OPEN ACCESS Freely available online

Modeling and Delay-Equalizing Loudspeaker Responses

AKI MÄKIVIRTA,[*1] *AES Fellow*, **JUHO LISKI,**[2] *AES Student Member*, **AND VESA VÄLIMÄKI,**[2] *AES Fellow*

(aki.makivirta@genelec.com)   (juho.liski@aalto.fi)   (vesa.valimaki@aalto.fi)

[1]*Genelec Oy, Olvitie 5, FI-74100 Iisalmi, Finland*
[2]*Aalto University, Acoustics Lab, Dept. of Signal Processing and Acoustics, FI-02150 Espoo, Finland*

The impulse response of a generalized multi-way loudspeaker is modeled and is delay-equalized using digital filters. The dominant features of a loudspeaker are its low and high-frequency roll-off characteristics and its behavior at the crossover points. The proposed loudspeaker model characterizes also the main effects of the mass-compliance resonant system. The impulse response, its logarithm and spectrogram, and the magnitude and group-delay responses are visualized and compared with those measured from a high-quality two-way loudspeaker. The model explains the typical local group-delay variations and magnitude-response deviations from a flat response in the passband. The group-delay equalization of a three-way loudspeaker is demonstrated in three different methods. Time-alignment of the tweeter and midrange elements using a bulk delay is shown to cause ripple in the magnitude response. The frequency-sampling method for the design of an FIR group-delay equalizer is detailed and is used to flatten the group delay of the speaker model in the whole audio range and in a limited range. The full-band equalization is shown to lead to preringing in the impulse response. However, group-delay equalization at mid and high frequencies only reduces the length of the loudspeaker impulse response without introducing preringing.

0 INTRODUCTION

This paper focuses on the modeling of the linear properties of loudspeakers. The impulse response can be measured on the acoustical axis of a loudspeaker as well as in other orientations to understand directional characteristics and to estimate the power response. Passive and active loudspeakers are systems containing electrical, mechanical, and acoustical subsystems, which create the time-frequency characteristics audible in the acoustic output and visible in the impulse response.

A loudspeaker can fundamentally be characterized as a bandpass system. The roll-off rate of the low- and high-frequency cut-offs define the global characteristics of such a system. Another main feature of a loudspeaker is its multi-way nature. Loudspeakers are typically two- or three-way systems, since a single element cannot radiate efficiently all audio frequencies over a ten-octave frequency range.

Most loudspeakers are dominantly minimum-phase systems. A minimum-phase system releases its energy with a minimal delay, which leads to an impulse response having large signal values in the beginning that decaying with

time. This is a typical behavior in natural causal and stable systems that do not store energy. The phase response of a minimum-phase system is linked to its magnitude response and is the Hilbert transform of the (natural) logarithmic magnitude response [2]. However, it is possible to modify the phase response of a system without affecting its magnitude response.

This paper sheds light on the fundamental characteristics of a multi-way loudspeaker system and adds to the understanding of group-delay equalization in loudspeakers. Usually, only the magnitude response of a loudspeaker is equalized [3, 4], but with time-domain system equalization the phase is also of interest. Another typical application for loudspeaker equalization is the room-response equalization [5–7]. However, in this paper we only analyze anechoic loudspeaker responses and study how completely the observable characteristics of a loudspeaker describe its impulse response. Additionally, this paper discusses delay equalization of loudspeakers. The frequency-sampling method for the design of an FIR group-delay equalizer is detailed. We show how delay equalization affects the impulse response length of a loudspeaker.

Earlier, the transducers of a multi-way speaker have been delayed relative to each other or filtered using allpass delay equalizers [8]. The use of an appropriately chosen delay has been proposed for synchronizing the elements of a

---

*To whom correspondence should be addressed. This paper is an extension of paper 9989 presented at the 144th Convention of the Audio Engineering Society in Milan, Italy in May 2018 [1].
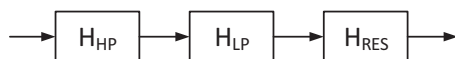
Fig. 1.   One-way loudspeaker model.

Fig. 2.   Magnitude response of the loudspeaker model.

multichannel speaker system [9]. A more refined method of time-domain equalization is to use frequency-dependent time-domain equalizers either for the whole system or for the individual driver channels. Preis has suggested a time-domain equalization, where the time-reversed impulse response of a loudspeaker is used as an FIR equalizer [10]. The phase linearization is achieved at the expense of the magnitude response flatness, since the magnitude, and thus the ripple, is doubled in this method.

Clarkson et al. have proposed flattening the magnitude response and linearizing the phase response of a loudspeaker over a limited frequency range due to hardware limitations using an FIR filter [11]. Such time-domain equalizers are also relatively easy to implement as digital allpass filters [12, 13] without affecting the magnitude response characteristics of the loudspeaker system or adding constraints to the design of the time-frequency equalizer. Adam and Benz [14] and Herzog and Hilsamer [15] have proposed using a time-reversed allpass filter to flatten the group delay of a loudspeaker.

This paper is organized as follows. Sec. 1 introduces the minimum-phase multi-way loudspeaker model. Sec. 2 studies how linear-phase crossover filters affect the impulse response. In Sec. 3, a measured impulse response of a two-way loudspeaker is compared with the model response. Sec. 4 focuses on the FIR group-delay equalizer design. Sec. 5 shows how delaying one driver channel or applying a delay equalizer to the loudspeaker modifies the impulse response. Finally, Sec. 6 concludes this paper.

1 LINEAR LOUDSPEAKER MODEL

In the following, the loudspeaker is modeled as a bandpass device consisting of a cascade of minimum-phase filters. A discrete-time system with a sample rate of 100 kHz is created to accurately simulate the whole operating range of high-quality loudspeakers. Even though the modeling is conducted in discrete time, the same principles apply to continuous-time systems, with the main difference being the compression of the high-frequency part of the spectrum close to the Nyquist limit related to the bilinear transform linking continuous-time to discrete-time representations.

Three system models are developed next: a one-way loudspeaker having no crossover filters but with the typical high- and low-frequency roll-off bandwidth limitation (see Fig. 1), a two-way loudspeaker including the effects of a crossover network and a tweeter-woofer system model, and a three-way loudspeaker modeling a tweeter-midrange-woofer system. Each model has the same low- and high-frequency limitations.

The low-corner-frequency model contains two highpass filters, the first modeling the low corner frequency roll-off towards low frequencies of the mechanic-acoustic system (H_{HP}) and the second modeling the possible bass reflex or
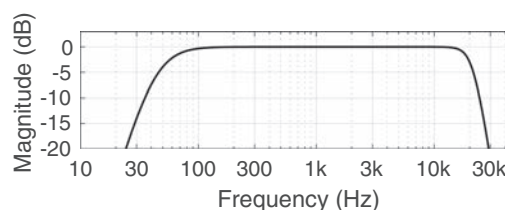
passive resonance effects in the magnitude response (H_{RES}). A bass-reflex port together with the cabinet volume creates a Helmholtz resonator. A passive resonator also creates a mass-compliance resonance system that can be tuned with the mechanical characteristics of the driver and the compliance of the enclosure volume. These are modeled using a second-order highpass filter H_{RES}. Acoustically the Helmholtz resonance opening is, by nature, a bandpass system, but, when considering the overall system magnitude response characteristics, the chosen highpass model describes the contribution of the Helmholtz resonance in the correct way. All filters are modeled as minimum-phase maximally-flat (Butterworth) filters. The one-way speaker model contains three filters in cascade (Fig. 1). In the two-way and three-way systems, the two highpass filters are located in the woofer channel.

The tweeter channel contains a fourth-order lowpass filter H_{LP} modeling the acoustic-electronic roll-off at the high corner frequency. The steep high-frequency roll-off is frequently seen when the natural tweeter roll-off is combined with an electronic bandwidth limitation and describes well the typical characteristics seen at the tweeter roll-off. The exact frequency of this roll-off varies depending on the overall system design. The roll-off typically takes place below 30 kHz due to tweeter driver characteristics.

In the loudspeaker model, the driver outputs are first generated independently and then summed together to obtain the total acoustic response. The summation is equivalent to measuring the system response on the acoustical axis in anechoic conditions. Ideally, this creates a flat response within the loudspeaker passband (Fig. 2). The −6-dB points of the modeled loudspeaker are set to 43 Hz and 22 kHz. These values describe a typical system behavior.

Differences in the loudspeaker responses occur in the impulse and phase responses. We compute the group delay, i.e., the negative derivative of the phase function [2], which describes the delay the system causes for an input signal. The group delay of a loudspeaker contains the contributions of the system's low and high corner frequencies [8], the crossover-filter responses [8], and the mass-compliance resonant system tuning at the woofer frequencies.

Figs. 3 and 4 show the impulse response and group delay, respectively, of a one-way system, defined only by its low and high corner frequencies and the shapes of their roll-off. The impulse response in Fig. 3(a) is normalized so that the maximum signal value is 1.0. The largest oscillation in the impulse response is created by the lowpass characteristic occurring at the high corner frequency (22 kHz). The highpass corner frequency and the mass-compliance
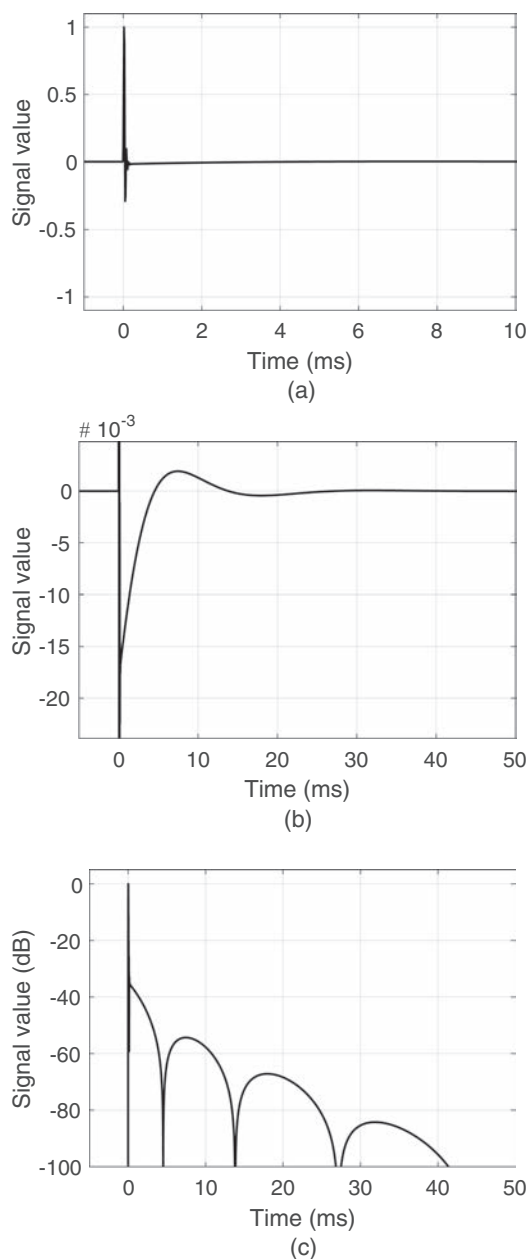
(a)

(b)

(c)

Fig. 3. (a) Normalized impulse response of the one-way speaker model and its low-frequency tail on (b) a linear and (c) a logarithmic scale. Note the different time scales.
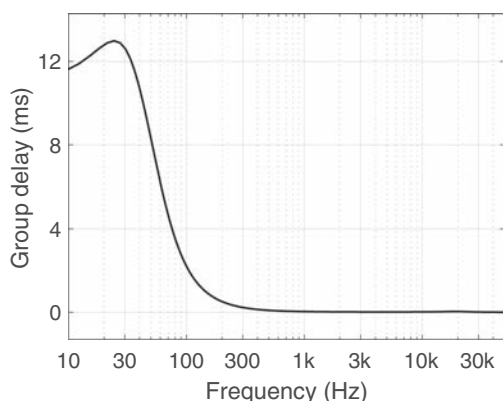

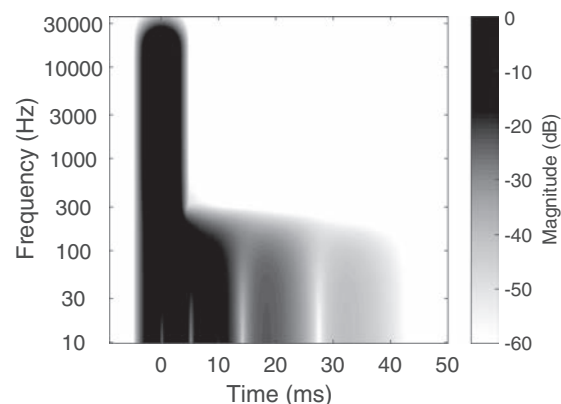Fig. 4. Group delay of the one-way speaker model.

Fig. 5. Spectrogram of the impulse response on a logarithmic frequency axis.

resonator contributions, on the other hand, have a low level in the impulse response but they ring over a long period of time, which is shown in Fig. 3(b) for a different scale and a longer segment of the response. Inspecting the impulse response on a logarithmic scale (Fig. 3(c)) provides a more detailed understanding of the level variation in the impulse response. The group-delay curve (Fig. 4) shows the typical increase in delay towards low frequencies of a bandlimited system with a peak around 30 Hz. The mass-compliance resonant system also contributes to the maximum value of the delay.

1.1 Spectrogram on a Log Frequency Scale

The spectrogram of the impulse response offers further understanding of the time-frequency distribution of the energy in a loudspeaker response. Fig. 5 presents the spectrogram of the model impulse response of Fig. 3. This spectrogram has been computed by splitting the impulse response into 10-ms segments (1000 samples at the 100-kHz sample rate). The hop size in the time domain is 1 sample. The segment is time-windowed using a 1000-sample Blackman window. The resulting segment data is evaluated using the discrete-time Fourier transform (DTFT) at 256 frequencies spaced logarithmically between 10 Hz and 50 kHz. The DTFT magnitude is displayed with the maximum scaled to 1.0 (0 dB), corresponding to the black color. Smaller levels are depicted in shades of gray on a logarithmic scale. This process is similar to the spectrogram computation used in [16].

The spectrogram (Fig. 5) shows that the highest impulse-response energy at frequencies above about 100 Hz appears around zero time. The plot is blurred because of time windowing. Increasing the system delay at low frequencies shows as a smearing of the signal energy over a wide range on the time scale. This is also suggested by Fig. 3(c).

1.2 Modeling Crossover Filters

Next, we add the contribution of the crossover filters to the model. The structure for the two-way speaker model with one crossover network is presented in Fig. 6. The crossover filters are modeled as the most typical Linkwitz-Riley filters [17, 18]. These are minimum-phase filters with
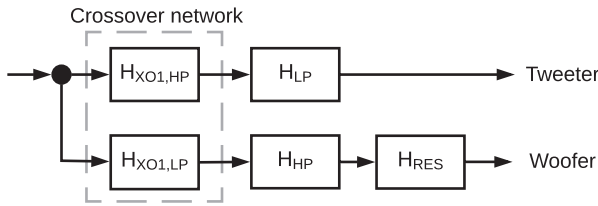
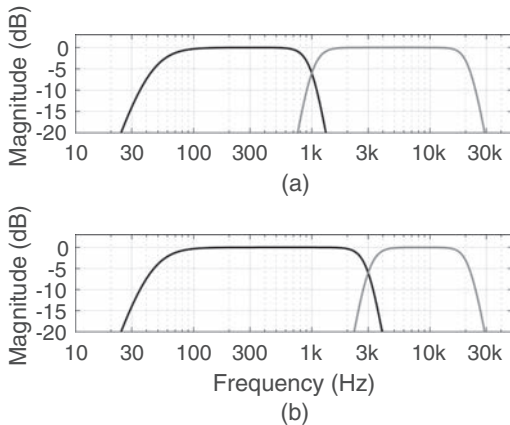Fig. 6.   Two-way loudspeaker model.

Fig. 7.   Magnitude responses of a woofer (black) and tweeter (gray line) using eighth-order crossover filters with the crossover frequency of (a) 1.0 kHz and (b) 3.0 kHz.
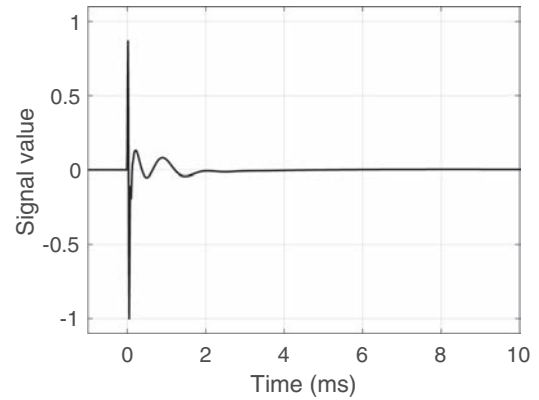

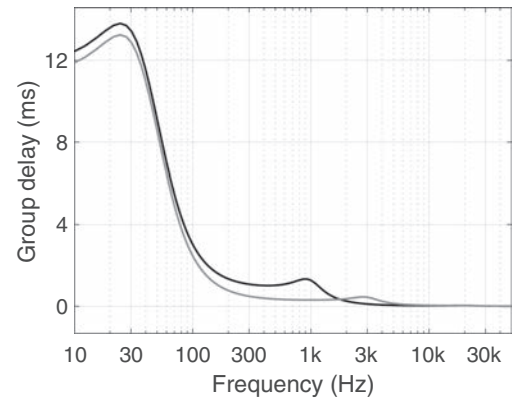Fig. 8.   Impulse response of the two-way speaker model.

Fig. 9.   Group delay of the two-way speaker model, showing an increase in delay near the crossover point at 1 (black) or 3 kHz (gray line). Cf., Fig. 4.

a 6-dB attenuation at the crossover point and an equal delay for the highpass and lowpass output. Each Linkwitz-Riley filter is created as a cascade of two identical fourth-order Butterworth filters so that the order of the crossover filters is eight. Typical values for the crossover frequency of a two-way system are between 1 and 3 kHz. Here, these two extreme crossover frequencies are used. The magnitude responses of the crossover filters together with the highpass and lowpass filters corresponding to the mechanical properties are shown in Fig. 7. The summation of different outputs in a multi-way loudspeaker system are typically set to occur in the correct phase on the design axis, the so-called acoustical axis. The overall magnitude response still remains the same as in Fig. 2 since the summation of the two channels with the crossover filters does not introduce any ripple.

The crossover filters form an allpass system but affect the phase response, and this is visible in the impulse response (see Fig. 8) and group delay: An increase is seen in the group delay in Fig. 9 around each crossover frequency. The lowpass branch delay is larger than the highpass branch. When the crossover frequency is increased, the delay variation at the crossover and the delay difference between the lowpass and highpass branches decreases, as also evident in Fig. 9. The tendency for the delay to increase towards low frequencies below 300 Hz still remains, as this is mainly set by the system characteristics at the low corner frequency.

The effect of the crossover filters of a three-way speaker are studied next. The use of filters in a three-way speaker system is illustrated in Fig. 10, where the low-frequency output of the first crossover network is seen to split into two subbands using a second crossover network, as is typical in

passive loudspeakers [18]. The two crossover frequencies are chosen as 500 Hz and 3.0 kHz. The high-frequency output of the first crossover network goes to the tweeter whereas the two outputs of the second crossover network feed the midrange and woofer elements. Additionally, we still include the mechanical lowpass filter in the signal path of the tweeter and the mechanical highpass and Helmholtz resonance filters in the woofer signal path. Additionally, we include a time-synchronization filter H_{SYNC} in the tweeter signal path in the structure in Fig. 10.

However, if the three-way loudspeaker structure shown in Fig. 10 is used without the synchronization filter (i.e., $H_{SYNC} = 1$), the magnitude response shown in Fig. 11(a) is obtained. The notch around the upper crossover frequency (3 kHz) is due to the phase differences between the tweeter and the midrange channel outputs during summation. This is caused by the second crossover filter since only the signal for the midrange channel passes through it. In order to obtain flat magnitude response over the entire passband, the phase properties of $H_{XO2, HP}$ must be replicated using the synchronization filter H_{SYNC} in the tweeter channel. This can be achieved with an allpass filter, but the simplest solution is to use the highpass filter of the second crossover filter network also in the tweeter channel, i.e., $H_{SYNC} = H_{XO2, HP}$, since the frequency range around and above 3.0 kHz
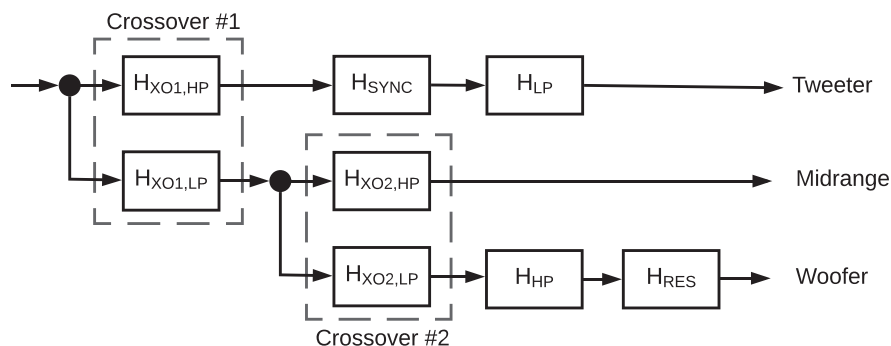
Fig. 10.   Three-way speaker model with two crossover networks and a synchronizing filter H_{SYNC}.
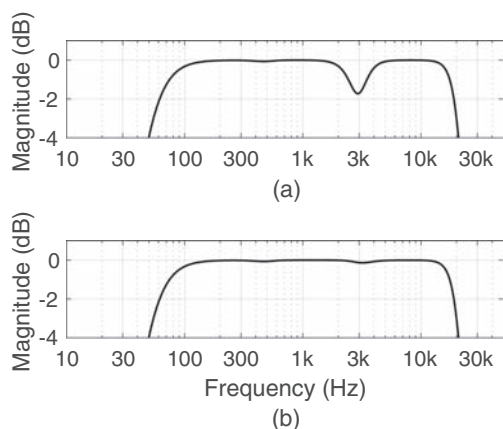

Fig. 11.   Magnitude responses of the three-way speaker model of Fig. 10 (a) without the synchronization filter and (b) using an extra midrange crossover filter as the synchronization filter in the tweeter channel.
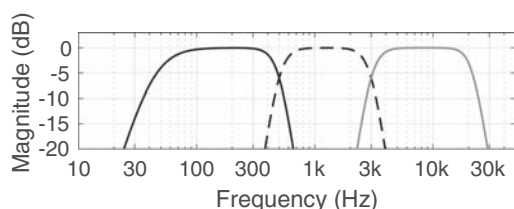

Fig. 12.   Magnitude responses of a woofer (black), midrange (dashed), and tweeter (gray line) signal paths of the three-way speaker model.

Fig. 13.   Impulse response of the three-way speaker model.

Fig. 14.   Group delay of the three-way speaker model.

is in its passband and, thus, the filter does not affect the magnitude response of the tweeter channel. The effect of using the midrange crossover filter as H_{SYNC} on the complete magnitude response of the speaker model is shown in Fig. 11(b), where the notch has practically vanished.

Fig. 12 shows the magnitude responses of the eighth-order Linkwitz-Riley filters when the crossover frequencies are 500 Hz and 3.0 kHz, as well as the highpass and lowpass filters corresponding to the mechanical properties of the loudspeaker. Figs. 13 and 14 present the impulse response and the group delay, respectively, of the three-way loudspeaker model. The impulse response is now seen to be longer than when only two crossover filters were used. The
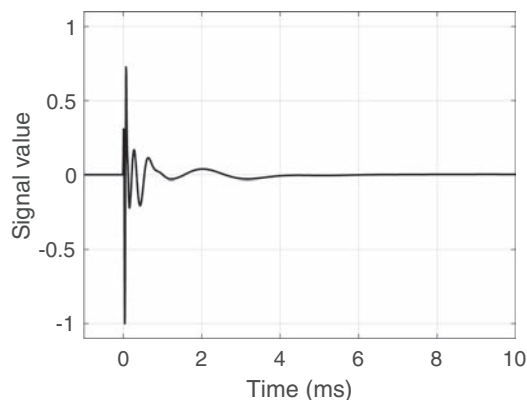
group delay is also larger in the 500 Hz region in Fig. 14 than it was in the case of the two-way speaker (Cf., Fig. 9).

2 LINEAR-PHASE CROSSOVER MODEL

In the digital domain the crossover filter can also be implemented as an FIR filter, which can have a linear phase response [19]. The lowpass and highpass outputs can be implemented as complementary filters. Such a filter can also have an exactly constant delay, and therefore does not contribute to the relative input-to-output timing of audio signal passing through the system.
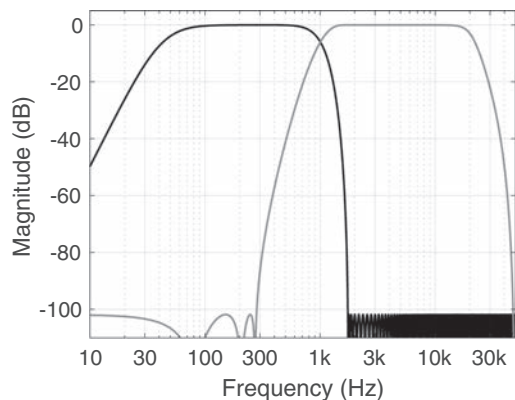
Fig. 15. Magnitude responses of the woofer (black) and tweeter (gray line) signal paths of the two-way speaker model using equiripple FIR crossover filters with a 100-dB stopband rejection.
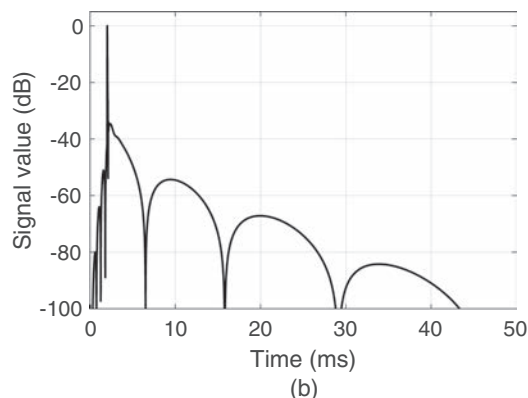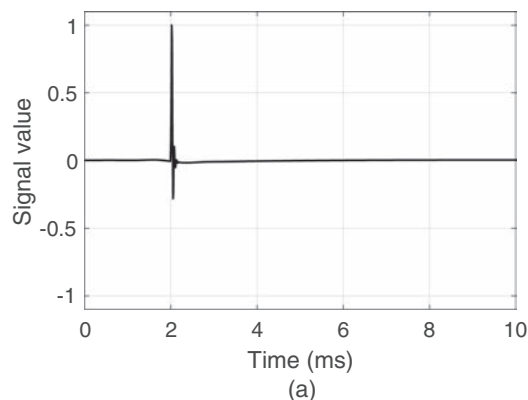

Fig. 16. Impulse response of the two-way speaker model with linear-phase FIR crossover filters showing the preringing on (a) a linear and (b) a logarithmic scale. For the latter, the peak has been normalized to 0 dB.

Fig. 17. Group delay of the two-way speaker model with linear-phase crossover filters. The dashed line indicates the group delay caused by the linear-phase FIR crossover filters.

The precision of the phase match of the crossover-driver system is in reality affected not only by the crossover filter but also by any other component, such as the drivers, having a non-flat magnitude response or filtering effect. For simplicity, in this work we assume the drivers to have a flat magnitude response on the passband of their associated crossover filter.

In the following, a two-way system is modeled with a linear-phase FIR filter. The overall magnitude response is practically the same for this and the previous examples. The crossover frequency is 1.0 kHz, defined with a transition ending at the 100-dB stopband attenuation at 1.7 kHz in the lowpass kernel filter. The FIR filters were designed using the Parks-McClellan optimization algorithm, which leads to an equiripple design [2]. The magnitude responses of the woofer and tweeter channels containing the FIR crossover filters are shown in Fig. 15. The FIR crossover creates a constant delay in the passband mid frequencies not affected by the low or high corner-frequency roll-off. The length of this delay is half of the FIR filter order because the kernel filter has a linear phase.

The linear-phase FIR crossover model shows the preringing typical of FIR filters. By preringing we mean the non-zero level in the impulse response appearing before the largest magnitude value in the impulse response. Since the filter order N is relatively small ($N = 400$ in this case), the preringing is short and, due to the shape of the filter's magnitude response, decays rapidly when moving to times earlier than the main peak in the impulse response. The contribution of the linear-phase crossover filters is visible in the impulse response of the loudspeaker model, which is delayed and also has some preringing on the left side of the main peak (see Fig. 16(a)). A causal linear-phase FIR filter introduces a constant input-to-output delay. The added group delay of 2.0 ms corresponds to 200 samples ($N/2$) at the 100-kHz sample rate. This is also visible as a time-offset of the main peak of the impulse response in Fig. 16(a). The magnitude of preringing becomes more visible when the same impulse response is displayed on a logarithmic scale (Fig. 16(b)). Fig. 17 shows the corresponding group-

delay curve, which has an approximately constant value of 2.0 ms at frequencies above approximately 300 Hz.
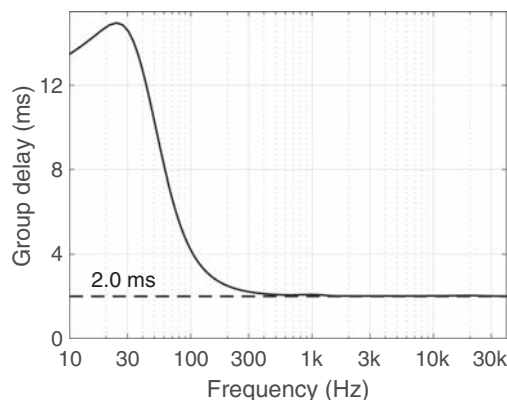
3 COMPARISON WITH LOUDSPEAKER MEASUREMENTS

The validity of the loudspeaker model was tested by comparing it with measurements from real loudspeakers. When two- and three-way loudspeakers were measured, the

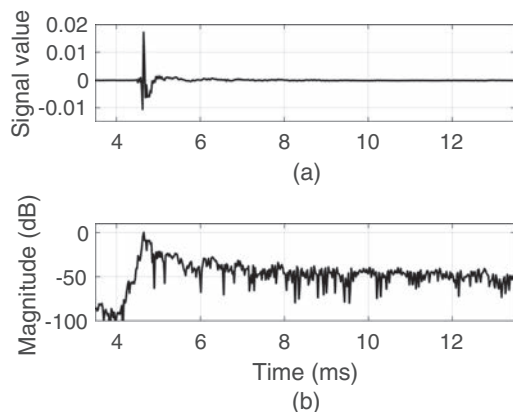Fig. 18. Impulse response of the measured two-way loudspeaker on (a) a linear scale and (b) a logarithmic scale.
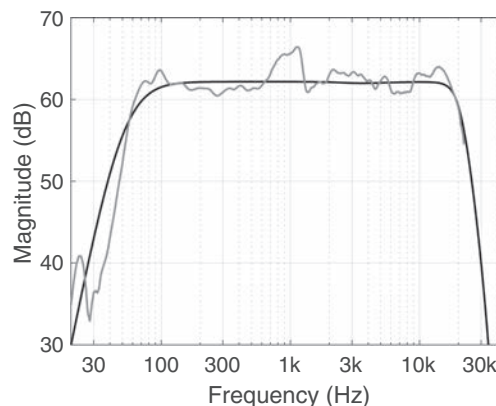

Fig. 19. Magnitude response of the two-way speaker model (black) and the measured two-way loudspeaker (gray line).
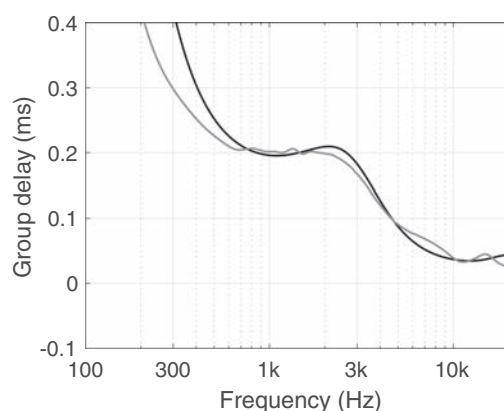

Fig. 20. Group delays of the two-way speaker model (black) and the measured two-way loudspeaker (gray line). (Note the different frequency limits.)

responses were found to be similar to the model responses. However, in three-way speakers, the increase of the measured group delay at the higher crossover frequency is so small that it is hard to observe. This may be caused by measurement noise and by the fact that speaker drivers do not have a flat frequency response, as assumed in this work. For this reason, comparison with a two-way loudspeaker is discussed here.

The measured loudspeaker is a small-sized loudspeaker with a bass-reflex port and crossover frequency of 3.2 kHz, according to the manufacturer's specifications. Its impulse response is shown in Fig. 18(a). Similarly to the digital model, the measured impulse response comprises a tall impulse and a long tail caused by the crossover filters. The measured impulse response is also shown on a logarithmic scale in Fig. 18(b).

The parameters of the model were adjusted to match the magnitude response and group delay of the model to the measured response. The low corner frequency of the model woofer element is adjusted to a value that produces the same -6-dB point of approximately 52 Hz. In addition, the crossover-filter order is set to four and the crossover frequency to 2.9 kHz to conform to the measured data. A slight alteration is needed in the manufacturer-specified crossover frequency.

The magnitude response of the model and the measured loudspeaker show good agreement everywhere except for low frequencies below 50 Hz (Fig. 19). The measured magnitude response is not perfectly flat. This is typically a result of the characteristics of the drivers. There are differences at low frequencies between the measured system and the model. A better match of the magnitude responses could be obtained if the low corner frequency of the model were increased. However, this would lead to an increasing mismatch in the group-delay responses. Alternatively, the order of the low-frequency highpass model could be increased.

The group delay estimated from the measured impulse response is compared with that of the model in Fig. 20. The group delays of the model and the real loudspeaker are in good agreement everywhere except at low frequencies. This is due to the aforementioned compromise between the ideal parameters for the magnitude-response match and

group-delay match. The small increase in the group delay caused by the crossover filters, however, is well modeled. The propagation time in the measured impulse response is accounted for in order to achieve the correct overall level (vertical offset) for the measured group delay.

4 FIR GROUP-DELAY EQUALIZER BASED ON FREQUENCY SAMPLING

In this section we detail the FIR group-delay equalizer designed using frequency sampling. Frequency sampling is often used to design linear-phase FIR filters [20], but here we apply it to obtain an FIR filter with arbitrary phase characteristics. The frequency-sampling method starts with defining the target magnitude and the phase of the equalizer.

The magnitude response of the group-delay equalizer should be flat (i.e., unity) in order not to introduce magnitude distortions. However, since loudspeakers are bandpass devices, the behavior of the equalizer near the Nyquist limit affects the result a little. At the Nyquist frequency, the target frequency response should be real for the impulse response to be real, but in practice the target frequency response is complex-valued generally everywhere except at the zero frequency. Modifying a single point in the
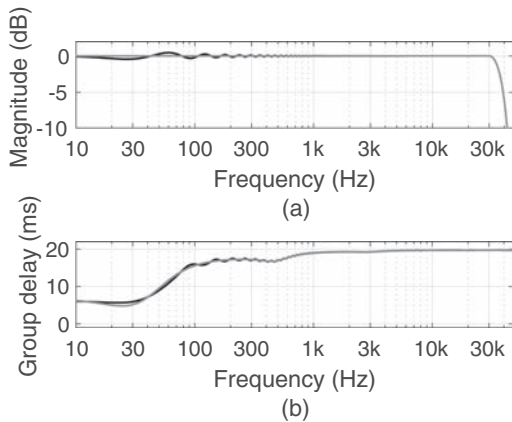
Fig. 21. (a) Magnitude and (b) group-delay response of an example FIR group-delay equalizer (black) and the target responses (gray).

Fig. 22. Impulse response of an FIR group-delay equalizer of order 2047.

target frequency response to be real, such as by zeroing its imaginary part, would lead to a small ripple in the filter response at high frequencies. An alternative way is to allow the target magnitude response to go to zero at the Nyquist frequency, as suggested by Lehtonen [21]. The resulting magnitude response of the equalizer is approximately all-pass in the passband of the loudspeaker, but above 30 kHz the response is attenuated using a raised cosine function, which is easily realizable with an FIR filter [20]. The target magnitude response is shown in Fig. 21(a).

The phase of the equalizer is obtained from the target group-delay response that, in turn, can be arbitrary. One way to define a target group delay is to modify the measured group delay of the loudspeaker to be equalized. For example, when flattening the group delay of the speaker is desired, the target can be set equal to the inverted and delay-adjusted version of the original group delay, which will result in a constant group delay. This is shown in Fig. 21(b), where the original group delay is that of the three-way model. After the group-delay response has been specified, the phase of the equalizer is its antiderivative, which is obtained using numerical integration. The final step in defining the magnitude and the phase of the equalizer frequency response is to mirror them to the negative frequencies everywhere except at the zero and the Nyquist frequency. At negative frequencies, the phase is the negative of the phase at positive frequencies.

The coefficients of the FIR equalizer are obtained with the inverse discrete Fourier transform (DFT). The zero point in time is located in the beginning of the resulting array, whereas the data appearing before the zero time is at the end. This is the consequence of the fact that time is circular in the DFT [2]. Thus, whenever the resulting coefficients of the FIR equalizer expand to the negative times, the two halves of the array must be swapped so that the zero time appears in the middle.

The length of the coefficient array is the same as the number of frequency (DFT) points used. In order to keep the frequency response ripple small, the length of the sampled frequency response must be large. In this work we have used the sampling frequency as the number of frequency
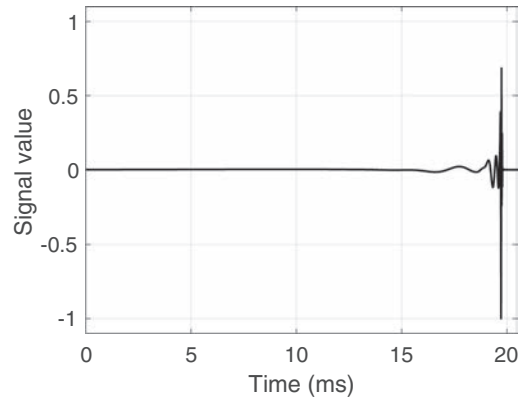
points (100,000). For a practical implementation, a rectangular window is utilized to further truncate the coefficient array, i.e., the impulse response of the equalizer. A truncated equalizer impulse response of 2048 samples is shown in Fig. 22. The rectangular window is suitable here, since the impulse response values are very small in the beginning and end of the response and since it is optimal in the least-squares sense [20].

The best position (start and end points) for the rectangular window function is found by sliding it over the coefficient array, seeking the position that results in the largest squared sum of the truncated array [21]. The length of the window function, or the FIR equalizer array, can be determined to be the smallest acceptable one giving a sufficiently small ripple in the magnitude and group-delay responses, such as less than 0.5 dB and 1.0 ms, respectively. As the filter length is shortened, the ripple starts to increase first at low frequencies. The magnitude ripple of the filter shown in Fig. 22 is less than 0.5 dB (Fig. 21(a)) and the group-delay ripple amounts to 0.8 ms (Fig. 21(b)), and thus the filter length is acceptable in this case.

5 DELAY EQUALIZATION

In this section we study three delay equalization methods that can be applied to a multi-way speaker. The primary motive for creating a loudspeaker response with constant input-to-acoustic output delay is to maintain the electronic signal waveforms accurately in the acoustic output pressure variation.

5.1 Time-Aligning Multi-Way Outputs Using Bulk Delay

The easiest approach to delay equalization is to delay the different outputs of a multi-way loudspeaker crossover such that the delays through all the outputs are equal. This leads to a system response in which the input-to-output latency is constant in most parts of the passband, excluding the frequency regions around the crossover points that show increased latency in the case where the crossover filter has minimum-phase characteristics. When this bulk-delay alignment method is used, achieving the same phase in the
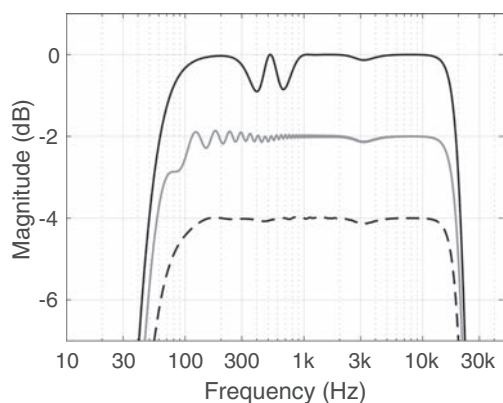
Fig. 23. Magnitude responses of the bulk-delay aligned (black, top trace), full-band delay-equalized (gray, middle), and limited-band delay-equalized model of the three-way speaker (dashed, bottom). The lower curves have been offset by −2 and −4 dB for clarity.
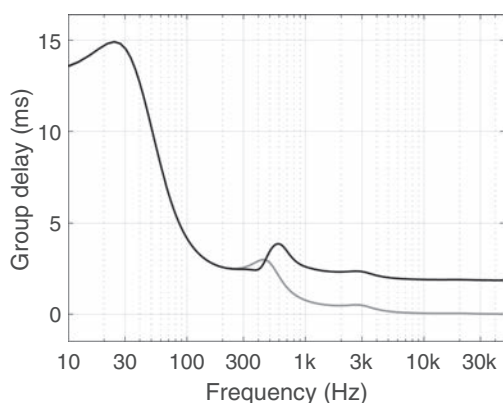

Fig. 24. Group delay of the bulk-delay aligned model (black) of the three-way speaker compared to the original group delay (gray line).

complementary highpass and lowpass output branches may be difficult, resulting in a non-flat system response across the crossover frequency region.

Figs. 23 and 24 show the effect of aligning the system branches using bulk delay in the magnitude response and group delay, respectively, of the three-way loudspeaker model. A bulk delay of 1.85 ms is added to the midrange and tweeter channels to time-align their latency across most of the audible frequency range and to produce the smallest variations in magnitude response. As is shown in Fig. 23, the bulk delay causes a ripple in the magnitude spectrum around the lower crossover frequency. The maximum value of the ripple is 0.9 dB, which is unacceptable, since we want to limit the ripple to 0.5 dB. In addition, as is seen in Fig. 24, the group delay is not precisely constant: At low frequencies, the bass-reflex port still causes a large increase in latency, and there is still an increase in the group delay around the lower crossover frequency (500 Hz).

The impulse response of the bulk-delay modified loudspeaker model on a logarithmic scale, shown in Fig. 25, resembles the impulse response of the original system. The obvious difference in the responses is the increase in the
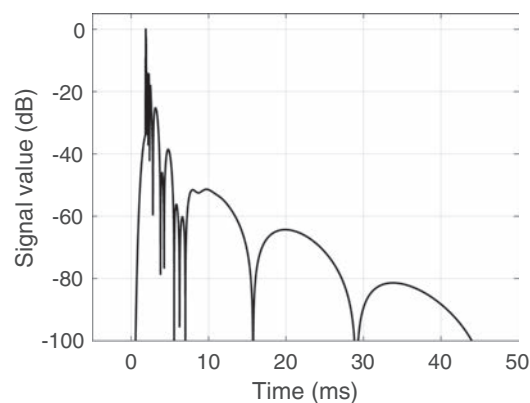

Fig. 25. Impulse response on a logarithmic scale of the bulk-delay aligned model of the three-way speaker.

latency due to the bulk delay (cf., Fig. 13). However, when the impulse responses are compared in dB, the differences in the impulse response length are apparent. Here, we have defined the length using a level of −60 dB in comparison to the maximum value of the impulse response. The length decreases by 0.94 ms, which corresponds to about 7.0% of the original impulse response length.

As the phase responses of the highpass and bandpass branches cannot be matched across the crossover region using this method, a separate filter would be needed to flatten the magnitude-response ripple in the on-axis response. Even if this was done, there would be frequency-dependent variations in the direction of the resulting system acoustical axis across the crossover region due to a phase mismatch of the two branches. In order to alleviate these problems, a more precise method to compensate the delay in the different system output branches, for example, using optimized allpass filters or complementary FIR filters is needed.

5.2 Full-Band Delay Equalization

Other approaches for creating constant-latency system responses with better control at the crossover regions exist, such as the symmetrical FIR filter design within the main audible range and one or more allpass filters [13] used as time-domain equalizers for traditional minimum-phase crossover designs. The FIR crossover filter described in Sec. 2 is an example of the first design approach.

Although design methods for linear-phase FIR filters are well known, the system design usually requires accounting for the frequency-dependent delay effects in the acoustical responses of the transducers, the electronic signal processing, and the amplifier. Optimization methods can be used to incorporate the delay from these different sources in the delay-equalizer design.

Typically, time-domain equalizer designs exclude very low frequencies because delay effects are known to be less audible at low frequencies [22, 10, 23]. Latency effects are created by modal resonances in listening rooms as well, and additionally, delaying the system output much to enable equalization of latency down to low frequencies is impractical. However, when using the three-way loudspeaker model
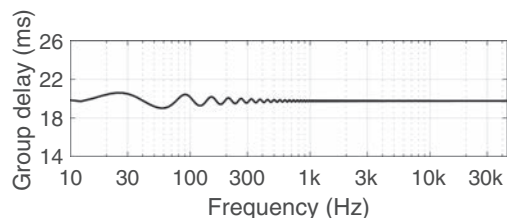
Fig. 26. Group delay of the delay-equalized three-way speaker model.

(a)

(b)

Fig. 27. Impulse response of the delay-equalized model of the three-way speaker on (a) a linear and (b) a logarithmic scale, showing the preringing.

Fig. 28. Group delay of the three-way speaker model when the group-delay equalization has been applied at frequencies above 100 Hz only.

presented in this paper as a starting point, the effects of the group-delay equalization on the entire audio range and a limited frequency range can be demonstrated.

When equalizing the whole frequency band, group-delay equalization is achieved with an FIR equalizer filter having inverted group-delay characteristics with respect to the three-way loudspeaker model shown in Fig. 14. The total group delay becomes constant, corresponding to the maximum value of the original group delay added to the delay of the equalizer. The magnitude and phase responses are defined as discussed in Sec. 4 and the filter coefficients are obtained with the inverse DFT. The length of the filter is 2048 samples, which corresponds to approximately 20 ms. The FIR delay equalizer is placed in the signal path before the crossover filters, and thus it processes the input signal of the loudspeaker.

Fig. 23 shows that the magnitude response remains almost unchanged. The maximum ripple caused by the equalizer is less than 0.5 dB, which is acceptable in hi-fi audio. An approximately constant system group delay is achieved, as desired (Fig. 26). However, this leads to dramatic changes in the system impulse response, as is seen in Fig. 27(a). (The original impulse response is shown in Fig. 13.) The length of the equalized impulse response is calculated similarly to the previous section and the achieved shortening amounts to about 2.4 ms, or 18%.

The equalized system impulse response of Fig. 27(a) resembles an ideal delayed unit impulse. The response has essentially a linear phase. The symmetry is also evident on the logarithmic scale in Fig. 27(b). With the ideal loudspeaker model used here, such group-delay equalization works extremely well. However, with real loudspeakers, the result would be worse, as the resulting system would have an excessive input-to-output delay and considerable preringing in the impulse response. In reality, the constant delay characteristics would have to be limited in frequency in order to keep the increase in the latency at a reasonable level and to reduce the extent of preringing in the time domain. This is demonstrated next.

5.3 Delay Equalizing Excluding Low Frequencies

Since the original three-way model shows large group-delay values below approximately 100 Hz, we use this frequency as the lower limit for the limited-range delay equalization. Thus, the target group delay of the equalizer is set to zero below 100 Hz resulting in a similar delay increase at low frequencies in the equalized response as is shown
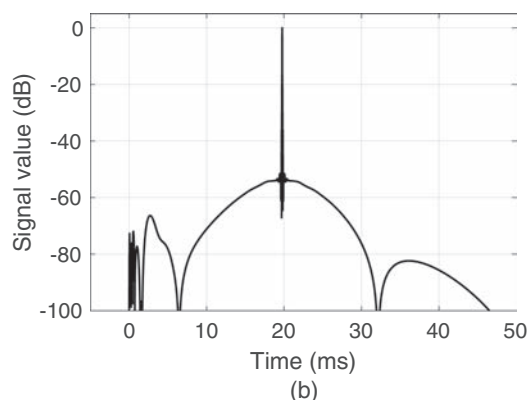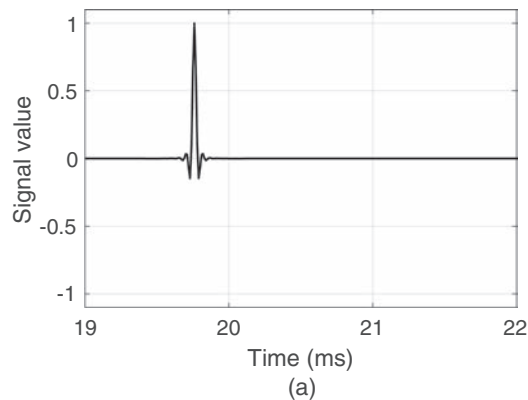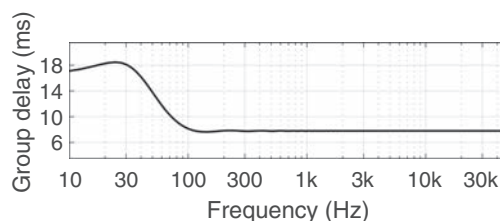
in Fig. 14 in the original response. At higher frequencies, the target group delay of the equalizer is specified so that a flat group delay is obtained. The FIR equalizer is designed using the frequency-sampling method and is placed before the crossover filters. The main difference to the previous example is the length of the window. Since the lowest frequencies are excluded, satisfactory results are obtained using a smaller filter order. Here, the filter length is 1024 samples, which corresponds to approximately 10 ms.

The resulting magnitude response is shown in Fig. 23. The resulting ripple is only approximately 0.1 dB and the smallest of the presented three delay equalization methods. The equalized group delay, shown in Fig. 28, matches the assigned target well. In addition, limiting the equalized frequency band leads to smaller changes in the impulse
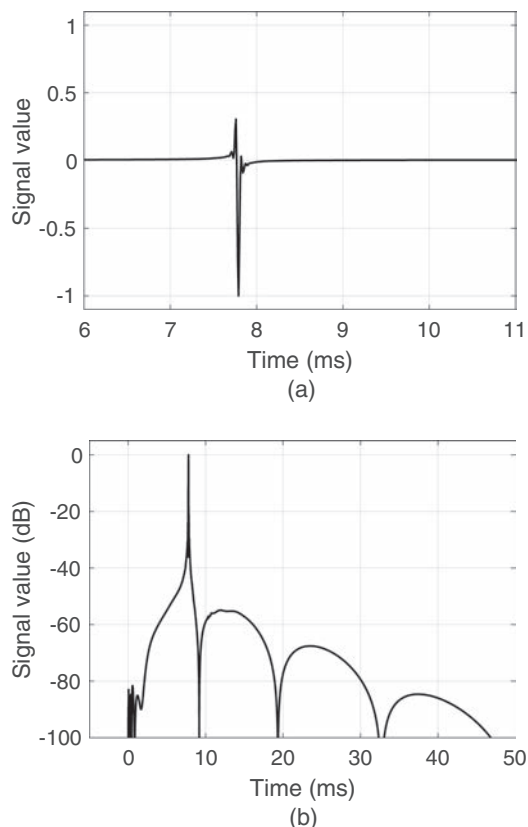
PAPERS

Fig. 29. Impulse response of the delay-equalized model of the three-way speaker, when only the group delay above 100 Hz is equalized, on (a) a linear and (b) a logarithmic scale.

response, as is seen in Fig. 29(a). The impulse response is not as symmetrical as in Fig. 27(a), but instead it retains the properties of the original impulse response (Fig. 13). The impulse response on a logarithmic scale is shown in Fig. 29(b) and the preringing is evidently shorter than in the previous case (Fig. 27(b)).

From Fig. 29(a), the length of the impulse response is estimated to be shortened by 1.2 ms, which corresponds to 8.8%. Since the delay equalization excluding the lowest frequencies avoids pre-ringing, maintains a flat magnitude response, and can still shorten the speaker impulse response, it appears to be the best method for group delay equalization.

5.4 Audibility of Impulse Response Characteristics

A constant-latency design can, in theory, reduce the latency deviation down to any accuracy, but the relevant goal for such an equalizer is to limit the delay variation to at least below the limit of audibility. The audibility limit of latency variation depends on the frequency and input signal character, and frequently the local delay increase of 2 ms in the mid frequencies has been found to describe the just audible difference [22, 10, 24, 25]. These tests are typically conducted over headphones.

Increasing the order of an equalizer system, for example, to reduce the delay variation or to increase the crossover fil-

ter roll-off rates, tends to increase the time-domain extent of the impulse response. The impulse response of a constant-latency system tends to have energy before the largest peak in the impulse response. This preringing in the impulse response has spawned discussions about the possibility of a "pre-echo," which is related to the speculation that parts of the system impulse response occurring early enough before the largest peak in the impulse response might become audible as separate auditory events, constituting unwanted changes in the character of the loudspeaker response where more than one auditory event is recognized.

The human auditory system presents a certain amount of masking before and after an auditory event, a phenomenon called temporal masking. The level of temporal masking before the auditory event (premasking or backward masking) is particularly relevant for the potential to recognize a pre-echo. Unfortunately, while it has been shown to exist, the agreement on the level of premasking as a function of time is not yet unanimous [24, 26]. In published work, the threshold increase in the backward direction depends on the test signals, method of presentation, and other factors. For example, for a narrow-band noise masker the threshold lift for sinusoidal bursts decreases from 5–15 dB at 2 ms before the noise burst to 5–12 dB at 50 ms [27]. For 20-μs duration impulsive stimuli, the backward masking threshold lift is 20 dB very close to the masker and decreases to 5 dB at 4 ms before the masker [28]. To understand the significance of premasking in the case of the loudspeaker impulse response structure, loudspeaker impulse responses will have to be specifically tested, but we can anticipate a potential audible effect.

6 CONCLUSION

This paper has demonstrated how the salient features of the two- and three-way loudspeaker impulse response are described by a model containing a highpass filter modeling the system characteristics at the low corner frequency that includes the effect of any mass-compliance resonant systems used to enhance the low-frequency output, a lowpass filter modeling the cutoff of the tweeter driver and the amplifier bandwidth, and the crossover filters. The effect on the impulse response by an increase in the group delay of conventional Linkwitz-Riley crossover filters as well as linear-phase FIR crossover filters using the modeling method was demonstrated. The measured response of a two-way speaker was compared with the modeled one to demonstrate the similarities.

The frequency-sampling method to design an FIR group-delay equalizer was described. The effects of three alternative delay equalization methods on the system impulse response were then compared. The first method used a bulk delay to align the crossover output branches in time. This method caused a ripple in the magnitude response in the vicinity of the upper crossover frequency. The second method used a general FIR group-delay equalizer, designed using the frequency-sampling method to flatten the entire group-delay response. The preringing caused in the system response in this case was demonstrated.

The third delay equalization method was similar to the second method with the difference that the group delay was only flattened above a certain frequency; in the case of our example, at frequencies above 100 Hz. This method largely avoided the preringing and did not introduce any ripple in the magnitude response. The method also reduced the length of the speaker's impulse response. Thus, the FIR equalization on a limited frequency range can be recommended for group-delay equalization of multi-way loudspeakers. However, delay equalizers can also be implemented using allpass IIR filters. In practical systems, while delay equalization increases the input-to-output delay of the system, this delay can be minimized by limiting the equalization bandwidth from extending down to low frequencies. In addition, delay equalization in real loudspeaker systems must also offer the capacity to deal with the delay variation introduced by the non-ideal response flatness in real transducers.

Future work will consist of listening tests to verify the benefits of the delay equalization in terms of sound quality, particularly considering the audibility of preringing in the system impulse response.

7 ACKNOWLEDGMENT

8 REFERENCES

[1] A. Mäkivirta, J. Liski, and V. Välimäki, "Effect of Delay Equalization on Loudspeaker Responses," presented at the *144th Convention of the Audio Engineering Society* (2018 May), convention paper 9998.

[2] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, 3rd ed. (Pearson Education, Upper Saddle River, NJ, USA, 2010).

[3] M. Karjalainen, E. Piirilä, A. Järvinen, and J. Huopaniemi, "Comparison of Loudspeaker Equalization Methods Based on DSP Techniques," *J. Audio Eng. Soc.*, vol. 47, pp. 15–31 (1999 Jan./Feb.).

[4] V. Välimäki and J. D. Reiss, "All About Audio Equalization: Solutions and Frontiers," *Appl. Sci.*, vol. 6, no. 129/5, pp. 1–46 (2016 May), https://doi.org/10.3390/app6050129.

[5] L. D. Fielder, "Analysis of Traditional and Reverberation-Reducing Methods of Room Equalization," *J. Audio Eng. Soc.*, vol. 51, pp. 3–26 (2003 Jan./Feb.).

[6] A. Mäkivirta, P. Antsalo, M. Karjalainen, and V. Välimäki, "Modal Equalization of Loudspeaker-Room Responses at Low Frequencies," *J. Audio Eng. Soc.*, vol. 51, pp. 324–343 (2003 May).

[7] S. Cecchi, A. Carini, and S. Spors, "Room Response Equalization—A Review," *Appl. Sci.*, vol. 8, no. 16/1, pp. 1–47 (2017 Dec.), https://doi.org/10.3390/app8010016.

[8] N. Thiele, "Phase Considerations in Loudspeaker Systems," presented at the *110th Convention of the Audio Engineering Society* (2001 May), convention paper 5307.

[9] S. Hosoi, H. Hamada, and N. Kameyama, "An Improvement in Sound Quality of LFE by Flattening Group Delay," presented at the *116th Convention of the Audio Engineering Society* (2004 May), convention paper 6115.

[10] D. Preis, "Phase Distortion and Phase Equalization in Audio Signal Processing—A Tutorial Review," *J. Audio Eng. Soc.*, vol. 30, pp. 774–794 (1982 Nov.).

[11] P. M. Clarkson, J. Mourjopoulos, and J. K. Hammond, "Spectral, Phase, and Transient Equalization for Audio Systems," *J. Audio Eng. Soc.*, vol. 33, pp. 127–132 (1985 Mar.).

[12] M. F. Quélhas, A. Petraglia, and M. R. Petraglia, "Efficient Group Delay Equalization of Discrete-Time IIR Filters," *Proceedings of the 12th European Signal Processing Conference*, pp. 125–128 (Vienna, Austria) (2004 Sep.).

[13] S. Bharitkar, C. Kyriakakis, and T. Holman, "Time-Alignment of Multi-Way Speakers with Group Delay Equalization—I," presented at the *124th Convention of the Audio Engineering Society* (2008 May), convention paper 7396.

[14] V. Adam and S. Benz, "Correction of Crossover Phase Distortion Using Reversed Time All-Pass IIR Filter," presented at the *122nd Convention of the Audio Engineering Society* (2007 May), convention paper 7111.

[15] S. Herzog and M. Hilsamer, "Low Frequency Group Delay Equalization of Vented Boxes Using Digital Correction Filters," *Proceedings of the International Conference on Digital Audio Effects (DAFx-14)*, pp. 57–64 (Erlangen, Germany) (2014 Sep.).

[16] V. Välimäki, J. S. Abel, and J. O. Smith, "Spectral Delay Filters," *J. Audio Eng. Soc.*, vol. 57, pp. 521–531 (2009 Jul./Aug.).

[17] S. H. Linkwitz, "Active Crossover Networks for Noncoincident Drivers," *J. Audio Eng. Soc.*, vol. 24, pp. 2–8 (1976 Jan.).

[18] A. V. Mäkivirta, "Loudspeaker Design and Performance Evaluation," in D. Havelock, S. Kuwano, M. Vorländer (Eds.), *Handbook of Signal Processing in Acoustics*, pp. 649–667 (Springer, New York, NY, USA, 2008), https://doi.org/10.1007/978-0-387-30441-0_33.

[19] R. Wilson, G. Adams, and J. Scott, "Application of Digital Filters to Loudspeaker Crossover Networks," *J. Audio Eng. Soc.*, vol. 37, pp. 455–464 (1989 June).

[20] T. W. Parks and C. S. Burrus, *Digital Filter Design* (John Wiley and Sons, Inc., New York, NY, USA, 1987).

[21] H.-M. Lehtonen, "Analysis of Piano Tones Using an Inharmonic Inverse Comb Filter," *Proceedings of the International Conference on Digital Audio Effects (DAFx-08)*, pp. 333–340 (Espoo, Finland) (2008 Sep.).

[22] J. Blauert and P. Laws, "Group Delay Distortions in Electroacoustic Systems," *J. Acoust. Soc. Am.*, vol. 63, no. 5, pp. 1478–1483 (1978 May), https://doi.org/10.1121/1.381841.

[23] B. M. Fazenda, M. Stephenson, and A. Goldberg, "Perceptual Thresholds for the Effects of Room Modes as a Function of Modal Decay," *J. Acoust.*

Soc. Am., vol. 137, no. 3, pp. 1088–1098 (2015 Mar.), https://doi.org/10.1121/1.4908217.

[24] T. G. Dolan and A. M. Small, "Frequency Effects in Backward Masking," *J. Acoust. Soc. Am.*, vol. 75, no. 3, pp. 932–936 (1984 Mar.), https://doi.org/10.1121/1.390540.

[25] J. Liski, A. Mäkivirta, and V. Välimäki, "Audibility of Loudspeaker Group-Delay Characteristics," presented at the *144th Convention of the Audio Engineering Society* (2018 May), convention paper 10008.

[26] H. Fastl and E. Zwicker, *Psychoacoustics: Facts and Models*, 3rd ed. (Springer, Berlin, Germany, 2006), https://doi.org/10.1007/978-3-540-68888-4.

[27] M. Ahmadi, *Central Versus Monotic Masking In Non-Simultaneous Masking Conditions*, Ph.D. thesis, The Ohio State University, Columbus, Ohio, USA (2010).

[28] H. Gaskella and G. Bruce Henning, "Forward and Backward Masking with Brief Impulsive Stimuli," *Hear. Res.*, vol. 129, no. 1–2, pp. 92–100 (1999 Mar.), https://doi.org/10.1016/S0378-5955(98)00228-7.

THE AUTHORS

Aki Mäkivirta   Juho Liski   Vesa Välimäki

Aki Mäkivirta received his Master of Science, Licentiate of Science, and Doctor of Science in Technology degrees in electrical engineering from Tampere University of Technology, Tampere, Finland, in 1985, 1989, and 1992, respectively. After positions in the Medical Engineering Laboratory of Research Centre of Finland and Nokia Corporation Research Center, he joined Genelec Oy in 1995. Since 2012, he has been holding the position of Technology Director at Genelec. Dr. Mäkivirta is a Fellow of the AES and a Life Member of the Acoustical Society of Finland.

●

Juho Liski was born in Helsinki, Finland, in 1989. He received his M.Sc. in Technology degree in electrical engineering from the Aalto University School of Electrical Engineering, Espoo, Finland, in 2016. He conducted his Master's thesis project in the research group of Prof. Vesa Välimäki with the topic "Adaptive Hear-Through Headset." Currently he is working for the doctoral degree, focusing on audio signal processing and equalization.

●

Vesa Välimäki is Professor of audio signal processing at Aalto University, Espoo, Finland. He is the Vice Dean for Research in the Aalto University School of Electrical Engineering. He received his M.Sc. in Technology and Doctor of Science in Technology degrees in electrical engineering from the Helsinki University of Technology (TKK) in 1992 and 1995, respectively. His doctoral dissertation dealt with fractional delay filters and physical modeling of musical instruments. In 1996, he was a Postdoctoral Research Fellow at the University of Westminster, London, UK. In 2001-2002, he was Professor of signal processing at the Pori School of Technology and Economics, Tampere University of Technology, Pori, Finland. In 2006–2007, he was the Head of the TKK Laboratory of Acoustics and Audio Signal Processing. During the academic year 2008–2009 he was a visiting scholar at the Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, Stanford, CA, USA. His research interests include digital filter design, loudspeaker and headphone signal processing, audio effects processing, and sound synthesis. Prof. Välimäki is a Fellow of the AES, a Fellow of the IEEE, and a Life Member of the Acoustical Society of Finland. He was the General Chair of the 11th International Conference on Digital Audio Effects DAFx-08 in 2008 and the General Chair of the 14th International Sound and Music Computing Conference SMC-17 in 2017.