A METHOD FOR ORTHOGONAL AMPLITUDE AND DELAY PROCESSING OF SUBJECTIVE LISTENING TEST MATERIAL

Aki Mäkivirta, Jan Abildgaard Pedersen[a]

Genelec OY
Olvitie 5, 74100 Iisalmi, Finland
Tel: +358 17 813311, Fax: +358 17 812267
aki.makivirta@genelec.com

(a) Bang & Olufsen a/s
DK-7600 Struer, Denmark
Tel: +45 96 84 48 97, Fax: +45 97 84 12 50
jgp@bang-olufsen.dk

Abstract – We present a method to change amplitude and delay responses in subjective listening test material independently of each other. This is necessary in subjective listening experiments to apply modern statistical methods treating simultaneously several statistical variables. A case study of producing audio test material with this method is presented. This is related to an experiment where the audibility of the amplitude response variation and the delay response variation are studied at low frequencies based on data obtained from impulse responses of a room.

0. INTRODUCTION

Experimental designs based on using orthogonal arrays change more than one variable simultaneously with the aim to effectively determine the importance of several variables with a single test run. Orthogonal arrays were invented by Jacques Hadamar in 1897 and first used for statistical experimental design by Plackett and Burman [1]. The motivation is to minimise the size of an experiment, to control the amount of information required to be evaluated and to minimise the time and cost to execute an experiment. Examples of statistical techniques using this approach include various factorial designs such as Latin square type designs, balanced incomplete block designs, response surface methods and Taguchi techniques [1,2,3]. In this work we report the methodology for producing listening test material for an experiment related to orthogonal application of amplitude and phase deviations.

A room impulse response defines the transfer function $H_n(e^{jw})$ from a loudspeaker location to a listening point. This transfer function, later called a kernel filter, typically does not have a flat amplitude response or a constant group delay,

$$\begin{aligned} m_n(e^{jw}) &= \left| H_n(e^{jw}) \right| \\ \mathrm{grd_n}(e^{jw}) &= -\frac{\partial \arg\left[H_n(e^{jw})\right]}{\partial w}. \end{aligned} \tag{1}$$

In our case, the response surface method [2,4] is applied to study the audibility of amplitude and delay variation at low frequencies, below 200Hz [5,6]. The source data are simulated room impulse responses of a fixed size room for three different reverberation times, and additionally the anechoic case.

1. ORTHOGONAL CONVOLUTION FILTERS

Testing the audibility of amplitude and delay the orthogonal experimental design exhaustively combines amplitude and group delay properties of room impulse responses to produce a set of filters representing their desired combinations. One way to achieve this is to design two filters. Filter M_n is responsible only for the amplitude response variation and has linear phase. Filter G_k is responsible only for the delay variation and has allpass characteristic. The desired combination of the amplitude and phase characteristics is then produced as a convolution of these two filters. The resulting orthogonal convolution filters

$$H_{nk} = H(m_n, grd_k) = M_n\ G_k\,, \qquad n,k = \{1,2,3,4,\dots\} \tag{2}$$

are then used to produce the listening test items.

2. AMPLITUDE PROPERTY

The object is to produce a filter $M_n(e^{jw})$ having the amplitude behaviour of the kernel filter $H_n(e^{jw})$ at frequencies $w < w_o$,

$$\left|M_n(e^{jw})\right| = \begin{cases} \left|H_k(e^{jw})\right|, & w < w_o \\ c, & w > w_o \end{cases} \tag{3}$$

and a linear phase property

$$\mathrm{grd}\left(M_n(e^{jw}) \right) = d, \qquad 0 < w < p\ . \tag{4}$$

The desired magnitude response given by Equation 3 is interpolated onto a dense, evenly spaced grid. A symmetric set of filter coefficients is obtained by applying an inverse fast Fourier transform to the grid and multiplying by a Hamming window, resulting in a linear phase FIR filter having a constant delay *d*.

In practice it is necessary to design the different filters M_n such that the constant gain levels *c* are the same for frequencies $w > w_o$ for all filters to maintain right gain relationships for frequencies $w < w_o$ and avoid false cues after the processed listening test signals have been level aligned at $w > w_o$. The linear phase design of the amplitude property filter has a constant delay, which will add the total delay of the final convolution filter.

Although we use the amplitude response characteristics of the kernel filter below a certain frequency (lowpass formulation), bandpass or highpass formulations of the amplitude response property are equally possible.

3. GROUP DELAY PROPERTY

The object is to produce a filter $H(e^{jw})$ having delay variation of the kernel filter $H_n(e^{jw})$ at frequencies $w < w_o$,

$$\mathrm{grd}\left[H(e^{jw})\right] = \begin{cases} \mathrm{grd}\left[H_n(e^{jw})\right] + d_0\,, & w < w_o \\ d_1\,, & w > w_o \end{cases} \tag{5}$$

and allpass amplitude characteristic

$$\left|H(e^{jw})\right| = 1, \qquad 0 < w < p\ . \tag{6}$$

To obtain this filter, we use a method where a lowpass filter $H_{LP}(z)$, having a delay N, selects part of the kernel filter response into $H_t(z)$.

$$H_t(z) = H(z) \cdot H_{LP}(z) + \left(z^{-N} - H_{LP}(z) \right) . \tag{7}$$

This filter has the desired delay variation characteristics but it does not have allpass characteristic. If the amplitude response of $H_t(e^{jw})$ does not contain zeros, we can design a whitening filter $H_i(z)$,

$$G_k(z) = H_t(z) \cdot \left[\frac{1}{\left| H_t(z) \right|} \right] = H_t(z) \cdot H_i(z) \tag{8}$$

such that

$$\left| G_k(e^{jw}) \right| = 1, \qquad 0 < w < p . \tag{9}$$

This whitening filter is designed using the method described for obtaining the amplitude response property, and it has linear phase.

We used FIR approximations that were forced causal by allowing enough constant delay d_0 to model the delay variation required. In this way, this method can model the delay variation within the selected frequency band, but the absolute delay of the room impulse response can not be modelled.

If in a listening test the delay property is switched in and out, it is necessary to compensate for the constant delay term d_1 to avoid the listeners from being mislead by the additional cue of the time shift created by this constant delay.

Although above we use the delay characteristics of the kernel filter below a certain frequency (lowpass formulation), bandpass or highpass formulations of the delay property are equally possible.

4. CASE STUDY

Our case study is an experiment [5,6] where the target it to investigate the importance of the amplitude and the delay variation, or ripple, at low frequencies. As the amplitude and delay ripple may also interact, the experimental design formulation exploiting orthogonal arrays is interesting because it allows us to study these interactions. The down side is that the preparation of listening test material is complicated by the requirement to dissociate the amplitude variation from the delay variation.

Room impulse responses used as the source data originate from simulation of an IEC 268-13 compliant room. The shape of the room is maintained in the simulation and the reflection coefficient of room boundaries is used to vary the reverberation time of the room. Four values for reverberation time are used, the anechoic case and reverberation times RT=200ms, 400ms and 800ms [5]. This results in four impulse responses (Fig. 1), each with their natural amplitude and delay variation. The frequency response of the radiating loudspeaker is eliminated by deconvolution with the anechoic impulse response. The anechoic response represents the properties of the radiating loudspeaker.

The resulting impulse responses show significant frequency dependent changes both in amplitude (Fig. 2) and delay (Fig. 3). The peak-to-peak amplitude variation in frequencies below 200Hz varies from 15dB (RT=200ms) to 35dB (RT=800ms) and we can recognise the characteristic peaks and notches of room amplitude response measurements. We also see the variations in delay typical to a room at these extremes.

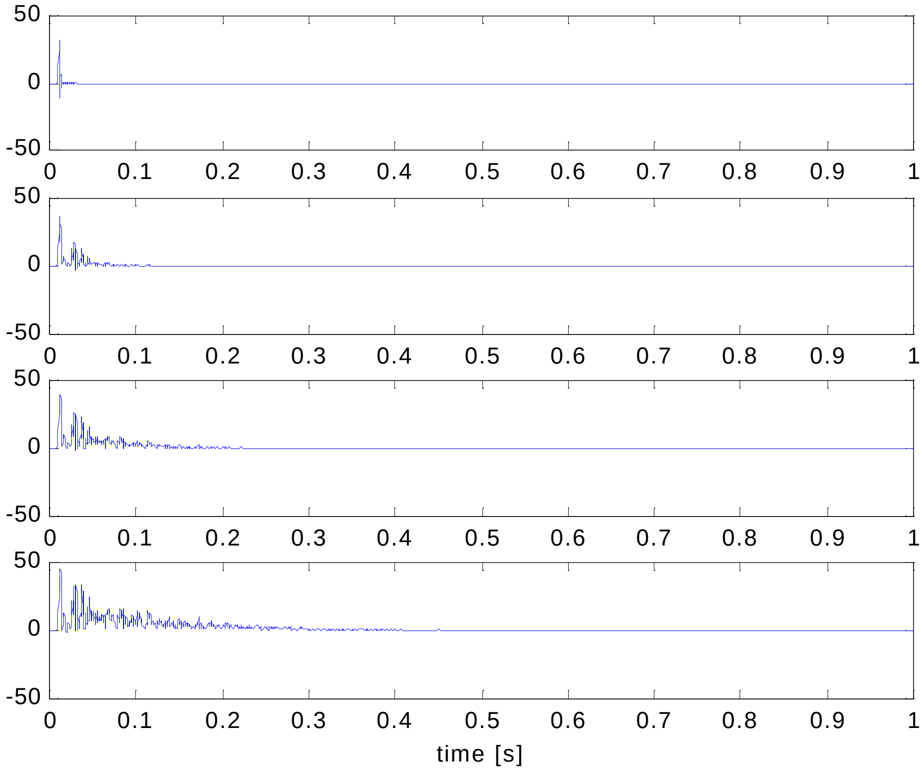


Fig. 1. Room impulse responses simulated by varying the absorbance of room surfaces, from top to bottom: anechoic condition (reverberation time 0ms), 200, 400 and 800ms.

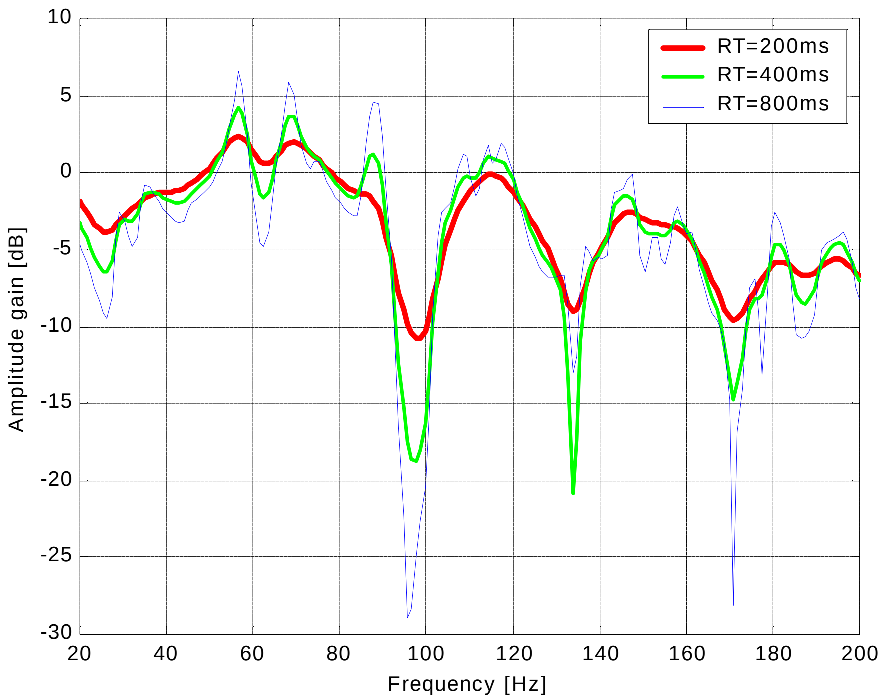


Fig. 2. Amplitude changes in the simulated room impulse responses.

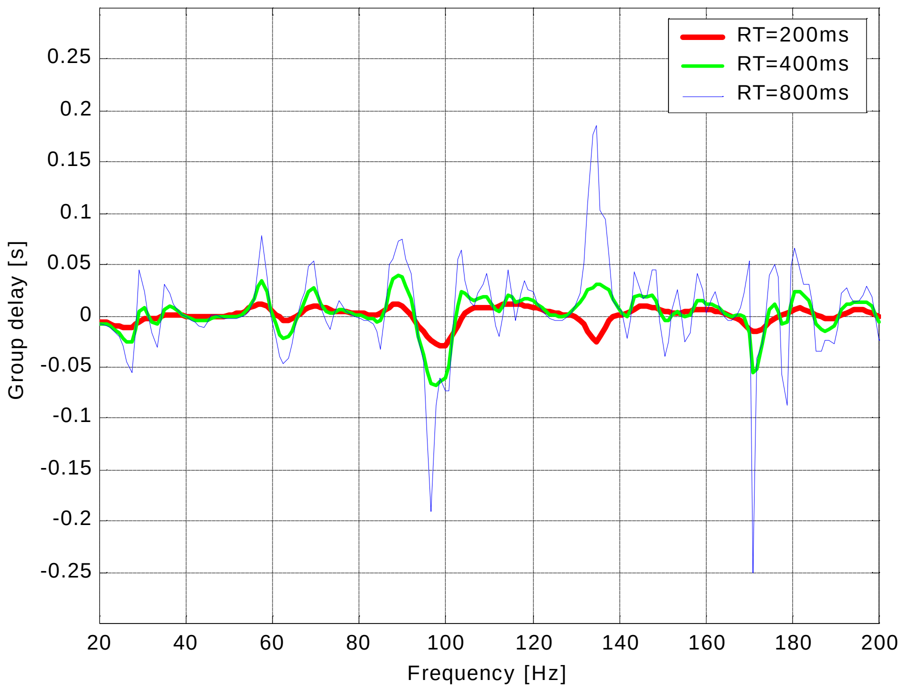


Fig. 3. Delay variation in the simulated room impulse responses.

Amplitude property

The amplitude property filter (Fig .4) is implemented as a linear phase FIR filter. The amplitude in the impulse responses is conditioned by removing frequencies <5Hz, essentially the DC component. The subjective effect is limited below a corner frequency f_c of either 80Hz or 120Hz by removing the amplitude ripple above these frequencies. The removal is implemented by modifying the filter design target. A filter is designed for each RT and corner frequency f_c, resulting in six filters (RT=200ms, 400ms and 800ms, f_c=80Hz and 120Hz).

Delay property

The filters to implement the delay property (Fig. 5) are designed by modifying the target function above the corner frequency f_c to have a constant delay. The design method produces a FIR filter with allpass characteristics. Note that the real time delay that makes the filters causal has not been removed in Fig. 5. For the case fc=80Hz the delay changes to a constant value between frequencies 70…100Hz, and for the case fc=120Hz the delay changes to a constant value between frequencies 80…120Hz.

Orthogonal filters

The experimental design produces a full permutation set of amplitudes and ripples. A total of 32 filters are produced for this experimental design. Fig. 6 depicts some examples of the resulting FIR filters, which are then convolved with the audio test material to produce the listening test items.

Typical modelling errors are depicted in Fig. 7 for the convolution filters in Fig. 6. The lower corner frequency for the modelling is 20Hz. The higher corner frequency for modelling is either 80Hz or 120Hz. The modelling error is controlled between these frequencies. Above the higher corner frequency the modelling error becomes very large because the amplitude and delay assume constant values at these frequencies.

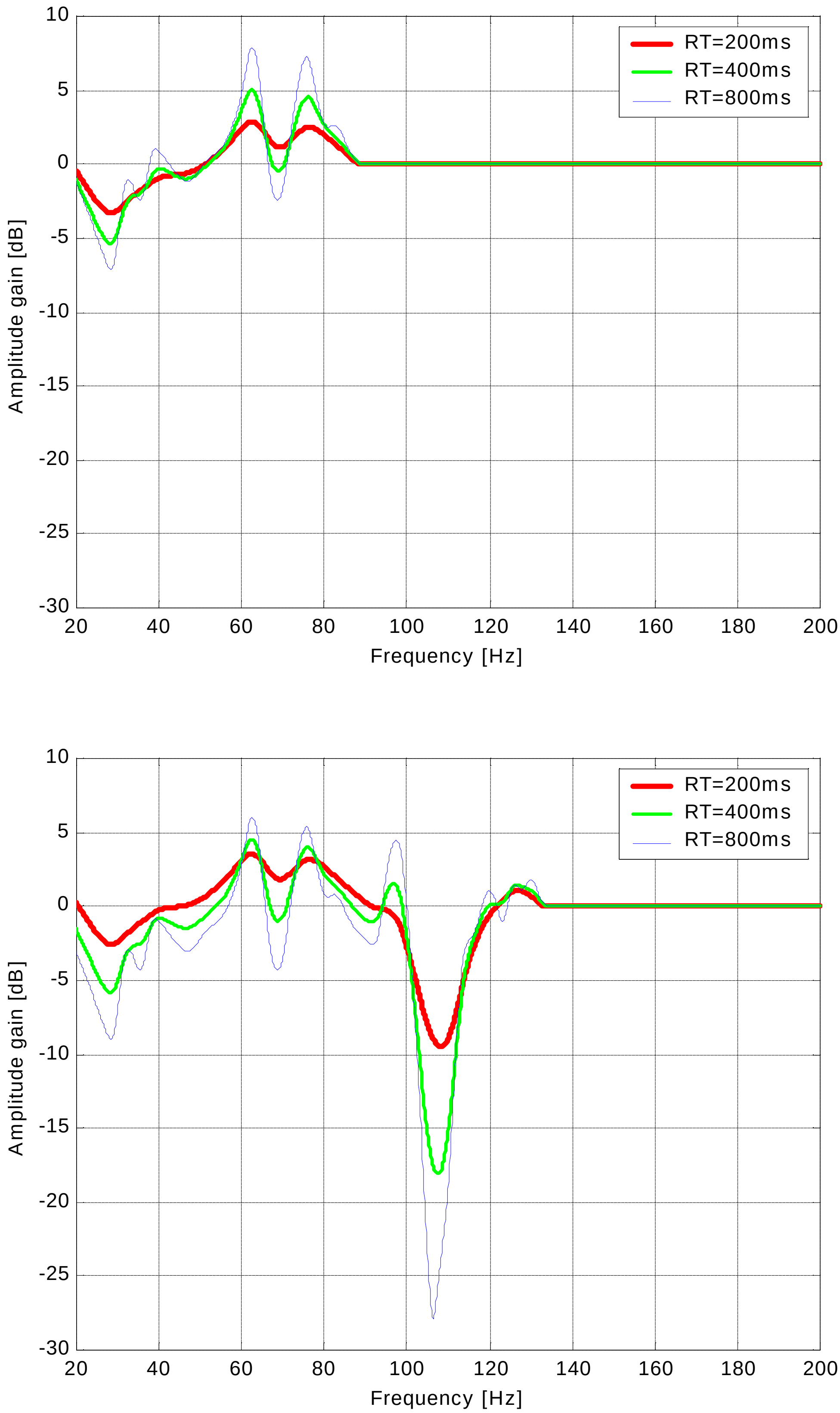
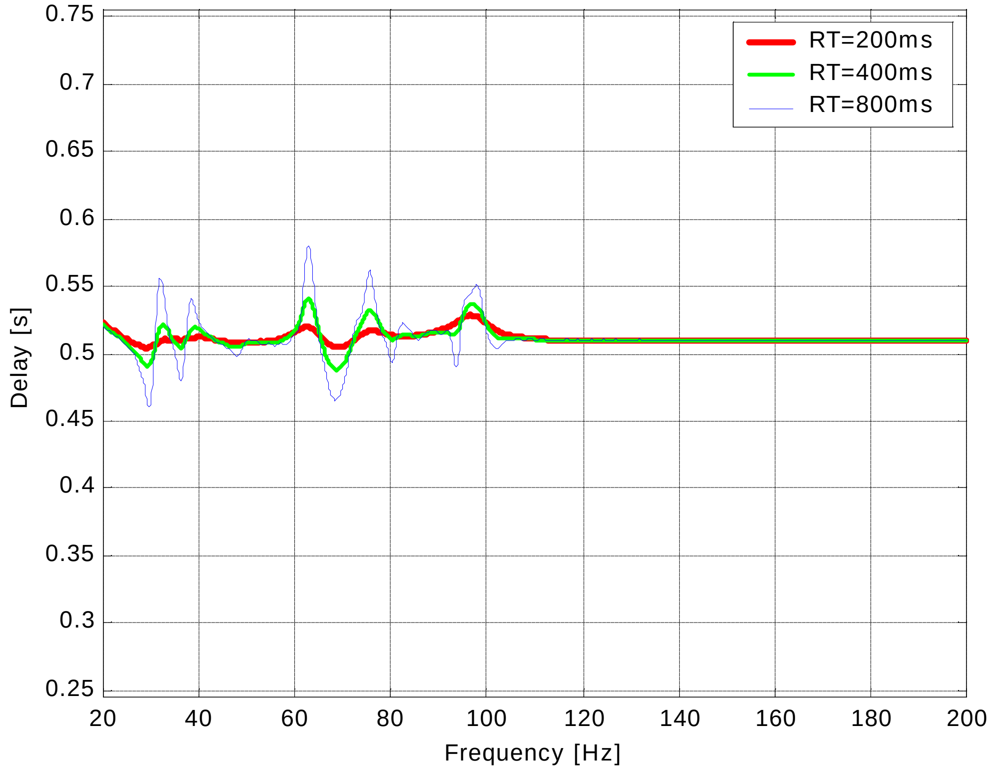


Fig 4. Amplitude filters for fc=80Hz (top) and fc=120Hz (bottom).



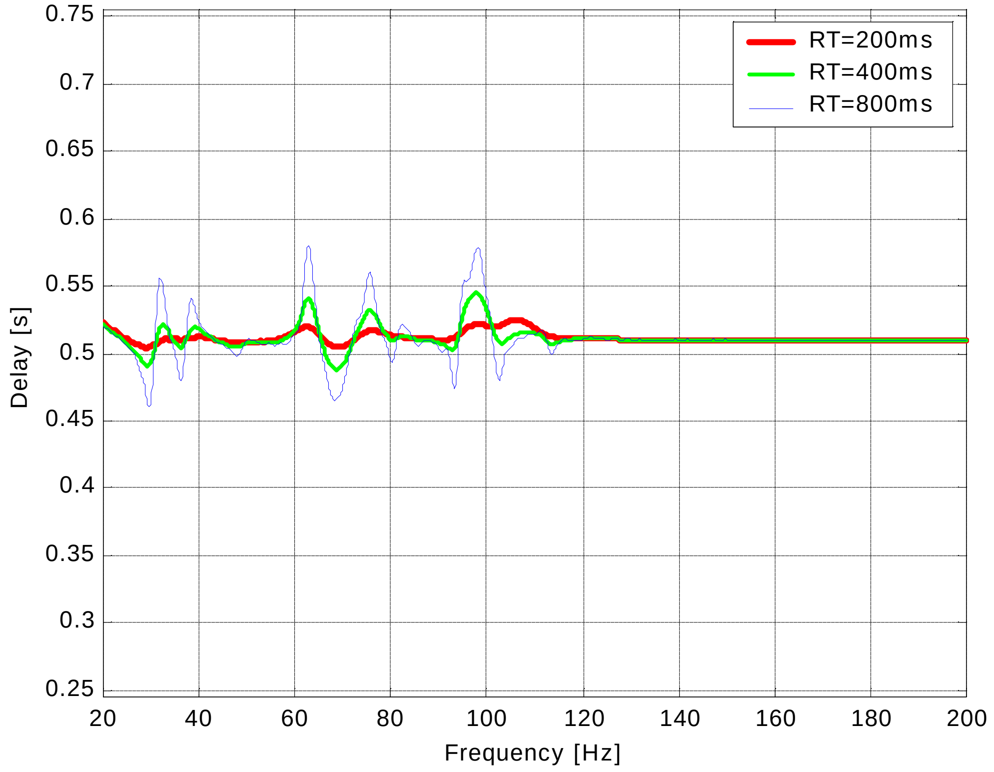


Fig. 5. Delay filters for fc=80Hz (top) and fc=120Hz (bottom).

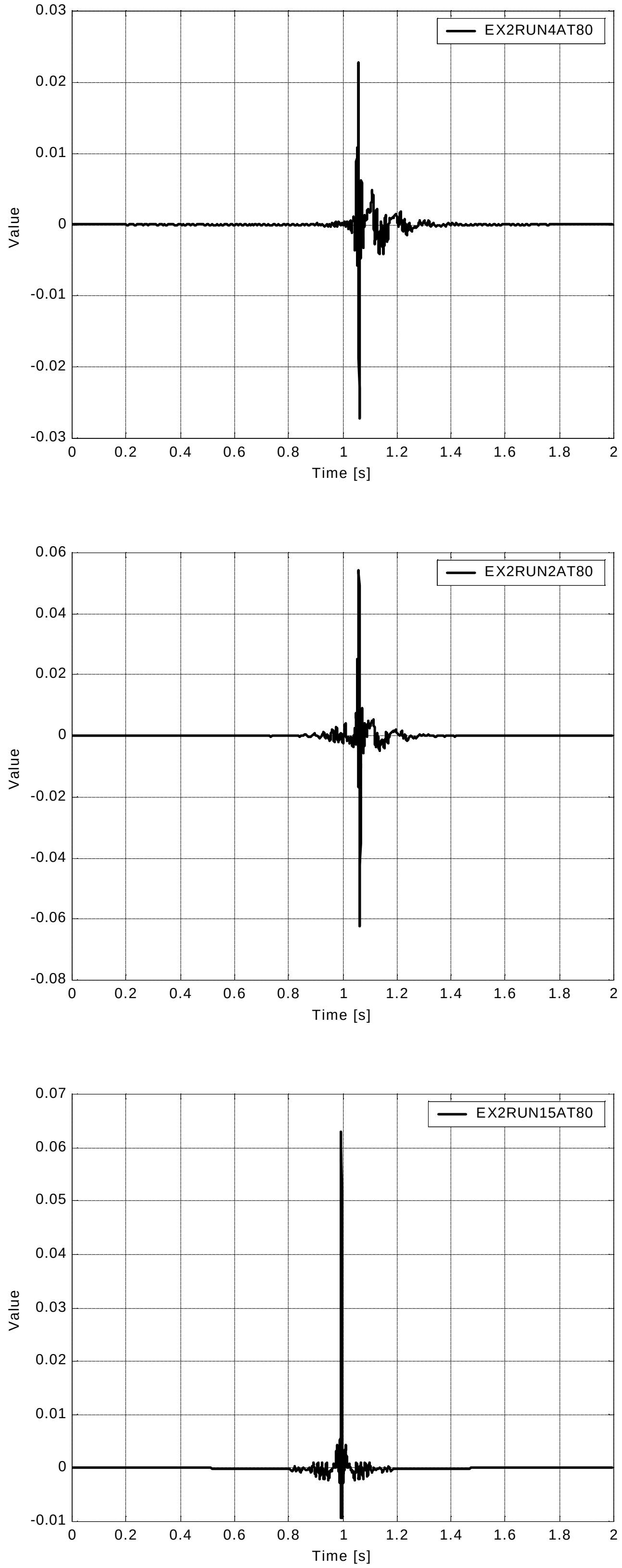


Fig. 6. An example of convolution FIR impulse responses of room impulse response RT=800ms, fc=80Hz. Both amplitude and phase (top), allpass or only delay (middle), linear phase or only amplitude (bottom).

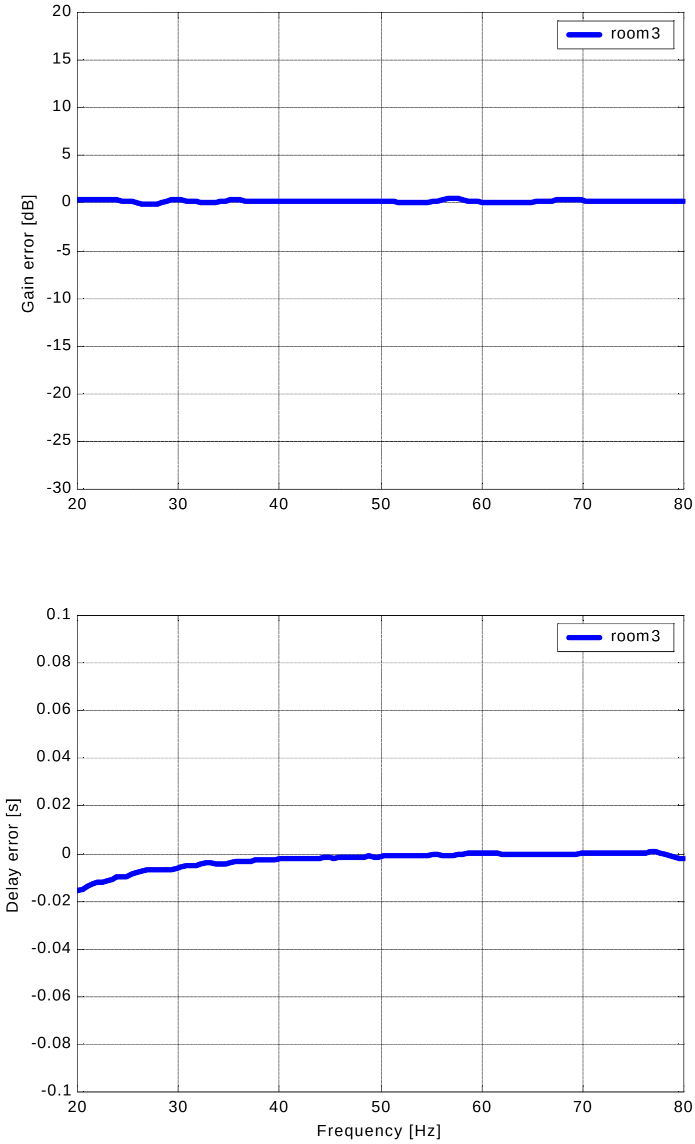


Fig 7. Typical amplitude modelling error (top) and delay modelling error (bottom).

Convolution processing

The orthogonal filters are synthesised at the sampling frequency of 4410Hz, 1/10 of the audio sampling frequency for the listening test items. The convolution process is depicted in Fig. 8. The audio material is crossover filtered at 150Hz by $H_{xo}(z)$. This crossover filter is a 12th order maximally flat zero-phase implementation. The roll-off speed is fairly fast, and the lowpass section has 100dB attenuation at 400Hz. The attenuation is sufficient to remove aliasing in the decimation process. Zero phase implementation is possible because the audio material is processed off-line in a computer, allowing convenient time-inversion.

The lower band is decimated by 10 to make the convolution processing more efficient, then interpolated, filtered and added with the delayed and level aligned high frequency band to reconstruct the full-bandwidth processed audio material. The anti-imaging filter $H_r(z)$ is needed as a part of the reconstruction process for removing the images of the fundamental frequency band generated by the interpolation process. The convolution is implemented off-line in a PC computer, and all signal processing is done in 64-bit floating point representation, but some intermediate storing is done in 16-bit PCM format. The conversion to 16-bit PCM is done with dithering before truncation to minimise any distortion effects.

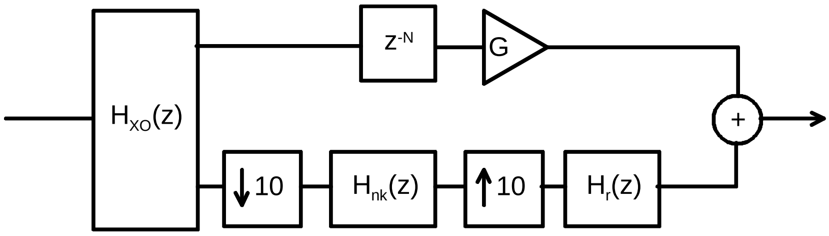


Fig. 8. Block diagram of the convolution process.

5. DISCUSSION

Listening test techniques try to identify not only the primary influences but also their possible interactions. Modern statistical techniques based on orthogonal arrays provide a method of achieving this goal. We have presented a method to apply these techniques to studying ripple in amplitude and delay. The amplitude and delay ripples are modelled separately as FIR filters. While FIR filters may not be computationally the most effective they are flexible and easy to design and provide the additional advantage of a precisely linear phase.

Multirate processing makes the convolution feasible in a rational amount of time and computer power. The disadvantage is that additional design steps are needed to align the highpass path with the lowpass path both in level and phase. This is made easier in our implementation by constraining the orthogonal filters to have a flat amplitude response and a constant integer delay at frequencies near the crossover.

6. ACKNOWLEDGEMENTS

This work is a part of the Eureka project EU1653 Medusa (Multichannel Enhancement of Domestic User Stereo Applications). The Medusa project is a 3.5-year joint research project of the following parties: Bang & Olufsen, British Broadcasting Corporation, Genelec, Institute of Sound Recording at the University of Surrey, and Nokia Research Center. The aim of this project is to examine the variables of the domestic multi-channel system, with and without picture, to carry out the essential optimisation leading to consumer end-products.

7. REFERENCES

1. Phillip Ross: "Taguchi Techniques for Quality Engineering: Loss Functions, Orthogonal Experiments, Parameter and Tolerance Design", McGraw-Hill, New York, 1988.
2. George E. P. Box, William G. Hunter, J. Stuart Hunter: "Statistics for Experimenters", John Wiley & Sons, New York, 1988.
3. G. Taguchi, Y. Yokoyama: "Taguchi Methods: Design of Experiments", American Supplier Institute, Dearborn MI, in conjunction with the Japanese Standards Association, Tokyo, Japan, 1994.
4. G. E. P. Box, N. R. Draper: "Empirical Model-Building and Response Surfaces", Wiley & Sons, New York, 1987.
5. Jan Abildgaard Pedersen and Aki Mäkivirta: "Quantification of Subwoofer Requirements, Part I: Generation of Stimuli and Listening System Equalization", AES 108th Convention, Paris 19-22 Feb, 2000.
6. Søren Bech: "Quantification of Subwoofer Requirements, Part II: The influence of Lower System Cut-off Frequency and Slope, and Pass-band Amplitude and Group Delay Ripple", AES 109th Convention, Los Angeles 22-25 Sep, 2000.