**Comments on Kirkeby Classical Inverse Filtering Method with Regularization**

By Bohdan Raczynski

De-reverberation method proposed by Kirkeby [2] in 1999 was based on transforming the room-loudspeaker transfer function into the frequency domain with the application of FFT, then performing simple division of spectral values and finally returning the result back to time domain via inverse FFT.

Assuming the following notation:

$$C(\omega) = a(\omega) + jb(\omega) = Measured\ \mathrm{Re}\,sponse$$
$$H(\omega) = Inverse\quad Filter$$
$$F(\omega) = Corrected\ \ System$$

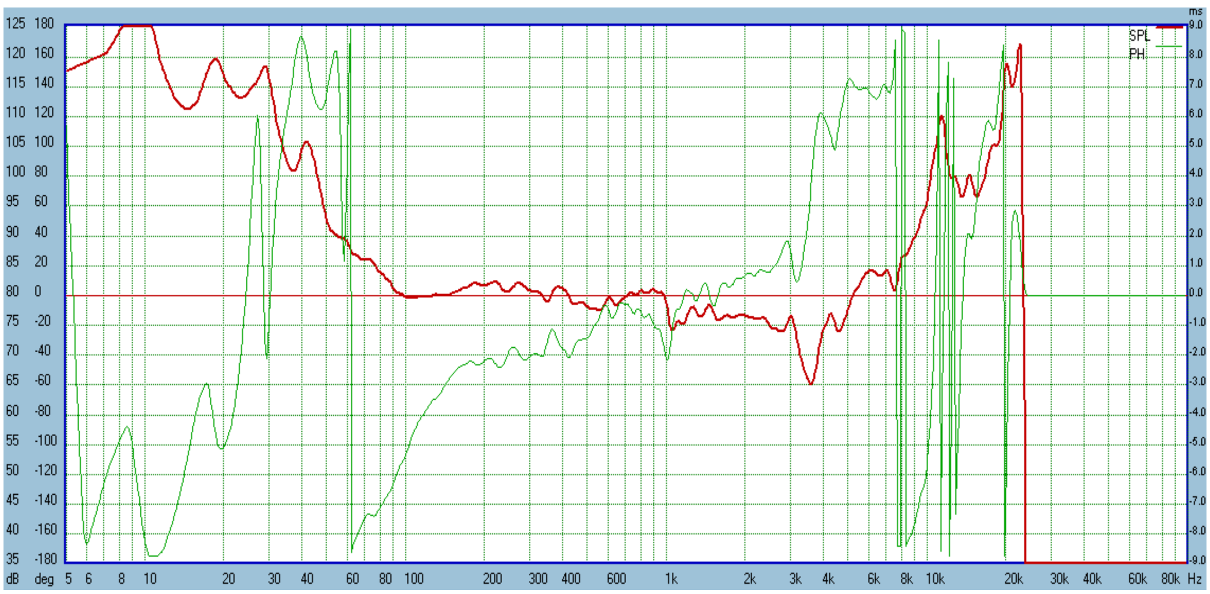


Figure 1. The simplest “inverse filter”.

The inverse filter created by simply inverting the original, measured frequency response would have major drawbacks – see Figure 1. It would require very large gain at low and high frequencies to equalize the system. A solution method was proposed by Kirkeby [1][2], and involved introduction of regularization mechanism to isolate frequency range of interest, where the equalization was to take place.

Commonly used expression for Kirkeby’s regularization method is shown below.

$$H(\omega) = \frac{C^*(\omega)}{\left|C(\omega)\right|^2 + \beta\,\mathrm{Re}\,g(\omega)} \qquad where \quad C^*(\omega) = a - jb \quad (conjugate), \quad \beta = adjustable\ \ strength$$

The idea behind regularization is, that when the term $\beta$Reg($\omega$) is small in comparison with $|C(\omega)|^2$, the regularization will have little effect and the inverse filter should fully correct the measured response. This will happen within the frequency range where the Reg($\omega$) = 0.

Conversely, when the term $\beta$Reg($\omega$) is large in comparison with $|C(\omega)|^2$, the regularization will exhibit dominant effect in the denominator and the inverse filter will not correct the measured response. This will happen within the frequency range where the Reg($\omega$) = 1.

By making the $\beta$Reg($\omega$) frequency dependant, we could therefore nominate some desirable frequency range, where we would wish the equalization to take place, and at the same time, leave the remaining spectrum unchanged. The strength of the regularization is determined by parameter $\beta$ and the operational frequency range is determined by Reg($\omega$).

This approach is used by Farina [5], Fielder [6],[7], Norcross, Bouchard and Soulodre [3].

**It appears, the there are four side-effects of classical regularization method:**

1. Frequency response outside the equalization range is affected by filter's operation.
2. Phase response of the corrected system is always forced to linear phase. Without some additional mathematical effort, there is no option to perform minimum-phase correction.
3. Linear-phase systems tend to have symmetrical impulse response. This leads to excessive equalizer latency. This effect will manifest itself more strongly when good low-frequency resolution is required for room equalization, therefore longer impulse response is required = large latency. Maximum tolerable audio-video latency is 185ms.
4. Increased pre-ringing. Since the frequency response outside equalization range is "squared" (double the slope rate), then the pre-ringing will be increased because it's dependent on the slope rate.

In this discussion, the regularization will attempt to use Reg(w) curve shown on Figure 2, for the purpose of equalizing the loudspeaker shown on Figure 3.

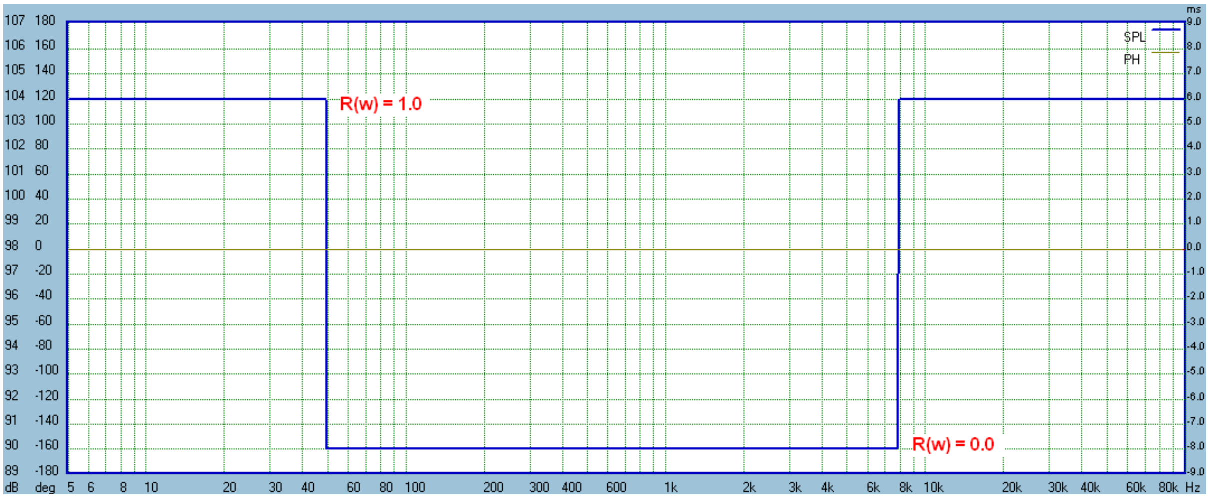


Figure 2. Regularization curve, Reg(w), determining EQ range: 50Hz-8000Hz.

Figure 3. Measured loudspeaker – ready for equalization.

As mentioned before, we start with small value of β, and increase this value up to such point, where **only the desired 50-8000Hz frequency range of the loudspeaker is corrected**, but without excessive distortion of the remaining frequency spectrum. It is also anticipated, that bass response (rolling off below 100Hz – blue curve on Figure 3) will be equalized from 50Hz up. The original low-frequency roll-off is 18dB/oct, and the enclosure alignment is QB3.

This process is depicted on Figure 4a – h. As expected, initially (**β = 0.0)** the inverse filter (green curve) fully corrects the loudspeaker (Figure 4a). The requested frequency range begins to manifest itself on Figure 4e, and is probably acceptable on Figure 4g – this is what is assumed to be the optimal strength of **β = 6.0**. This regularization strength gives the near-perfect alignment of corrected and uncorrected SPL curves between 8kHz and 10kHz - just outside the filter's operating band. By the same token, it is assumed, that the system shown on Figure 4h is over compensated.

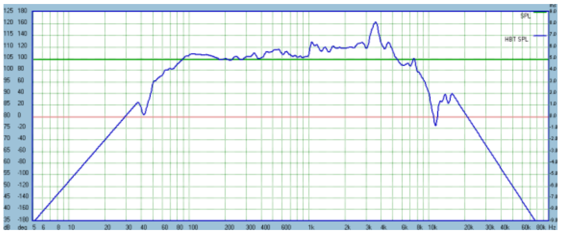


Figure 4a

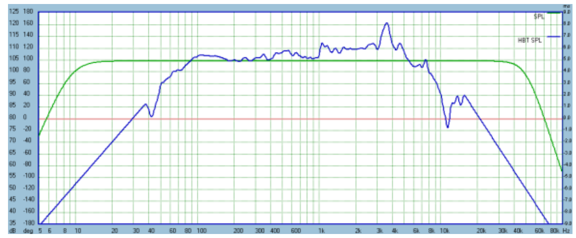


Figure 4b

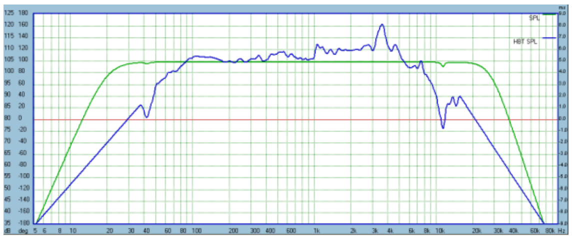


Figure 4c

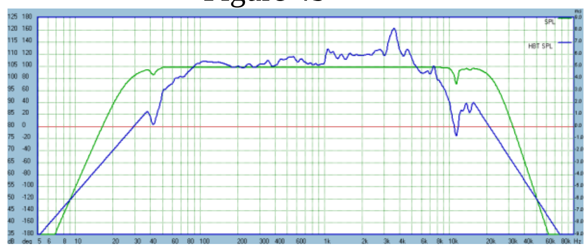


Figure 4d

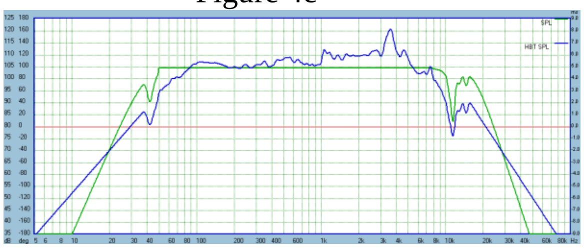


Figure 4e

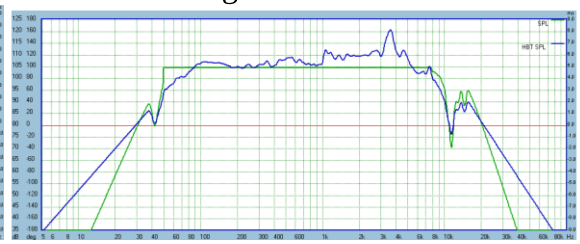


Figure 4f

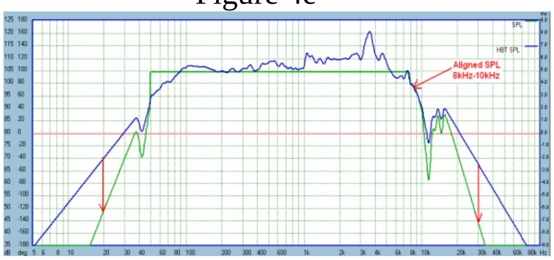



Figure 4g

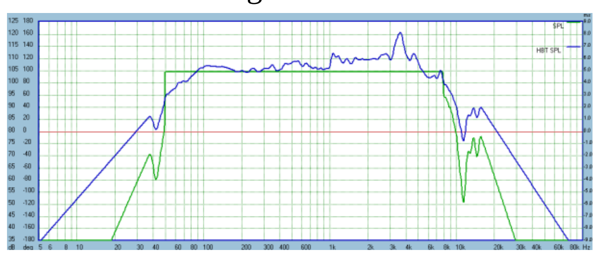


Figure 4h

**Examination of Issue 1, presented on page 2.**

It is observable on Figure 5 below, that the loudspeaker (blue curve) has been equalized within the desired frequency range of 50-8000Hz., However, it appears that the corrected system (green curve) now has faster roll-off and more “bumpy” response outside the desired frequency range. Steeper slopes are visible for all values of β - see Figure 4b-h. Steeper slopes and bumps are highlighted by red arrows on Figure 5.

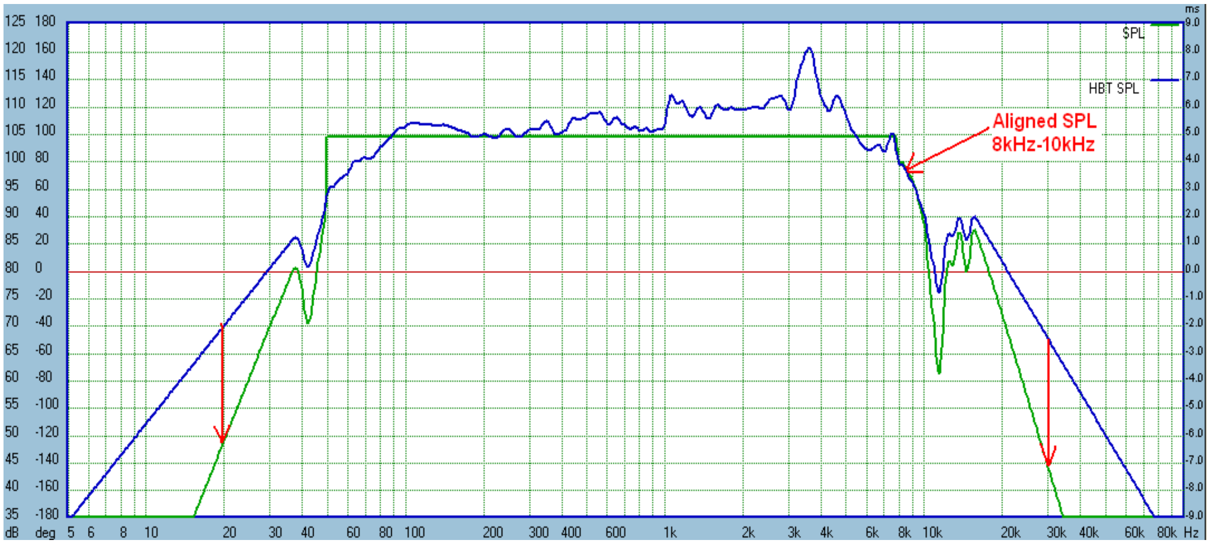


Figure 5. Comparison of equalization effects outside the inverse filter's bandwidth.

Some insight into this mechanism is offered by comparing magnitude responses of the uncorrected and corrected sections. The "raw", uncorrected loudspeaker magnitude response is:

$$|C(\omega)| = \sqrt{a^2 + b^2} \qquad \textbf{Formula (1)}$$

The corrected system looks as follows:

$$|F(\omega)| = |C(\omega)H(\omega)| = |(a+jb)A(a-jb)| \quad where \quad A = \frac{1}{a^2 + b^2 + \beta \operatorname{Re} g(\omega)}$$

**A.** When the filter operates within it's intended frequency range, βReg(w) = 0. Therefore,

$$|F(\omega)| = \frac{(a+jb)(a-jb)}{a^2+b^2} = \frac{a^2 - jab + jba + b^2}{a^2+b^2} = \frac{a^2+b^2}{a^2+b^2} = 1 + j0 = 1$$

In this case the imaginary part is zero, and the magnitude response of the corrected system equates to unity – this is what is depicted by the green curve on Figure 5 within 50-8000Hz.

**B.** When the filter operates outside it's intended frequency range, $a^2+b^2$ can be neglected and βReg(w) >> $a^2+b^2$ is **a constant value** (=1 on Figure 2). Therefore,

$$F(\omega) = \frac{(a+jb)(a-jb)}{a^2+b^2+\beta \operatorname{Re} g(\omega)} = \frac{a^2 - jab + jba + b^2}{\beta \operatorname{Re} g(\omega)} = \frac{a^2+b^2}{\beta \operatorname{Re} g(\omega)} = A(a^2+b^2+j0)$$

Since we are only interested in the frequency response fluctuations, we can disregard the constant A in the above equation. The value of A is independent of frequency below 50Hz and above 8000Hz, and will not affect the frequency response irregularities. Now, we can finally evaluate the magnitude response irregularities outside the filter's operating range:

$$|F(\omega)| = A\sqrt{(a^2+b^2)^2+0^2} = A(a^2+b^2) \qquad \textbf{Formula 2}$$

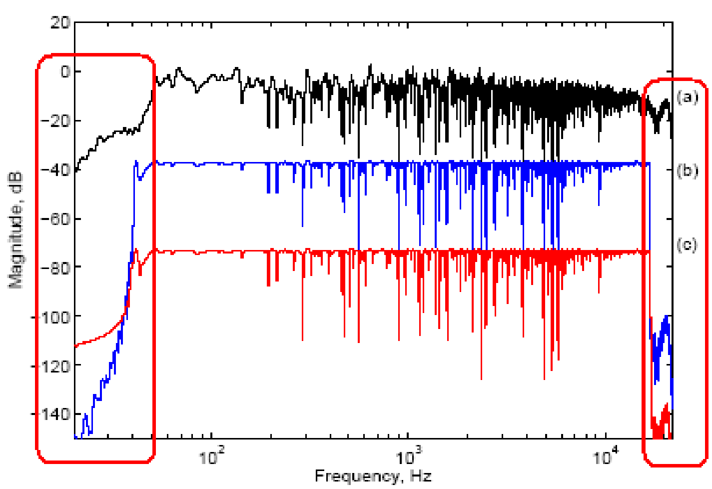
We can now compare Formula 1 and Formula 2, and conclude, **that outside the filter's operating bandwidth, the regularization method causes "squaring" of the original amplitude response.** Therefore –18dB/oct slope will become –36dB/oct slope, 12dB peak becomes 24dB peak, and so on.

The same issue is visible in Farina's [5] measurements – see below.


Fig. 16 – measured frequency response of the artificial mouth system


Fig. 19 – measured frequency response of the artificial mouth system after equalization

And also in Norcross, Bouchard and Soulodre [3].



It may also be beneficial to compare the results from Kirkeby algorithm (Figure 6) to the same equalization process, but performed by Inverse HBT equalization process depicted on Figure 7 below and described in [7].

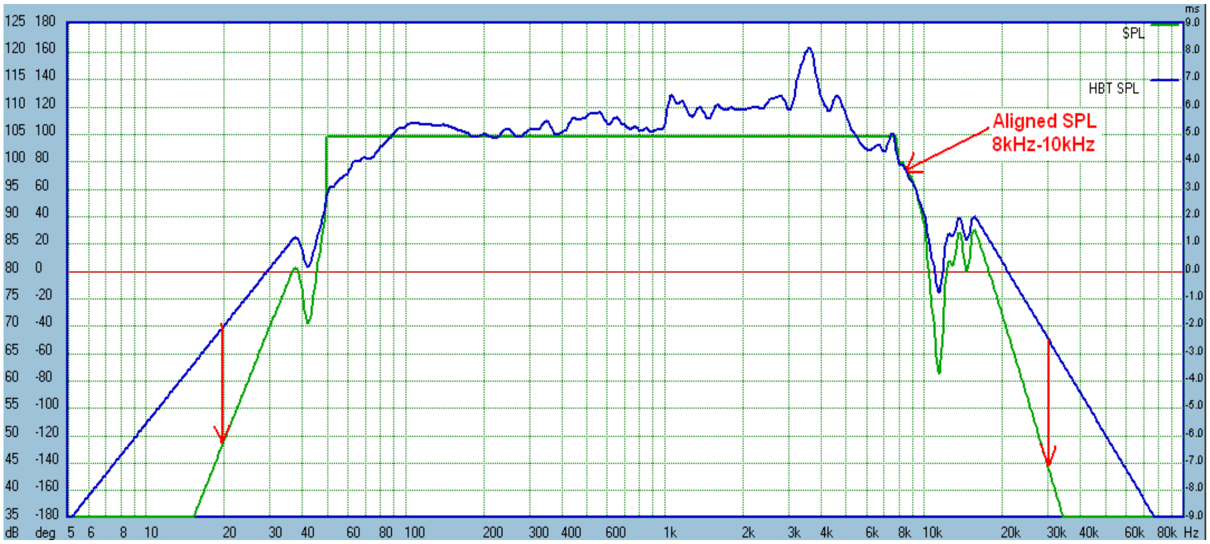



Figure 6. Loudspeaker equalized within 50Hz and 8000Hz using Kirkeby Inverse Filter. **Note the rapid step at 50Hz**.

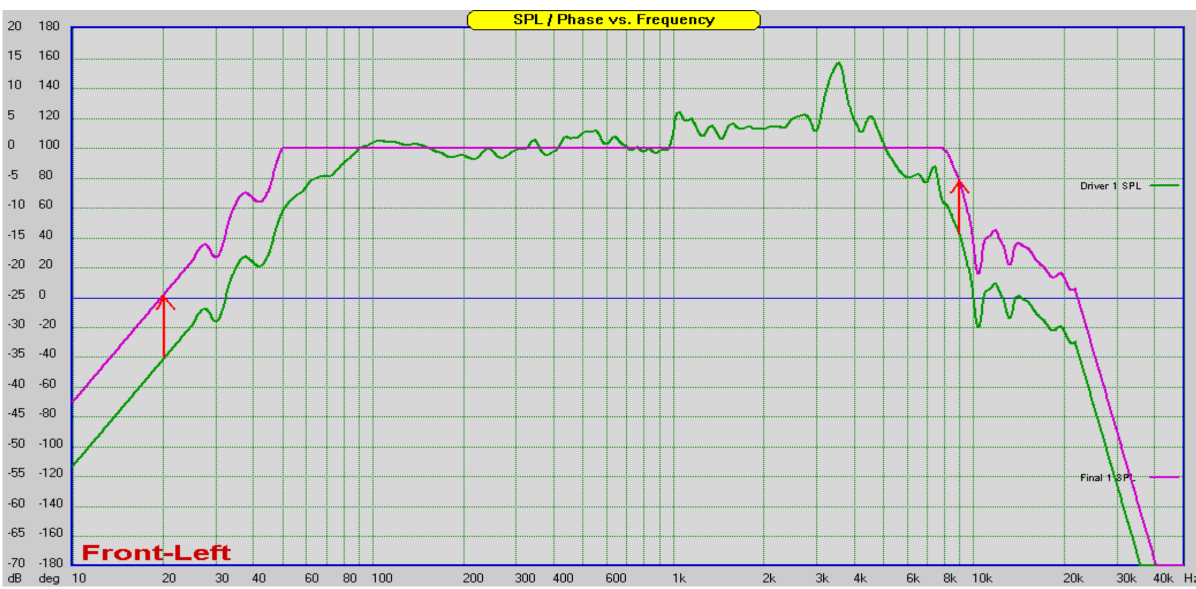



Figure 7. Loudspeaker equalized within 50Hz and 8000Hz using Inverse HBT method.

It is clearly observable, that the Inverse HBT equalization process will maintain both: the original slopes, and the shape of the SPL curve outside the equalization range. In other words – there is no “amplitude squaring” effect. In addition, the frequencies outside the equalization range are actually boosted, rather than progressively attenuated.

**Examination of Issue 2, presented on page 2.**

Phase response of the corrected system is always forced to linear phase. Without some additional mathematical effort, there is no option to perform minimum-phase correction. Linear-phase is generally not a detriment in loudspeaker systems – quite the opposite. It has advantages resulting in tighter bass and helps with localization accuracy, as it produces excellent transient response. Time-domain responses of such system are excellent.

On the negative side, in some instances, the pre-ringing effect may manifest itself audibly. Some cases were discussed in [8].

It appears, that without some mathematical intervention into the classical Kirkeby's method, the process may attract some criticism for being forced to be only a linear-phase system. This is exactly what happened in Norcross, Bouchard and Soulodre [3]. With the help of Hilbert Transform, they proposed new regularization function, which is capable of elimination pre-response to same degree. It is unknown, if this new approach will also correct system's latency to the same level as exhibited by standard minimum-phase systems.

The equalizer should be able to perform both: minimum-phase and linear-phase equalization.

**Examination of Issue 3, presented on page 2.**

Linear-phase systems tend to have symmetrical impulse response. This leads to excessive equalizer latency. This effect will manifest itself more strongly when good low-frequency resolution is required for room equalization, therefore longer impulse response is required = large latency. When equalizing 5.1HT or 7.1HT system, one needs to observe maximum tolerable audio-video latency, which is 185ms.

Low-frequency resolution (LF_Res) is linked to sampling frequency and impulse response length by:

$$LF\ \mathrm{Re}s = \frac{Sampling\ Frequency}{IR\ Length}$$

For example, when sampling with 48000Hz, and using impulse response length of 8192, we can obtain LF_Res of 48000/8192 = 5.86Hz. This bass resolution is sufficient for most of the loudspeaker equalization, but for more accurate room equalization, it may have to be improved.

Additionally, we need to determine partition length, which the convolution engine will use in real-time. If we assume partition length of 1024, we will need 8 partitions to cover the impulse response above. If we use linear-phase approach, the center of the impulse response will be located at 4 x 21.33ms = 85.36ms. Since we are using "block processing" approach, we need to read-in some data block before they can be processed. This takes time and when we add typical sound card buffer processing delay, and IRQ processing delays, we are likely to end up with total latency of 145ms (60ms more than the shift of the impulse response peak). This latency is still quite acceptable, as we are well within the tolerable limit of 185ms.

So, what would happen if we decide to double the LF_res?.

Now, the LF_Res = 2.93Hz. The impulse response length is now 16384, and we need to double the number of partitions to 16, having the center of the impulse response located at 8 x 21.33ms = 170.72ms. On the top of this delay, we need about 60ms to cover for all other delays, and the result will be 230.72ms – well outside the tolerable limit. If nothing else changes in the example above, and we need to improve bass resolution again (LF_Res=1.47Hz), we are likely to incur system latency of 460ms.

However, none of these issues matter when minimum-phase approach to equalization/filtering is used. The peak of the impulse response in minimum-phase system is typically located very close to the start of the impulse. In this case, improving the bass resolution and increasing the length of the tail of the impulse response by allowing more partitions works very well, as it does not increase latency at all. However more CPU power will be required though.

**Examination of Issue 4, presented on page 2.**

This is possibly the most detrimental aspect of the regularization method. Linear-phase systems are characterized by symmetrical impulse response. Therefore, the impulse response will exhibit pre-ringing. Pre-ringing can be controlled to large degree, and relevant explanations have been given in the following papers:

http://www.bodziosoftware.com.au/Pre_Post_Ringing_IR_And_Pulses.pdf

http://www.bodziosoftware.com.au/HT_Loudspeakers_SPL_Polar_Comparison.pdf

The slopes are doubled with regularization method, therefore the pre-ringing will be increased. Is highly undesirable, to increase pre-ringing with regularization method, while the HBT method does not increase the pre-ringing.

## Conclusions

Classical regularization method for inverse filtering was briefly examined. Three issues were identified, with some evidence, that other audio researchers are also aware of the problems.

The linear-phase system's impulse response issues are not unique to the discussed algorithm, and it seems, that any such audio filtering scheme, will eventually have to support it's existence with Hilbert Transform or the "Law of Bode" approach (also known as HBT in DIY circles).

The distortions of the amplitude response outside filter's bandwidth appear to be the consequence of the discussed algorithm. The distortions can be filtered out for band-pass loudspeakers, but it will be difficult to avoid them on low-end of the subwoofer and high-end of the tweeter drivers, as these tend to be unfiltered. It would be undesirable to have a vented subwoofer +24dB/oct LF roll-off increased to +48dB/oct simply by trying to slightly extend it's frequency response outside it's operating bandwidth.

If bass response is to be equalized, the discussed equalizer may produce rapid step in the frequency response – as shown of Figure 6. Since the undesired pre-response is related to steepness of the slopes, there is a possibility, that the algorithm will increase pre-ringing. Fortunately, these issues can be completely avoided, as described in [9].

Audio-video latency issues are typically difficult to resolve. In linear-phase filtering approach it is the video signal that needs to be delayed. In many instances, the complex video signal processing performed by flat-panel display screens, does provide some welcoming delays. Other than that, the filtering scheme can be run in minimum-phase mode, where latencies well under 100ms are quite feasible. This would further reinforce the need for both types of equalization: minimum-phase and linear-phase.

Obviously, if the audio system is only intended for music playback, then the latency issue does not come into focus at all. With the help of partitioned convolution, one is able to employ really long impulse responses and achieve outstanding bass resolution, going into fractions of a Hertz.

Thank you for reading

Bohdan

**References:**

[1] "**Digital Filter Design for Inversion Problems in Sound Reproduction**" Ole Kirkeby and Philip A. Nelson, JAES Vol 47, No 7/8 1999 July/August.

[2] "**Inverse Filter of Sound Reproduction Systems Using Regularization**", Hironori Tokuno, Ole Kirkeby, Philip Nelson and Hareo Hamada, IEICE Trans. Fundamentals, VolE80-A, No5 May 1997.

[3] "**Inverse Filtering Design Using a Minimal-Phase Target Function from Regularization**". Scott G. Norcross, Martin Bouchard, and Gilbert A. Soulodre. Convention Paper 6929 Presented at the 121st Convention 2006 October 5–8 San Francisco, CA, USA

[4] "**Fast Deconvolution of Multi-Channel Systems Using Regularisation**" by Ole Kirkeby, Philip A. Nelson, Hareo Hamada*, and Felipe Orduna-Bustamante, Institute of Sound & Vibration Research, Southampton University, SO17 1BJ, UK *Department of Electrical and Communications Engineering, Tokyo Denki University, Tokyo 101, Japan

[5] "**Advancements in impulse response measurements by sine sweeps**", Angelo Farina, Convention Paper Presented at the 122nd Convention 2007 May 5–8 Vienna, Austria

[6] "**Practical Limits for Room Equalization**", Louis D. Fielder, Dolby Laboratories Inc. Convention Paper 5481 Presented at the 111th Convention 2001 September 21–24 New York, NY, USA

[7] "**Analysis of Traditional and Reverberation-Reducing Methods of Room Equalization**", Louis D. Fielder, JAES Vol 51, No ½ 2003 January/February.

[8] **http://www.bodziosoftware.com.au/Pre_Post_Ringing_IR_And_Pulses.pdf**

[9] **http://www.bodziosoftware.com.au/Square_Wave.pdf.**