but note that it is redistributed in time.

4.5. Filtering in a Digital Audio Chain.

A complete Digital Audio Chain consists in its simplest form of an anti-alias filter, an A/D converter, a digital recorder, a digital play-back system, a D/A converter and a reconstruction filter. As far as filtering is concerned, only the anti-aliasing and reconstruction filtering are of interest. In all discussion about the filtering in CD players it is mostly forgotten that digital recordings have to be made with anti-aliasing filters which are often not of a digital nature (ref. 6). We can thus get an impression of the overall behaviour by putting the filter discussed in 4.3 and 4.4 in series. The thus obtained response is shown in fig. 14, from which we can see that most of the undesirable behaviour has already happened and that the reconstruction filter hasn't made things much worse at best. Because the overall steepness of the combined filter is increased, the temporal decay of the overall system is further decreased and the t_{90} increased.

4.6. Minimised time-smear (MTS-)filter.

A completely different type of filters are the ones optimised for temporal impulse response. This is achieved by a cleverly chosen roll-off rate of the filter, together with a phase shift proportional to the frequency over the most important part of the frequency range. An example of the thus obtained impulse response is shown in fig. 15, which is a recent development of the author.

Due to the shape of the impulse response the temporal decay of the filter *increases* with time, which severely reduces the time required for the envelope of the signal to reach the -90 dB level, which is only 100 μsec. The initial value of the temporal decay (over the first 10 μsec.) is 0.78 dB/μsec.

These values are very close to the first order low-pass filter values. Yet its spectral characteristic is far superior to that of the first order low-pass filter (it develops into a 10^{th} order filter). Hence this filter can be regarded as an excellent compromise between temporal and spectral behaviour.

5. DISCUSSION.

The examples presented in the previous section show that the filtering in audio systems can severely degrade the temporal behaviour of the systems although this is not reflected at all in the amplitude response curve between 20 Hz and 20 kHz. To the audibility of these effects a few remarks will be devoted in the next section, but a few important observations can already be made. These examples also show that the temporal decay has a clear correlation with the characteristics of (low-pass) filters and can thus in a semi-quantitative way be used to compare the temporal behaviour of audio systems and thus an aspect of their perceived quality.

If only phase linear filtering is applied in a digital audio system for reconstruction filtering, the overall result is not much better than that of the anti-aliasing part itself.

If linear phase filters would be used both during recording and playback in digital audio systems, as is the trend (ref. 6, 7, 8 and 9) the pre-ringing would occur in the overall system. One could have doubts if this would be attractive because all natural sounds have the tendency to have a sharp increase in strength at the start and then to decay slowly. The unnatural increase in volume before the actual instrument starts playing (which can be longer than the temporal resolution of our ears, see above) might, for the perceived quality of a system, be worse than the phase distortion introduced by analog filters. To answer this question, we should know in what way our ears are sensitive for phase effects. An alternative interpretation of the results of ref. 4 could point that this is done indirectly by the non-linear properties of our ears. The use of MTS filters avoids this problem completely.

The relatively long time that the filters need to reach a level below the noise level of digital audio systems requires re-definition of the signal-to-noise ratio of a system, because the system produces "noise" of itself rather long after it has been excited by a signal, at least during periods of time which are in the same order as the temporal resolution of the human ear. Such artefacts can, in my view, not be regarded as part of the original signal.

The temporal decay of high-end analog audio systems is higher than the decay of digital systems *in their present version* and consequently the temporal "smearing" of the formers is less.

A side remark can be made about the influence of cross-over filters. Many loudspeaker systems use filters which have a temporal behaviour far from ideal, as illustrated in fig. 16. If this is not paid the required attention, the temporal decay of the total audio system is very low.

The remark that I've heard long ago that filtering is nothing but a choice from different kinds of misery still holds.

6. EXPERIMENTAL EVIDENCE.

The superior sound quality of moving coil cartridges over moving magnet ones is at least partly due to the extended frequency response and higher temporal decay. Moving magnet cartridges with extended frequency responses approach the perceived quality of the moving coil cartridges, especially those which produce a higher output signal (and thus generally speaking have a lower mechanical resonance frequency). Compensation of the mechanisms that create the low temporal decay of moving magnet elements leads to significant improvement of their perceived quality (ref. 1, 2).

One of the better ways to compare analog and digital systems is by listening to a good copy of an analog recording on disc and the CD made of the same master tape. If the digital re-processing would not audibly effect the signal, no difference would be perceivable. Yet, on a high-end audio system, using e.g. electrostatic loudspeakers for the midrange and high frequencies, the transparency and clarity of the analog version (half-speed master copies) invariably showed to be better.

Comparing loudspeaker systems is one of the most difficult and tricky aspects of audio. Yet, generally speaking, the loudspeakers sounding best are those with the highest temporal decay. To mention some examples: electrostatics, ribbon tweeters and last-but-not-least ionophones. Also, loudspeakers that show a high temporal decay in Wigner distributions generally sound best (ref 3).

7. SUGGESTIONS FOR EXPERIMENTS.

As the work reported here is partly based on theory, partly based on experience, further experiments should determine if temporal decay can be used as a semi-quantative parameter for the perceived sound quality. It is not within my possibilities to do much experimental work on a scientific basis. I therefore invite investigators in this field to do the following experiments. These experiments require:

1. Recorded sound with a frequency response at least up to 40 kHz in order to maintain the envelope of the original signal as much as possible and to have a temporal decay of the recording system that is sufficiently high.

2. The loudspeakers used in the experiments should at least have a temporal decay significantly higher than that of common digital systems.

3. Low-pass filters, both of analog and digital nature, with different temporal decays.

4. An experienced listening panel.

Experiments that could be performed are those in which the recorded sound is reproduced for the listening panel with and without filtering. The listening panel should then judge which sounds show the highest levels of definition and transparency. Although such experiments come close to the experiments reported in literature (ref. 8) the experiments suggested here would be an improvement, because in my view the use of real sounds is to be preferred over unnatural synthetic sounds and the temporal decay of the headphones used in the experiments of (8) has not been established.

8. CONCLUSIONS.

The temporal decay seems to be a useful "handle" to get grip on the temporal behaviour of audio systems and to make a semi-quantitative comparison. It is an excerpt of the impulse response of a system, which tells more about a system than its frequency response between 20 Hz and 20 kHz.

High-end audio systems often sound better with analog recordings than with digital ones. This is at first surprising because of the very high quality specifications of digital systems. But the temporal decay is one of the few points at which analog systems beat their digital counterparts and it is thus a clear hint of its importance.

The behaviour of the amplitude and phase characteristic of an audio system above 20 kHz. is of importance to its temporal decay and can thus be of influence on its perceived quality.

The use of linear phase filters in (digital) recording and playback might give a lower perceived quality than non-linear phase filters because of the unnatural pre-ringing of such systems.

The use of MTS-type filtering, both for recording (anti-aliasing) and playback (reconstruction) can give an increased perceived quality because of the high temporal decay and low t_{90} values of such filters.

Further experiments are required to establish the importance of temporal decay, although practical experience clearly indicates that it is of importance for the perceived quality of audio systems.

Acknowledgements.

The author wants to thank Difa Measuring Systems B.V. from Breda, Netherlands for their kind permission of the use of their filter characteristics and Ir. A. Verbruggen of Difa Measuring Systems for the many fruitful discussions on the subject of filter behaviour.

References.

1. H.R.E. van Maanen, "Compensatie van Mechanische Resonantie bij Pick-up Elementen", Radio Elektronica 79, 15/16, pp. 25-29 and 17, pp. 35-41 (1979).

2. H.R.E. van Maanen, "Signaalprocessor voor Magneto-dynamische Pick-up elementen", Radio Elektronica 80, 19, pp. 39-53 (1980).

3. C.P. Janse and A.J.M. Kaizer, "Time-Frequency Distributions of Loudspeakers: The application of the Wigner Distribution", J. Audio Engng. Soc. 31-10 (1983 Apr.), pp. 198-223.

4. D. Preis, F. Hlawatch, P.J. Bloom and J.A. Deer, "Wigner Distribution Analysis of Filters with Perceptable Phase Distortion", J. Audio Engng. Soc. 35-12 (1987 Dec.), pp. 1004-1012.

5. J. Vanderkooy and S.P. Lipshitz, "Uses and Abuses of the Energy-Time Curve", J. Audio Engng. Soc. 38-11 (1990 Nov.), pp. 819-836.

6. M.W. Hauser, "Principles of Oversampling A/D Conversion", J. Audio Engng. Soc. 39-1/2 (1991 Jan./Feb.), pp. 3-26.

7. D.R. Welland, B.P. del Signore, E.J. Swanson, T. Tanaka, K. Hamashita, S. Hara and T. Takasuka, "A Stereo 16-Bit Delta-Sigma A/D Converter for Digital Audio", J. Audio Engng. Soc. 37-6 (1989 June), pp. 476-485.

8. D. Preis and P.J. Bloom, "Perception of Phase Distortion in Anti-Alias Filters," J. Audio Engng. Soc. 32-11 (1984 Nov.), pp. 842-848.

9. T.F. Darling, "Oversampled Analog-to-Digital Conversion for Digital Audio Systems. J. Audio Engng. Soc. 38-12 (1990 Dec.), pp. 924-943.

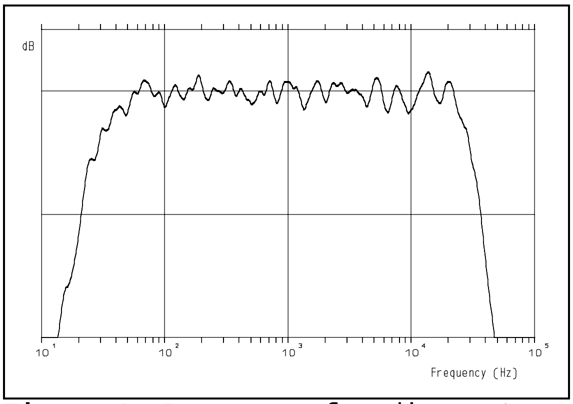

Figure 1: Response of audio system.

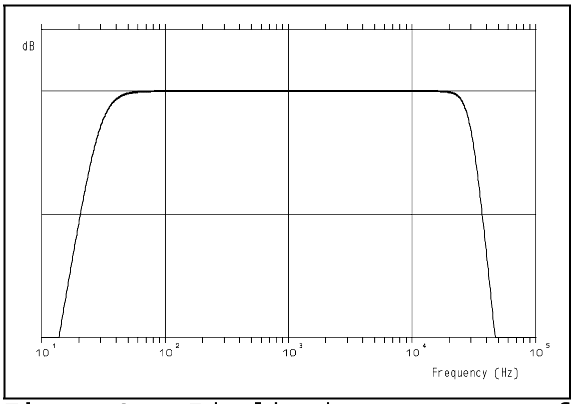

Figure 2: Idealised response of audio system.

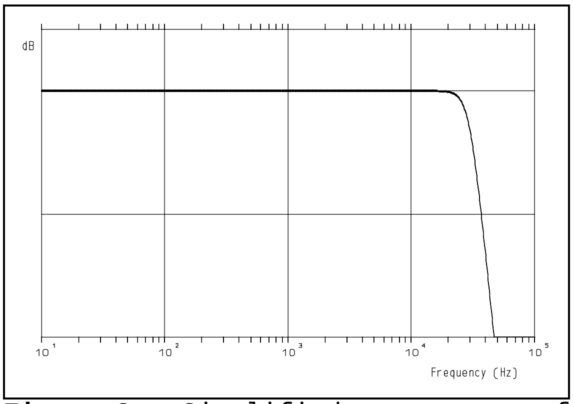

Figure 3: Simplified response of audio system.

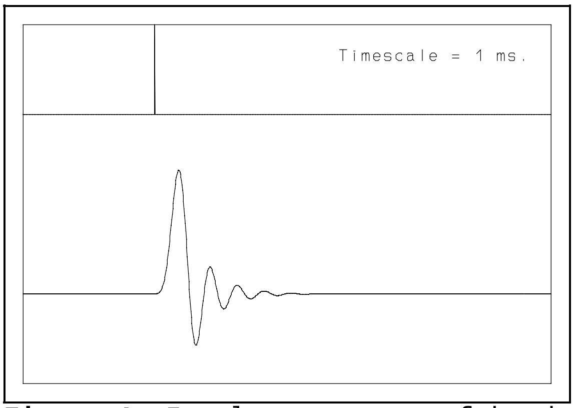

Figure 4: Impulse response of band-limited audio system.

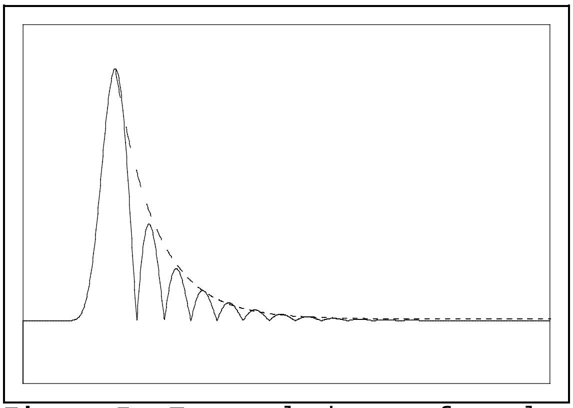

Figure 5: Temporal decay of analog filter.

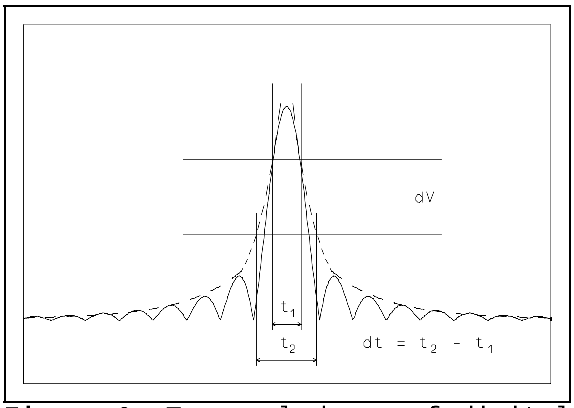

Figure 6: Temporal decay of digital filter.

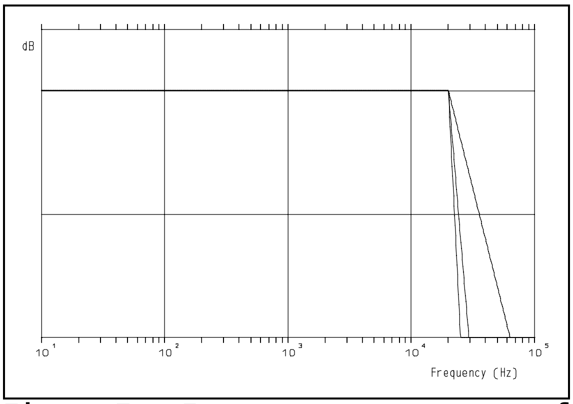

Figure 7: Frequency responses of phase linear audio systems with roll-off rates of order 2, 6 and 10.

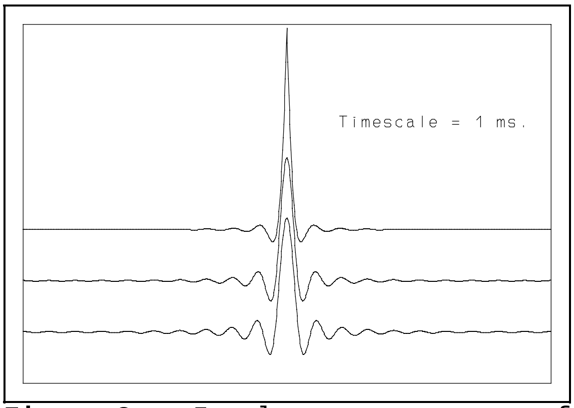

Figure 8: Impulse responses of audio systems with 20 kHz bandwidth but with roll-off orders of 2, 6 and 10.

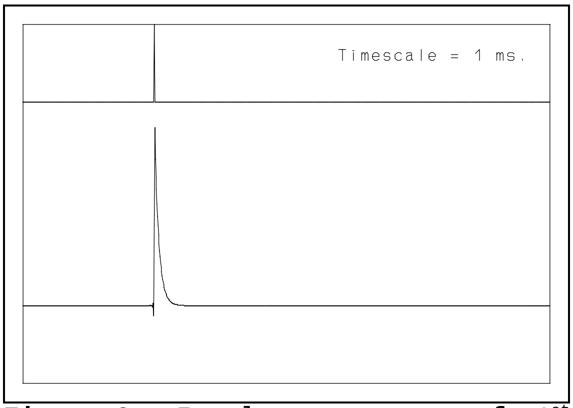

Figure 9: Impulse response of 1st order low-pass filter of 20 kHz.

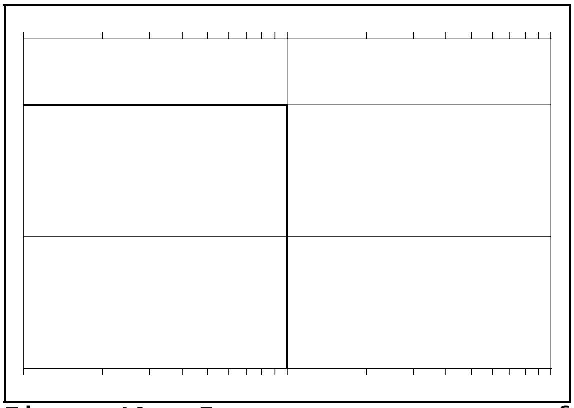

Figure 10: Frequency response of Brick-Wall filter.

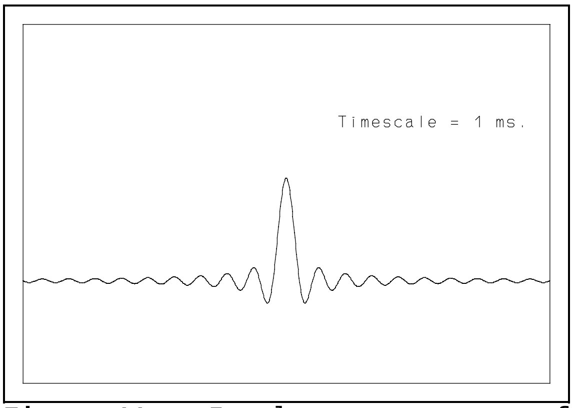

Figure 11: Impulse response of Brick-Wall filter.

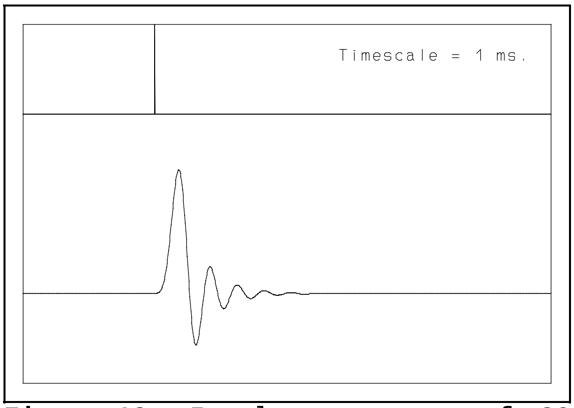

Figure 12: Impulse response of 80 dB/oct. elliptical filter at 20 kHz.

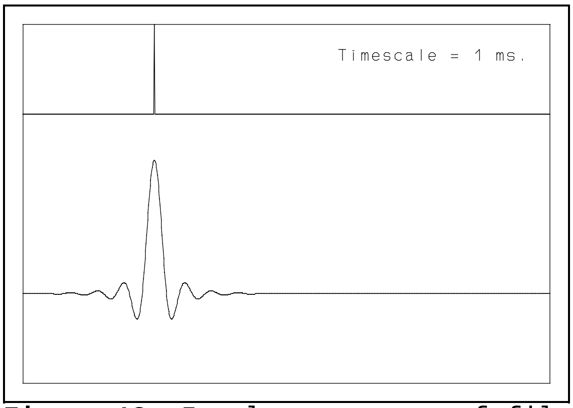

Figure 13: Impulse response of filter of fig. 12, but with the phase shift set to zero for all frequencies.

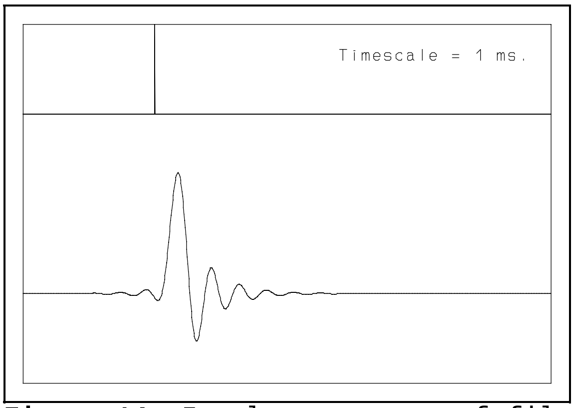

Figure 14: Impulse response of filters of fig. 12 and 13 in series, simulating the effect of a digital audio chain (recording + playback).

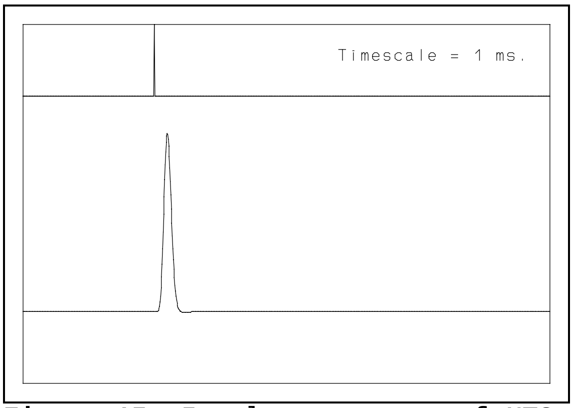

Figure 15: Impulse response of MTS-filter, optimising responses in time and frequency domains.

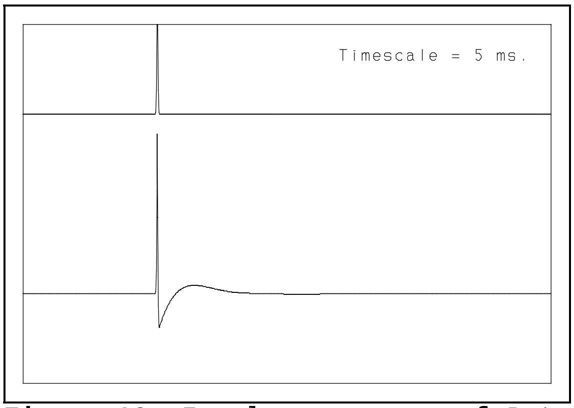

Figure 16: Impulse response of Butterworth cross-over filter of 4^{th} order at 1 kHz.