Modal Equalization by Temporal Shaping of Room Response

Matti Karjalainen[1], Poju Antsalo[1], and Aki Mäkivirta[2]

[1] *Helsinki Univ. of Tech., Lab. of Acoustics and Audio Signal Processing, P.O. Box 3000, FIN-02015 HUT, Finland*

[2] *Genelec Oy, Olvitie 5, FIN-74100 Iisalmi, Finland*

Correspondence should be addressed to Matti Karjalainen (`matti.karjalainen@hut.fi`)

ABSTRACT

The low-frequency behavior of sound reproduction in listening rooms is often problematic due to long-ringing modes that are difficult and expensive to control by acoustic means. Modal equalization has been proposed recently to correct the low-frequency problems by shortening the decay times of problematic modes through modification of transfer function poles. While the previous methods were based on the estimation of isolated modes and their parameters, the new method proposed here is a technique to change the time-domain response more directly. It is an advanced windowing technique where the temporal shaping of a given impulse response can be done in a frequency-dependent manner. The method is compared with previous modal equalization techniques.

1 INTRODUCTION

In a listening room it is desirable to have a balanced reverberation time behavior at all frequencies. Low-frequency modes, however, are often not well controlled, showing too slow decay times compared to mid- and high-frequency reverberation time, although some increase toward lowest frequencies is allowed [1, 2]. Designing and constructing a carefully damped listening room is acoustically challenging, expensive, and demands much space. Particularly frequencies around and below 100 Hz are problematic in this sense.

Decay time correction by signal processing in the electronic reproduction channel is an attractive alternative to acoustical improvement of a listening space, particularly at low frequencies where the wavelength is large enough. Traditional magnitude response equalization, however, cannot control the reverberation time nor the decay time of modes in detail. Also, DSP equalization by the inverse filter derived from the measured response, while possible at low frequencies, may shorten decay times too aggressively and may also lead to system overload by boosting notched frequency areas.

We have recently proposed modal equalization of low-frequency behavior in listening room conditions [3, 4]. Contrary to conventional equalization, where the aim is to flatten the magnitude response, modal equalization attempts to balance the decay time of low-frequency modes to correspond to the reverberation time at mid-frequencies.

The previously proposed methods of modal equalization are realized as follows. First, the system response at a receiving point or area is measured. Then the modal decay time has to be estimated as a function of frequency. The maximum allowed decay time is determined according to the reverberation time at mid frequencies, allowing a slight increase towards the lowest frequencies. A problematic mode is equalized by designing a pole-zero correction filter whenever the measured decay time exceeds the target for any low-frequency mode (typically below 200 Hz). This means using a zero pair to cancel out a pole pair associated with a physical mode that is too close to the unit circle, and to insert a new pole pair with the desired faster decay characteristics. This is iterated for problematic modes until the decay time is below the allowed limit value at all frequencies, or until a predefined maximum number of modes has been corrected.

Two methods were proposed for the critical task of modal parameter estimation. The first one [3] is based on time-frequency analysis (such as short-time Fourier analysis) to track the envelope trajectory of each prominent mode. The second method [4] is based on parametric estimation through AR or ARMA modeling whereby the overall response of a selected frequency band is modeled as an all-pole or pole-zero filter. Both of these methods work well when the problematic modes to be equalized are clearly isolated in frequency. The AR and ARMA techniques are more powerful than the time-frequency analysis when the modal density increases, but they also show problems with a high modal density. The new technique proposed in this paper shows more promise in such cases.

The new method is based on frequency-dependent windowing. The measured impulse response of the reproduction channel is first analyzed to obtain an estimate of the decay time as a function of frequency. The allowed upper limit is determined as described above. Next the impulse response is windowed by an exponentially decaying window at those frequencies where the decay time has to be shortened, and at the rest of frequencies the response remains in its original decay shape. The frequency-dependent windowing can be realized in different ways, as described in [5]. The equalization filter is obtained by deconvolution of the desired and the measured response. It is inherently of the FIR type and of relatively high order, but in a multirate implementation it is computationally efficient.

The relevance of modal equalization at low frequencies is dependent on the perceptual effects of modal behavior. Dips and peaks in the magnitude response properties of sound reproduction have been discussed for example in [6, 7]. The effects of temporal response properties have not been studied until recently. In an investigation on the incremental improvement of using modal equalization in addition to conventional magnitude equalization it was shown that a detailed modal equalization is only of secondary importance and often of no noticeable difference [8]. However, careful control of modal behavior guarantees that equalization is well done and may contribute to perceived sound quality in critical conditions.

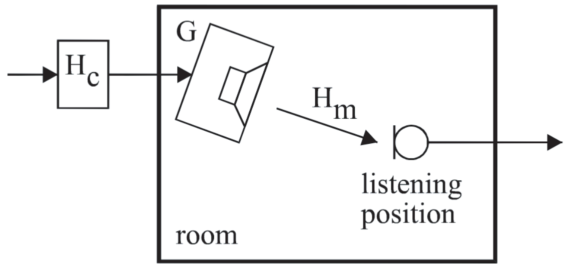

Fig. 1: Basic setup for modal equalization.

This paper is organized as follows. Section 2 is a short description of the modal equalization principle and its functioning. Section 3 introduces the theory of the new method, the frequency-dependent windowing of impulse responses. In Section 4 we show how the new method works for a real room response, and finally in Section 5 we include a summary and conclusions.

2 MODAL EQUALIZATION

Figure 1 depicts the basic setup for equalizing loudspeaker reproduction in a room, in the present case for modal equalization. Another setup is to use a secondary loudspeaker channel to produce the equalizing sound to the listening space [3], but here we will discuss only the case of Fig. 1 with primary loudspeaker equalization.

The total transfer function from the signal input to the listening position in Fig. 1, represented in the z-domain, is

$$H_{\mathrm{t}}(z) = H_{\mathrm{c}}(z)\, G(z)\, H_{\mathrm{m}}(z) \tag{1}$$

where $H_{\mathrm{c}}(z)$ is the equalizing filter, $G(z)$ is the transfer function of the sound radiator from electric input to acoustic output, and $H_{\mathrm{m}}(z)$ is the transfer function of the path from the sound radiator to the listening position. When a desired target response $H_{\mathrm{t}}(z)$ is specified, the response for the equalizer $H_{\mathrm{c}}(z)$ is achieved by

$$H_{\mathrm{c}}(z) = \frac{H_{\mathrm{t}}(z)}{G(z)\, H_{\mathrm{m}}(z)} \tag{2}$$

Now modal equalization is defined as a process that modifies the rate of modal decay in a room. Room modes are created by standing wave resonances. If there is only little damping, which is typical at low frequencies in a hard-walled room, the energy of the modal resonance decays slowly. The idea of modal equalization is to make the decay faster[1] by modifying the overall transfer function of the sound reproduction channel. This means that the (passive) acoustics of the room is not changed, but the reproduced sound behaves in a more controlled way. A simplified explanation on how the decay is made faster is that after the onset of a modal vibration the loudspeaker starts to radiate sound properly in opposite phase, thus actively reducing modal sound energy in the room.

To understand the modal equalization from a systemic point of view, the behavior of a single mode can be considered. Each mode can be understood as a time function of exponential decay

$$h_{\mathrm{m}}(t) = A_{\mathrm{m}} e^{-\tau_{\mathrm{m}} t} \sin(\omega_{\mathrm{m}} t + \phi_{\mathrm{m}}) \tag{3}$$

where A_{m} is the initial envelope amplitude of the decaying sinusoid, τ_{m} is a coefficient that denotes the decay rate, ω_{m} is the angular frequency of the mode, and ϕ_{m} is the initial phase of oscillation. The modal resonance is represented in the z-domain by a transfer function having a pole pair with pole radius r and pole angle θ

$$H_{\mathrm{m}}(z) = \frac{1}{(1 - re^{j\theta} z^{-1})(1 - re^{-j\theta} z^{-1})} \tag{4}$$

To shorten the decay time, i.e., to increase the decay rate and to decrease the Q-value of the resonance, the pole pair $re^{\pm j\theta}$ related to the mode must be moved closer to the origin by reducing radius r.

The impulse response of the acoustic path measured from the loudspeaker to the listening position is needed in order to design a modal equalization filter. In practice modal equalization concentrates to frequencies below 200 Hz and within a restricted listening area.

The process of modal equalization starts with the estimation of mid-range reverberation time, for octave bands between 500 Hz and 2 kHz, by calculating the mean of them. This mean decay time T_{60} at mid-frequencies is the basis for maximum allowed mode decay time at low-frequencies, but allowing the decay time to increase maximally by for example 0.2 seconds as the frequency decreases from 300 Hz down to 50 Hz.

We have previously developed two methods to design modal equalizer filters [3, 4]. The first method (called AMK) attempts to directly identify mode frequencies in the magnitude response and then to obtain the decay rate τ_{m} and mode (angular) frequency ω_{m} by using a time-frequency presentation of the impulse response at frequencies below 200 Hz [9, 3]. The decay rate for each identified mode frequency is calculated using a nonlinear fitting technique modeling the data as a sum of an exponential decay and background noise. The modal equalizer filter is then designed using the mode parameter data. The longest decay time is first corrected by filtering with the equalizer filter, and the process is then iterated until all decay rates are within desired bounds.

The second method (called ARMA) [4] identifies the pole and zero pairs describing a modal resonance by fitting a pole-zero least-squares model directly to the room impulse response. Similarly to the AMK method, the longest decay rate detected by finding the pole closest to the unit circle is compensated by designing an equalizer filter for it, and the method is then iteratively applied until all decay rates are within the desired bound. This second method does not require the explicit stage of fitting to the decay rate of a mode, and is better able to estimate the parameters of closely spaced modes.

The new method proposed in this paper deviates from the two previous ones in the sense that there is no need to deal with individual modes directly, but to shape the entire time-frequency properties of the measured impulse response. Before introducing this frequency-dependent windowing method, a synthetic example of modal equalization is shown to

[1] We may also lengthen the modal decay, or even create new artificial modes, but such boosting must be made with caution to avoid overloading the reproduction system.

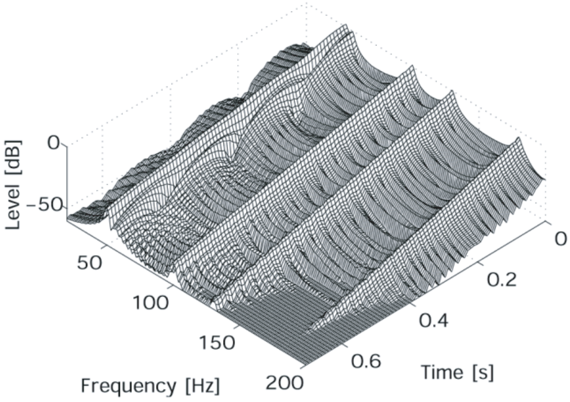

Fig. 2: Waterfall plot of a time-frequency response with five synthetic modes.

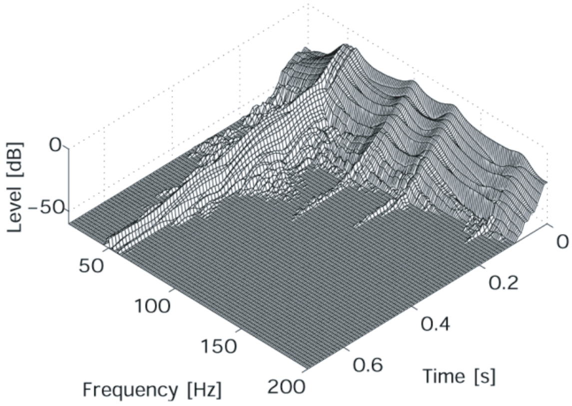

Fig. 3: Waterfall plot of the five synthetic mode case after modal equalization.

demonstrate the effect achieved by modal equalization.

2.1 Synthetic example of modal equalization

A case study with synthetic modes presented here clarifies the idea of modal equalization. Figure 2 shows the waterfall plot of a loudspeaker response measured in an anechoic chamber, with five synthetic modes added at frequencies 50, 55, 100, 130, and 180 Hz. The corresponding decay times are 1.4, 0.8, 1.0, 0.8, and 0.7 s. Now we set up a modal equalizer design target to reduce these decay times to 0.30, 0.30, 0.26, 0.24, and 0.20 s. After processing the synthetic response with an equalizer designed by the AMK algorithm, the decay times have been reduced, see Fig. 3. However, the decay at 50 and 55 Hz continues with the original rate after an initial rapid decay of 15-20 dB. This shows that the method is able to control the decay rate of individual modes, but its performance is limited when modes are close to each other in frequency.

The ARMA method performs better in such cases, but for a dense distribution of modes, often found in real rooms, both the AMK and ARMA methods are problematic. Only the beginning of the decay is then accelerated through modal equalization, but soon the decay rate starts to follow the unequalized rate, as it occured in Fig. 3 for the modes at 50 and 55 Hz.

3 THEORY OF FREQUENCY-DEPENDENT WINDOWING

Signal windowing is one of the most common operations in signal processing. It is motivated by the fact that finite support (span, range) is needed either due to limited processing capacity or most often due to non-stationarity of signals whereby they have to be processed frame by frame to obtain temporally localized representations.

Windowing in its traditional form is a multiplicative operation $(\cdot)$ in the time domain and thus corresponds to convolution $(\star)$ in the frequency domain, i.e.,

$$\begin{aligned} y(t) &= w(t)x(t) \\ Y(\omega) &= W(\omega) \star X(\omega) \end{aligned} \tag{5}$$

where $x(t)$ is a signal to be windowed, $w(t)$ is a windowing function, and $y(t)$ is the windowed signal. Upper case symbols denote Fourier transforms, respectively.

The form of a windowing function $w(t)$ can be any, but most often it is monotonically decreasing towards positive and negative time directions from its maximum value point, and for practical reasons achieves value 0 outside a specific span. Among most common symmetrical window functions are Hamming, Hann(ing), Blackman, Kaiser, and rectangular (or boxcar) window [10]. An asymmetrical window may be for example an exponentially decaying (and truncated) window. The selection of window type and possible parameters associated with it depends on the criteria of each specific application at hand.

Traditional signal windowing works in a frequency-independent manner. This is a straightforward solution, and with proper overlap-add or concatenation processing yields perfect reconstruction from consecutive windowed signal slices. In audio and acoustics applications, however, it is often useful to apply nonuniform time-frequency resolution. In this paper we discuss ways of controlling resolution, particularly methods that are based on time-frequency shaping of system responses.

The two different basic methods of frequency-dependent windowing are: (a) frequency-domain and (b) time-domain techniques. The first one modifies the convolution rule of Eq. (5). The second one applies warping techniques, and it can be done either through frequency-domain resampling or through warped mappings.

3.1 Frequency-domain formulation of generalized windowing

According to the frequency-domain version of windowing [5] in Eq. (5), the Fourier transform of a windowed signal is obtained by (complex) convolution. When this is rewritten in the discrete Fourier transform case for frequency bin m by

$$Y(m) = \sum_{k=0}^{N-1} W_m(k) X(m-k) \tag{6}$$

where indices are modulo N, it becomes obvious that if we wish to obtain frequency-dependent windowing, the windowing term $W_m(k)$ cannot be a fixed vector but should be made bin-dependent. Each bin m can be treated separately by different 'window spreading' vector, designed according to the desired window for that specific bin. These vectors can be composed into an $N \times N$ matrix $\mathbf{W}$ for operation

$$\mathbf{y} = \mathbf{W}\mathbf{x} \tag{7}$$

where $\mathbf{x}$ and $\mathbf{y}$ are the original and the windowed DFT vectors, respectively. Furthermore, the DFT and IDFT transforms can be formulated as matrix operations $\mathbf{q} = \mathbf{F}\mathbf{p}$ and $\mathbf{p} = \mathbf{G}\mathbf{q}$ where $\mathbf{p}$ is a signal vector, $\mathbf{q}$ is a spectrum vector, and $\mathbf{F}$ and $\mathbf{G}$ are the transform and its inverse transform composed of complex exponentials, respectively. Thus the whole chain of Fourier transform of signal $\mathbf{s}$, frequency-dependent windowing, and inverse transform to windowed signal $\mathbf{t}$, can be formulated [5] as

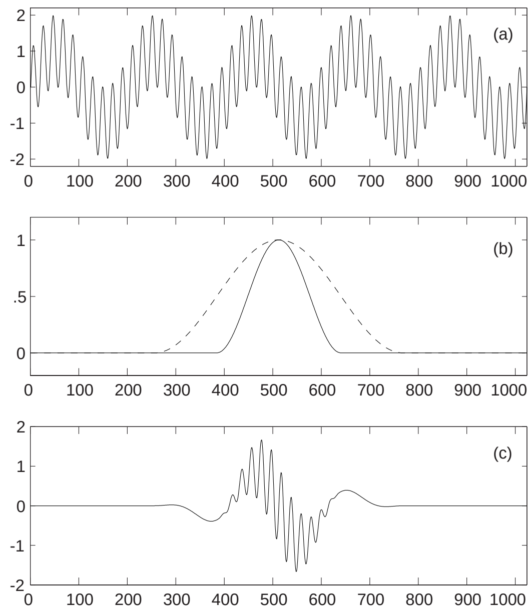

Fig. 4: Frequency-dependent windowing of a sum of two sinusoids: (a) sum of sinusoids, (b) two Hann windows (solid line for high frequencies, dashed line for low frequencies), and (c) result of frequency-dependent windowing.

$$\mathbf{t} = \mathbf{G}\mathbf{W}\mathbf{F}\mathbf{s} = \mathbf{M}\mathbf{s} \tag{8}$$

where $\mathbf{M}$ is an $N \times N$ matrix when the length of signal span $\mathbf{s}$ to be windowed is N, and the frequency-domain window matrix $\mathbf{W}$ is also $N \times N$. For a traditional frequency-independent window, $\mathbf{M}$ reduces to a diagonal matrix, i.e., to a sample-by-sample product of signal and window.

Due to the circularity of DFT and IDFT, frequency-dependent windowing may cause folding problems if not applied properly.

An example of using this method is given in Fig. 4 where the sum of two sinusoids is windowed so that a Hann window applied is shorter for the higher frequency signal component than for the lower frequency one.

3.2 Generalized windowing by time-frequency warping

The idea of *time-frequency warping* [11] can be characterized as follows. A sinusoid is scaled (expanded or compressed) in time by factor $\beta(f)$ depending on frequency f, which operation in the frequency domain corresponds to remapping of frequencies (resampling of frequency to keep uniform bin distribution):

$$A\sin(2\pi f[\beta(f)t]+\varphi) \leftrightarrow A\sin(2\pi[\beta(f)f]t+\varphi) \quad (9)$$

Notice that warping is not a shift-invariant operation, thus the selection of time origin is of special importance. For impulse responses, which we are interested in here, time origin $t=0$ is an inherently determined moment.

Time-frequency warping can be realized in two ways, (a) by frequency-domain resampling or (b) by analytical mappings. Frequency resampling [12] of signal $x(t)$ is realized by a sequence of operations

$$y(t)=\mathcal{F}^{-1}\{\mathcal{R}\{\mathcal{F}(x(t))\}\}$$

where $\mathcal{R}(\cdot)$ is a resampling operator, and $\mathcal{F}(\cdot)$ and $\mathcal{F}^{-1}(\cdot)$ are the Fourier transform and the inverse Fourier transform, respectively. While frequency resampling yields a high degree of freedom for the warping function $\beta(f)$, the method is complicated and computationally expensive, and is not discussed further in this paper.

The second method of warping is based on transforms that map the complex z-domain unit disk onto itself. This approach is known from the design of warped digital filters [11]. Since the bilinear mapping of warped filters has only one degree of freedom, it may not be suitable in modal equalization tasks if high frequency resolution needed. Modeling by Kautz filters [14, 5] is a more attractive technique for this purpose, but this approach remains outside of the scope of this paper.

4 MODAL EQUALIZATION BY THE WINDOWING TECHNIQUE

In this section we will explore the use of the frequency-dependent windowing in modal equalization, applied to a case study of a specific room (dimensions 5.5 x 6.5 x 2.7 m^3) with relatively long modal decay at low frequencies.

The process starts with a measurement of the impulse response from a loudspeaker to the desired listening position. Any modern measurement technique such as the MLS method may be applied; we

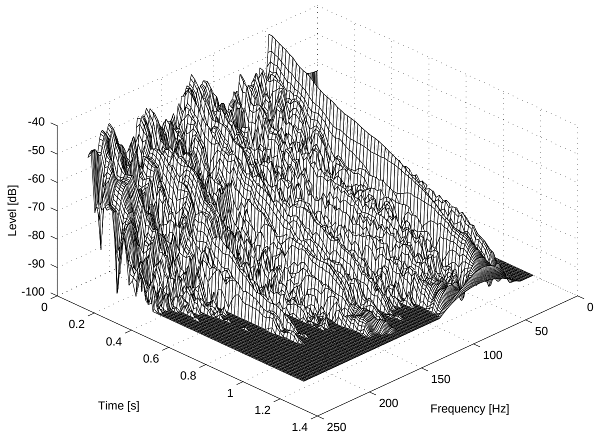

Fig. 5: Waterfall plot of the measured and preprocessed impulse response for the case study room.

have preferred the logarithmic sweep proposed in [15] due to its good tolerance of nonlinearities. The measured response is then preprocessed to remove the bulk delay due to acoustic and measurement system delay, as well as applying high-pass filtering to remove possible infrasound components. Figure 5 shows a waterfall plot computed from the measured and preprocessed impulse response that is used as a case study here.

4.1 Estimation of decay time

The next step of the new modal equalization is to estimate the decay rate in the impulse response as a function of frequency. We have applied first a backward in time (to avoid lengthening of decay) bandpass filtering in frequency bands of about 10–20 Hz width between 20 Hz and 220 Hz and then a nonlinear fitting of an exponential decay plus noise model [3] to estimate the decay time in each frequency band. There is no need to solve mode frequencies and only a moderately accurate estimate of decay rate as a function of frequency is required in this approach. The solid curve in Fig. 6 depicts the decay time vs. frequency curve of the case study response.

4.2 Target response computation

The allowed decay time can be specified by starting from the mid-frequency reverberation time of the room. The mean value of octave bands 500-1000 Hz and 1000-2000 Hz is taken as a reference, and a lin-

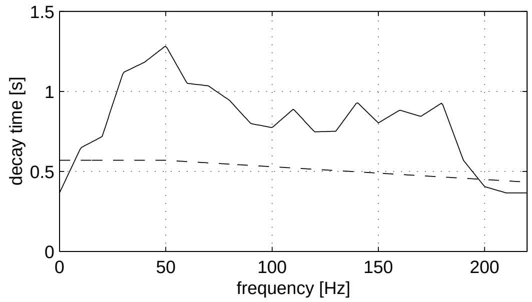

Fig. 6: Decay time as a function of frequency for the measured response (solid curve) and the target maximum decay time (dashed line).

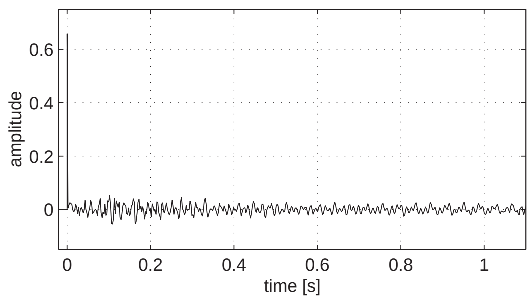

Fig. 7: Impulse response of modal equalizer computed by Eq. 13 for the case study response in the frequency range of 0-220 Hz.

ear increase of modal decay time can be allowed at low frequencies. In Fig. 6 the dotted line shows the maximum allowed decay time for the response of the present case study. It is specified to grow from the mid-frequency value $T_{60} \approx 0.35$ s allowed at 200 Hz by 0.3 s down to 30 Hz. As seen from the figure, there is need to shorten the decay time at frequencies up to 200 Hz.

4.3 Frequency-dependent exponential windowing

The next phase is to determine the frequency-dependent window that is able to modify the measured response to meet the target requirements. In each frequency bin the most natural window type is an exponential one since it retains the exponentially decaying character of each single mode, i.e.,

$$e^{-\tau_{\mathrm{d}} t} = e^{-\tau_{\mathrm{m}} t}\, e^{-\tau_{\mathrm{e}} t} = e^{-(\tau_{\mathrm{m}}+\tau_{\mathrm{w}})t} \tag{10}$$

where τ_{d} is the desired decay rate, τ_{m} is the measured decay rate, and τ_{w} is the decay rate of the exponential equalization window. Now τ_{w} can be computed by

$$\tau_{\mathrm{w}} = \tau_{\mathrm{d}} - \tau_{\mathrm{m}} = \frac{-\log(0.001)}{T_{60\mathrm{d}}} - \frac{-\log(0.001)}{T_{60\mathrm{m}}} \tag{11}$$

where $T_{60\mathrm{d}}$ and $T_{60\mathrm{m}}$ are the desired and measured T_{60} values, respectively, using the relation between T_{60} and decay rate τ.

Now the frequency-dependent exponential windowing can be carried out through the discrete Fourier transform in the following way. The computation can be done for the frequency range of interest, such as 0–220 Hz in our case. The impulse response $x(n)$ to be windowed is first DFT transformed by $X(k) = DFT(x(n))$ to form the vector $\mathbf{x}$ of Eq. 7.

For each frequency bin a new exponential window $e^{-\tau_{\mathrm{m}} t}$ is formed as a vector that is then DFT transformed, and these bin-related transformed vectors are combined to matrix $\mathbf{W}$ of Eq. 7. The windowed impulse response is achieved now simply by multiplication and inverse DFT as

$$\mathbf{y} = \mathbf{W}\mathbf{x}, \quad y(n) = DFT^{-1}(y(k)) \tag{12}$$

This yields the target impulse response of modal equalization.

4.4 Equalizer filter design

In the process above we have obtained the measured impulse and the target impulse response, $x(n)$ and $y(n)$, respectively. The next problem is to realize a digital filter that properly approximates the required modal equalizer transfer function

$$H_{\mathrm{c}}(z) = Y(z)/X(z), \quad h_{\mathrm{c}}(n) = DFT^{-1} H_{\mathrm{c}}(z) \tag{13}$$

This equation actually defines the way to obtain a full accuracy FIR version of the desired equalizer directly by inverse DFT of $H_{\mathrm{c}}(z)$. Figure 7 shows the impulse response obtained in this way for our case study. The response consists of an initial impulse and a slowly decaying ripple. This response cannot simply be shortened by truncation, otherwise the effect of temporal equalization is lost after the length of the response used, as can be seen in Fig. 8, where the impulse response achieved by the equalizer of

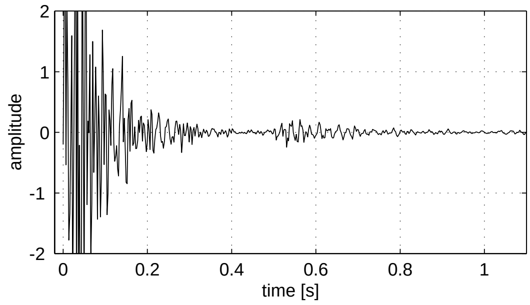

Fig. 8: Equalized response when only part (500 ms) of the equalizer impulse response is used. Notice the echo-like increase in level after 0.5 s.

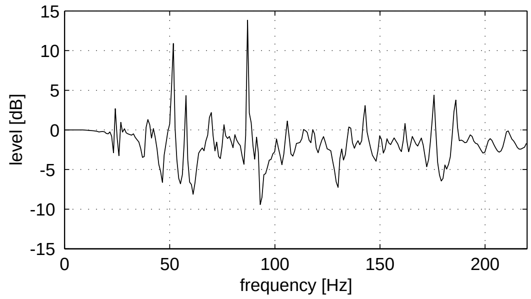

Fig. 9: Magnitude response of the modal equalizer designed for the case of the study.

Fig. 7 is shortened to 500 ms length. Such sudden increase in response is perceived as a disturbing echo effect and must be avoided[2].

An important question from the practical point of view is that the frame size for DFT-based windowing has to be long enough in order to avoid truncation and folding effects. The frequency bin related exponential windows as well as the measured impulse response become truncated, which introduces error, and the circularity of DFT may bring problems unless properly dealt with[3].

Another problem with the equalizer response $H_c(z)$, as obtained from Eq. 13, is that it may introduce boosting of the magnitude response at some frequencies. This can be seen in Fig. 9, which plots the magnitude response of the equalizer shown in Fig. 7. Narrow-band peaks higher than a few dB can be problematic by driving the loudspeaker and amplifier to overloading. A simple way to avoid the boosting is to limit the magnitude level before inverse DFT in Eq. 13. This is problematic, however, because it degrades the Hilbert relation of magnitude vs. phase in minimum-phase systems. A better way to control this detail is thus needed.

[2]Since the temporal shape of the window used in the method is free, a properly adjusted window with a temporal dip could be used to counteract an echo in a given impulse response. Furthermore, this windowing technique to generate a new impulse response from a given impulse response can be used also for audio effects, whereby the peculiarities of a response may be a useful feature.

[3]In fact, the windowed and inverse transformed target response $y(n)$ typically contains a small imaginary component that has to be removed to get a real-valued signal. However, the effect of this is typically negligible.

The question in techical focus is that of finding an efficient yet accurate realization for the derived equalization response. The direct solution of FIR filtering with $h_c(n)$ of Eq. 13 is of prohibitively high order for practical real-time processing at full sample rate, unless some type of fast convolution is used instead, which has a tendency to introduce latency.

Fortunately the need of filtering is here for the low-frequency range only, which means that by a properly decimated multirate implementation the computational load remains light. For example in our case study an FIR of order 512 at a sample rate of 440 Hz can do the frequency range of interest up to 200 Hz with a filter response length of 1.16 seconds. Using of most efficient filter structures means, however, a demanding design and additional latency compared to more direct computation at a higher sampling rate.

A more detailed comparison of different DSP realizations of the modal equalizer is left for future study. It is good to notice that for best results FIR type solutions are necessary, since the main task of the equalizer is to counteract modal pole pairs, and this can be done properly only by using enough zeros (zero pairs).

4.5 Analysis of equalization results

The result of applying the designed modal equalizer to the measured response is illustrated in the waterfall plot of Fig. 10 and the decay rate profile of Fig. 11. Figure 10 shows clearly an accelerated modal decay at most frequencies, when compared to the original response in Fig. 5. Only some of the lowest frequences remain long-ringing. The decay

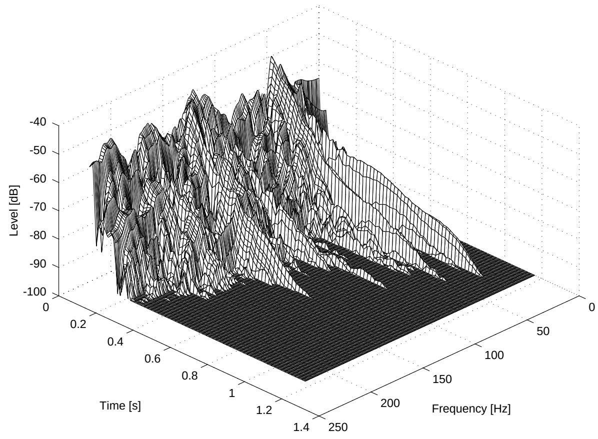

Fig. 10: Waterfall plot of the modal equalized response of the system under study.

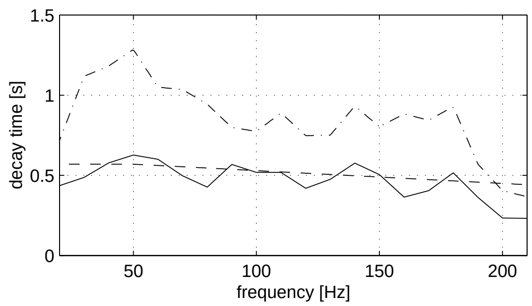

Fig. 11: Decay time as function of frequency for original response (dash-dot line), the target curve (dashed line), and for the equalized response (solid line) using the windowing technique.

rate profile in Fig. 11 gives more information on the performance of the equalization and nonidealities in it.

From Fig. 11 we can find that the proposed windowing technique has shortened the (early) decay times almost as desired. At some frequencies it still exceeds the allowed maximum but only slightly. This nonideality is due to several reasons. First, the decay time estimation, applied here both to the original and to the equalized response, is not ideal. Secondly the FIR equalizer was only 0.7 seconds long instead of 1.3 seconds of the original response. One undesirable feature is that the decay time is shortened also around 200 Hz although it should not be

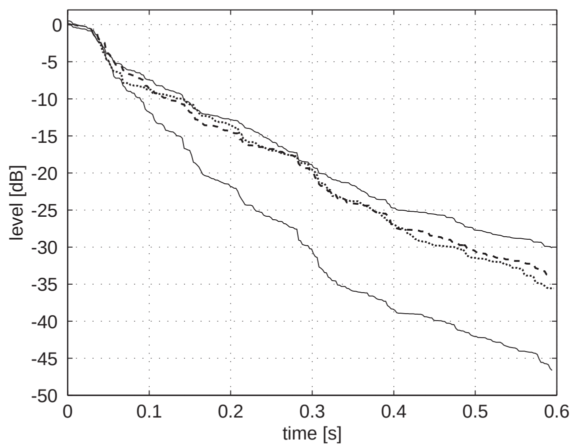

Fig. 12: Decay envelopes (Schroeder-integrated) for the original impulse response (upper solid line), for the AMK equalized (dotted line), for the ARMA equalized (dashed line), and for the windowing mode equalized response (lower solid line).

reduced. However, the overall accuracy achieved is good from a practical point of view.

From Fig. 10 it can be learnt that the decay rate remains slow at 50 Hz after about 0.5 seconds and at about 25 Hz already much earlier. This is no problem in practice, because the loudspeaker used hardly radiates at 25 Hz, and the reduced sensitivity of hearing at low frequencies relaxes the tolerance for long decay time.

4.6 Comparison to prior methods of modal equalization

We applied the AMK and ARMA methods of modal equalization to the case of study to compare the performance of the windowing method to them. When no hand-tuning of design parameters was done, both the AMK and the ARMA method showed only little if any improvement in this demanding modal equalization case. The reason to this is that both of these metods try to correct the response mode by mode. When a prominent mode has been detected and equalized, the magnitude spectrum is also reduced around this peak remarkably, whereby the neighbouring modes gain in relative prominence. If there is a high density of problematic modes, these techniques are limited in their ability to detect and correct all the modes.

Figure 12 shows very clearly how the windowing method is able to control the total decay envelope of the low-frequency part (below 200 Hz) in our case, while the AMK and the ARMA methods without hand-tuning of their design parameters show hardly any improvement. The curves are achieved by backwards Schroeder integration of their impulse responses. By careful adjustment the AMK and ARMA methods may achieve about 5–10 dB of fast initial decay and then the decay rate follows the original one [3, 4]. Thus the new windowing method is found more capable and robust particularly in cases where a densely distributed set of problematically long ringing modes are to be corrected.

5 SUMMARY AND CONCLUSIONS

In this paper we have proposed a new method to realize modal equalization of loudspeaker-room responses, i.e., to control by DSP the temporal decay of system response in sound reproduction for rooms with too low level of absorption at low frequencies. The method is based on shaping the time-frequency properties of a given impulse response by applying a frequency-dependent windowing technique.

The method is able to implement temporal response shaping for complex impulse responses, also for non-exponential decay envelopes, although in this paper we have only applied exponential windowing to shorten excessive modal decay times. The windowing method is found more accurate and robust than the AMK and ARMA methods proposed before.

The critical phases in the realization of the method are: first the measurement of system impulse response with good enough signal-to-noise ratio at low frequencies is needed, then the decay time profile must be estimated reliably, and finally an accurate yet efficient implementation for the equalization filter is to be found. Tradeoff between processing latency and computational efficiency must be done in filter design, since efficient multirate or frame-based structures introduce processing delay, while direct filter implementation may be beyond present processor capabilities. A remaining research task is also to gain deeper understanding of perceptual aspects related to room mode behavior.

6 ACKGNOWLEDGMENTS

This study has been part of the VÄRE technology program, project TAKU (Control of Closed Space Acoustics), funded by the National Technology Agency of Finland. A part of the work was conducted in the SARA Project funded by the Academy of Finland (project number 201050).

7 REFERENCES

[1] AES Technical Committee on Multichannel and Binaural Audio Technology (TC-MBAT), “Multichannel Surround Sound Systems and Operations,” Technical Document, version 1.5 (2001).

[2] EBU Document Tech. 3276-1998 (second ed.), “Listening Condition for the Assessment of Sound Programme Material: Monophonic and Two-Channel Stereophonic,” (1998).

[3] A. Mäkivirta, P. Antsalo, M. Karjalainen, and V. Välimäki, “Low-Frequency Modal Equalization of Loudspeaker-Room Responses,” *Preprint 5480, AES 111th Convention*, New York, USA, 2001 Nov/Dec.

[4] M. Karjalainen, P. A. A. Esquef, P. Antsalo, A. Mäkivirta, and V. Välimäki, ”Frequency-Zooming ARMA Modeling of Resonant and Reverberant Systems,” *J. Audio Eng. Soc.*, vol. 50, no. 12, pp. 1012-1029, 2002 Dec.

[5] M. Karjalainen and T. Paatero, “Frequency-Dependent Signal Windowing,” in *Proc. Workshop Applications of Signal Processing to Audio and Acoustics (WASPAA'2001)*, New Paltz, NY., USA, pp. 35–38, 2001 Oct.

[6] F. E. Toole and S. E. Olive, “The Modification of Timbre by Resonances: Perception and Measurement,” *J. Audio Eng. Soc.*, vol. 36, no. 3, pp. 122–141, 1988 March.

[7] S. E. Olive, P. L. Schuck, J. G. Ryan, S. L. Sally, and M. E. Bonneville, “The Detection Thresholds of Resonances at Low Frequencies,” *J. Audio Eng. Soc.*, vol. 45, no. 3, pp. 116–127, 1997 March.

[8] P. Antsalo, M. Karjalainen, A. Mäkivirta, and V. Välimäki, “Comparison of Modal Equalizer Design Methods,” Accepted to *AES 114th Convention*, Amsterdam, 2003 March.

[9] M. Karjalainen, P. Antsalo, A. Mäkivirta, T. Peltonen, and V. Välimäki, "Estimation of Modal Decay Parameters from Noisy Response Measurements," *J. Audio Eng. Soc.*, vol. 50, pp. 867–878, Nov. 2002.

[10] T. Saramäki, "Finite impulse response filter design," Ch. 4 in *Handbook for Digital Signal Processing*, ed. S. K. Mitra and J. F. Kaiser, Wiley 1993.

[11] A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U. K. Laine, and J. Huopaniemi, "Frequency-warped signal processing for audio applications," *J. Audio Eng. Soc.*, vol. 48, no. 11, pp. 1011-1031, Nov. 2000.

[12] J. O. Smith, *Techniques for digital filter design and system identification with application to the violin*, Ph.D. thesis, Elec. Eng. Dept., Stanford University, 1983.

[13] Churchill R. V., *Complex Variables and Applications.* McGraw-Hill, New York, 1960.

[14] T. Paatero and M. Karjalainen, "Kautz Filters and Generalized Frequency Resolution: Theory and Audio Applications," *J. Audio Eng. Soc.*, vol. 51, no. 1/2, pp. 27–44, Jan./Feb. 2003.

[15] A. Farina, "Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique," Preprint 5093, *AES 108th Convention*, Paris, France, Feb. 2000.