Experimental validation of loudspeaker equalization inside car cockpits

G. Cibelli[◇], A. Bellini [*], E. Ugolotti [◇], A. Farina[**], C. Morandi[*]

[*◇]ASK Industries S.p.A.

Via F.lli Cervi, 79, I-42100 Reggio Emilia - ITALY

email dodo@ee.unipr.it, UgolottiE@askgroup.it

[*] Dip. di Ing. dell'Inform. - Università di Parma

Parco Area delle Scienze, n.181A, I-43100 Parma - ITALY

email a_bellini@ee.unipr.it, c_morandi@ee.unipr.it

[**] Dip. di Ing. Industriale - Università di Parma

email farina@pcfarina.eng.unipr.it

Abstract

This paper deals with the definition of an automatic tool to develop digital equalization systems for car cockpits. The main purpose is to design a digital audio processor suitable for any car inside, whose programming parameters can be computed with a development tool, which relies only on on standard acoustic measurements performed in the target car cockpit. Experimental results of the defined tools are presented using as target a Fiat Lancia Delta.

1 Introduction

Sound reproduction within a car inside is a difficult task. Reverberation, reflection, echo, noise and vibration are some of the issues to account for. A first step in the direction of increasing sound comfort is that of equalizing the acoustic pressure response in the frequency domain. To accomplish this task the inversion of the measured Sound Pressure Level (SPL) should be performed [1], [2]. The main target of the equalization procedure is to increase sound comfort at the driver position, option which does not directly results in an homogeneous quality in the whole car cockpit.

The constaints to be accounted for are the quality of the input signal, for instance produced by a CD player, and the computational capability of the DSP to be used real-time as the hardware platform. Moreover the quality and the resolution of the equalization must be tailored to human earing system sensitivity. In fact it is uneffective to improve sounds in the range of frequencies which the human ears cannot sense, or to perform a too accurate correction. Therefore only listening test can be an objective evaluation procedure of the equalization system designed.

[*]This work was supported by ASK Industries and by EC under ESPRIT ESD project 28.583.

Figure 1: Hi-fi car system.

Another desired feature of the tool to be designed is its generality. The procedure should be independent of both target environment and the instrumentation used to perform acoustic measurements. The proposed procedure relies on automatic tools which elaborate the acoustic measurements and implement the filter structure and produce the suitable coefficients for the adopted hardware platform. In this paper the results presented are obtained for a Fiat Lancia Delta, using acoustic measurements performed with AURORA [3].

The following items will be detailed in the paper.

1. Choice of the suitable car audio system and analysis of the problems related with the measurements set-up.

2. Software tools for the elaboration of measurements data, and for the computation of the filter coefficients.

3. Design of the equalizer system, and its implementation on a commercial DSP.

4. Listening tests and conclusion.

2 Hi-fi car system model

Our target environment system is an audio system composed by a stereo audio source, an amplifier, a set of loudspeakers and the car cockpit, as sketched in fig. 1.

The audio source, the amplifier and the car cockpit can be considered linear with fine approximation. In the following the non-linear behavior of loudspeakers will be neglected. There-

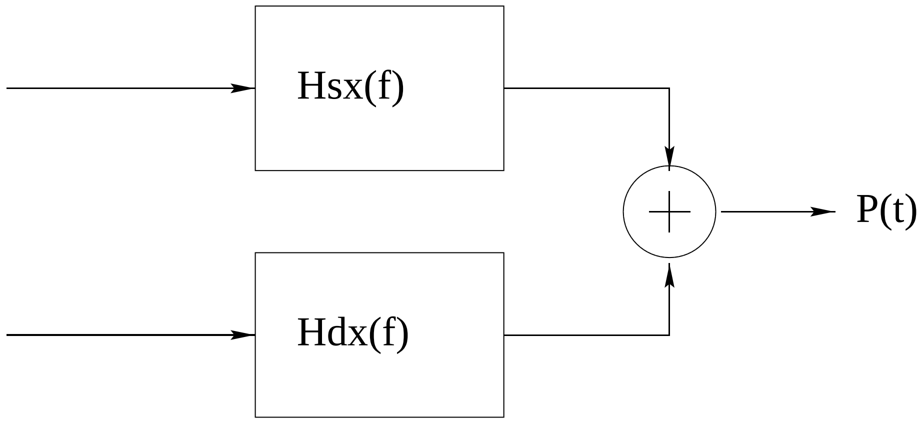

Figure 2: Linear model of the Hi-fi car system.

fore the whole sound reproduction chain can be considered a two-input one-output linear system, as sketched in fig. 2.

Under these assumption the equalization of two single channels is equivalent to the equalization of the whole system. Therefore we will measure separately the two channels, in order to characteriza their frequency response or their impulse response. Usually these kind of measurements are performed by means of MLS signals or sinuosoidal signals sweeps [3]. The two methods produce satisfactory results respecting a few points. The number of sinusoidal signals used must be high and increasing linearly with frequency. A sweep of sinusoidal signals increasing with frequency would be more consistent with human earing system, but though more effective to characterize low frequencies contents it would result in errors at higher frequencies. Therefore the variation of $H(f)$ between to adjacent frequency $f_i e f_{i+1}$ must be limited, thus since the phase response is of type $e^{j2\pi f T_d}$, the condition $2\pi f_i T_d - 2\pi f_{i+1} T_d < 2\pi$ must be respected, where T_d is the sampling period. If this fails the information retrieved from an "unwrapped" phase response are lost. If the digital method is used, the sampling frequency must of course be higher than $F_{Nyquist}$ and the length of the MLS sequence must be consistent. These constraints are often automatically forced by the adopted tool.

3 Elaboration of measurement data

As reported the Aurora tool was used to to perform acoustic measurements of the audio reproduction system of a Fiat Lancia Delta. In the following data regarding the left channel are reported, without losing generality. The impulse response was obtained through a microphone mounted on a dummy head positioned in the driving position, see fig. 3. The used microphone however is especially fitted for riverberating fields, and thus smooths heavily the high frequencies contents. The impulse response, whose duration is of about 50 ms, provides all the information concerning reflections, multiple path, ... i.e. all the events which affect the acoustic behavior of the system. The trasformation of the inverse of the impulse response is reported in amplitude and phase in figs. 4 and 5 and in fig. 6 as a function of time. In the following $H_{inv}(f)$ will indicate the trasformation of the inverse of the impulse response, and $h_{inv}(n)$ the corresponding time-domain function. A standard sampling frequency of $F_s = 44100Hz$ will be considered and as usual "Tap" will indicate the number of non vanishing impulse response terms. Their time duration can be computed as $L = \frac{Tap}{F_s}$. It must be pointed out that the

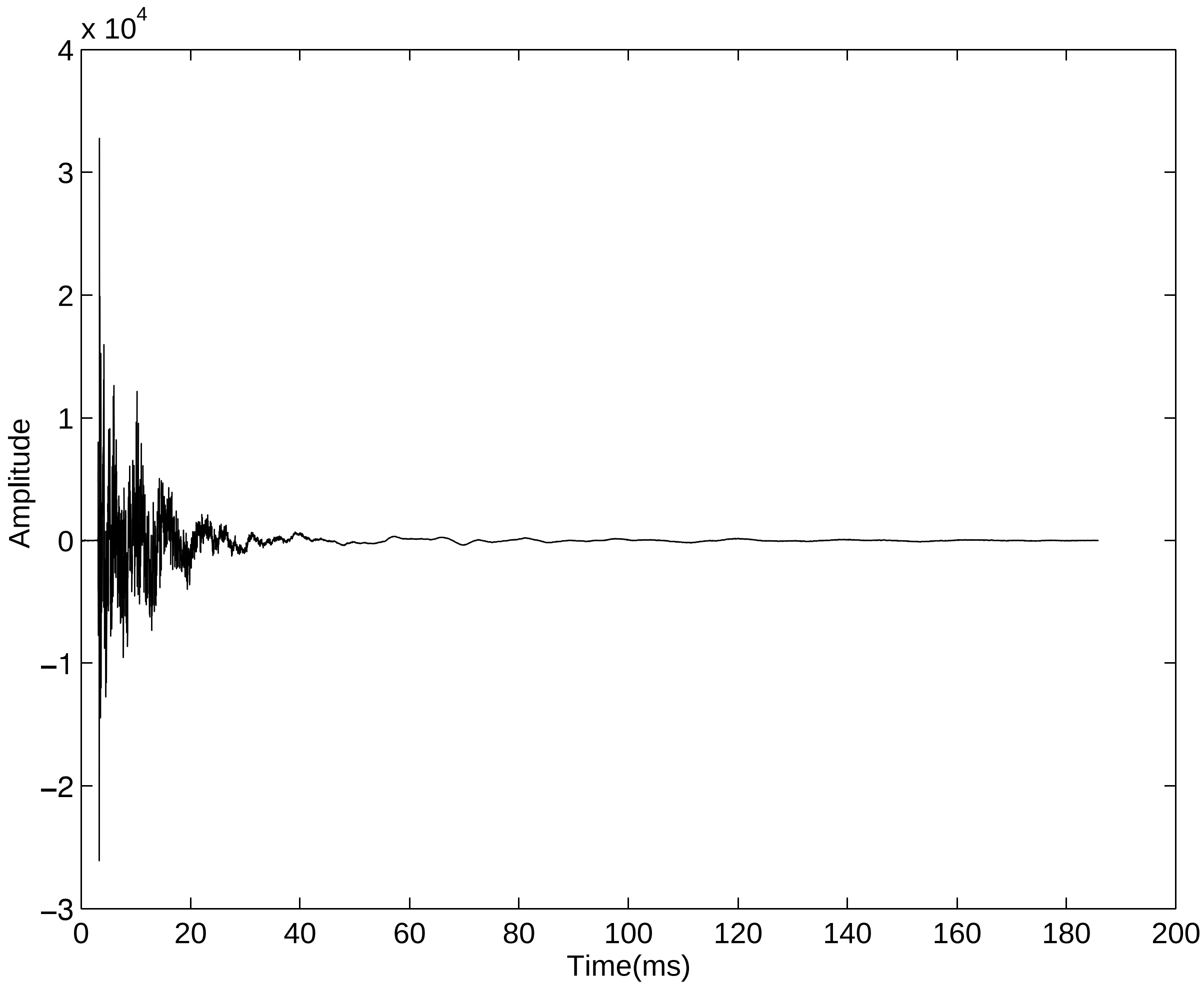

Figure 3: Impulse response of the left channel of a Fiat Lancia Delta (normalized data).

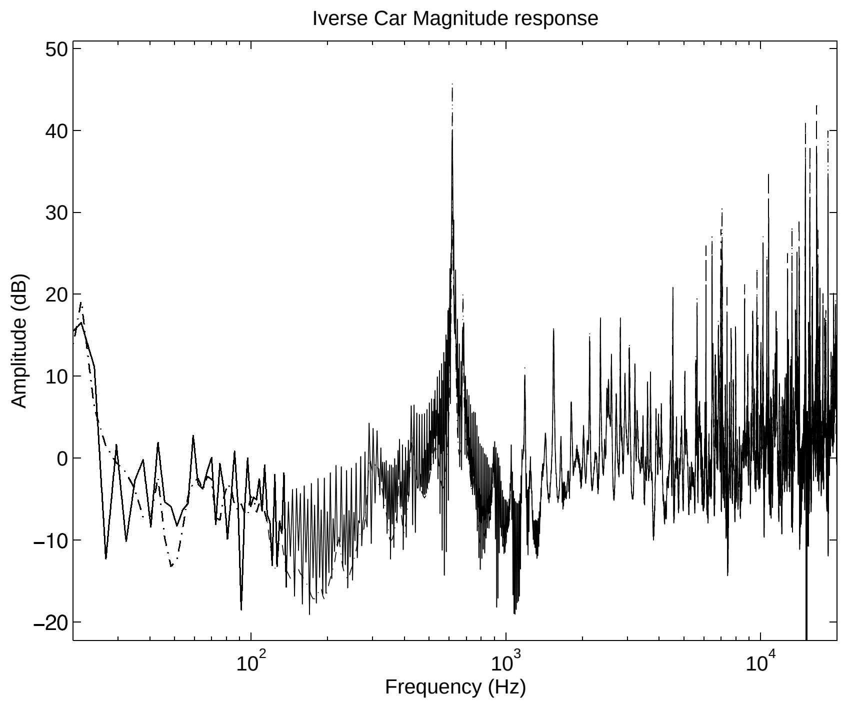

Figure 4: Amplitude of $H_{inv}(f)$

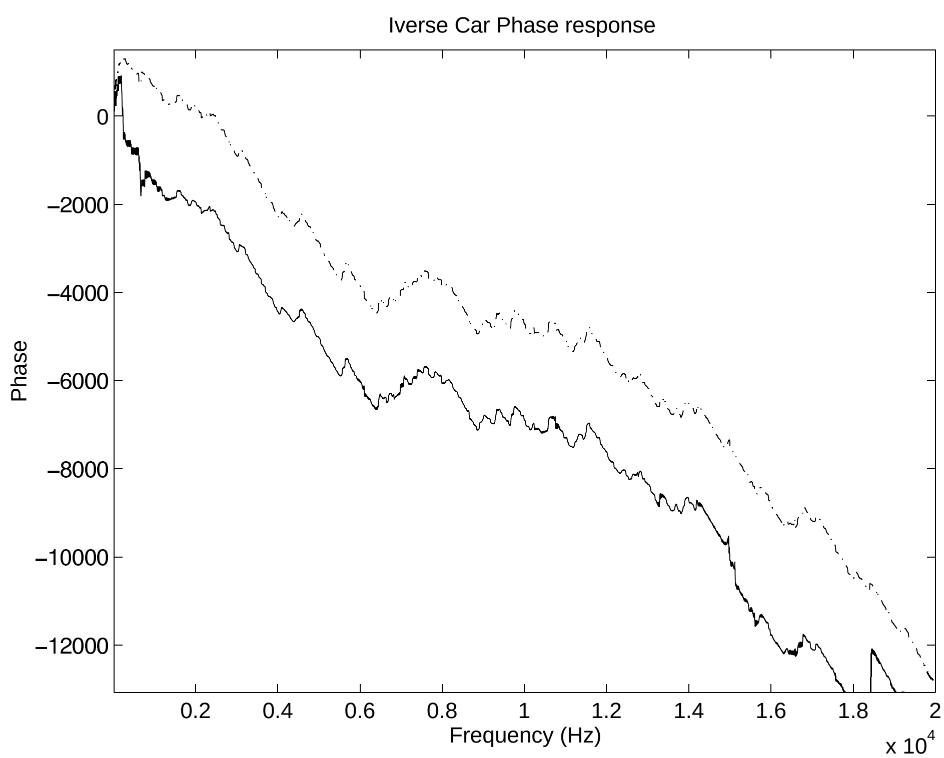

Figure 5: Phase of $H_{inv}(f)$

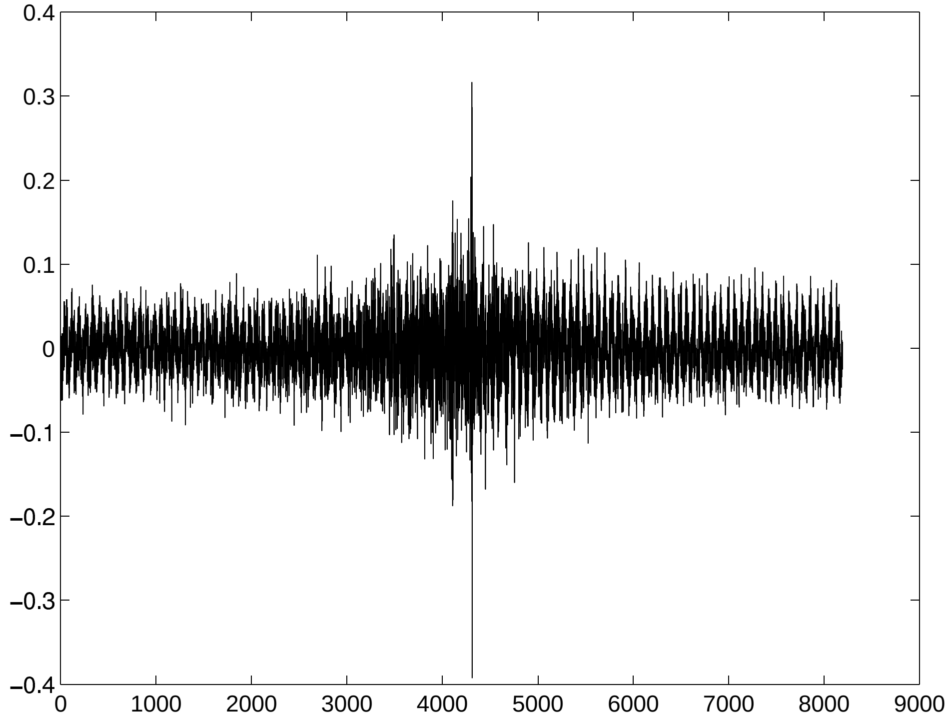

Figure 6: Impulse response $h_{inv}(n)$

duration of $h_{inv}(n)$ is of more than 8000 Taps, and thus it does not respect real-time constraints of a commercial DSP. In fact, since two channel must be operated in parallel a DSP capable of processing 16000 for each sample is required. In this work the AD SHARC DSP 21061 was used. It provides a computation power of around 40 Mips, with which 880 Taps can be computed for each sample at $Fs = 44100Hz$, respecting real-time constraints. Therefore this will be the our constraint considering all the program overheads. Therefore the $h_{inv}(n)$ is not directly implementable. Then a few approximation must be applied. Moreover it could be even dangerous to use an equalizing filter which employs the $H_{inv}(f)$ directly. In fact psycoacoustic considerations on human earing system suggest the following items

1. An ideal flat band reproduction system is not necessarily pleasant. In general an increased low frequency response is desirable, especially as far as car-audio systems are concerned, where running car noises must be masked.

2. Human earing system resolution is logarithmic and limited at about one third of octave.

¿From these it stems that the target frequency response function should be not exactly flat, but consistent with some sort of tailored equalization response. This should be confirmed by listening test and by expert considerations, see fig. 7. ¿From human earing resolution it stems that is uneffective to equalize all the frequency range with the same resolution. In fact at low frequencies the human earing system can detect as different sounds which are effectively near in frequency, while at high frequencies neither the most expert musicist cannot. Therefore the equalization of very narrow peaks or holes in the frequency response would be inefficient and even unpleasant. Therefore to reproduce human earing resolution the inverse harmonic response $H_{inv}(f)$ is averaged on thirds of octave starting from 20 Hz to 20000 Hz so as to get a stair wave. Nevertheless abrupt changes in the frequency response can results in artificial and non-natural

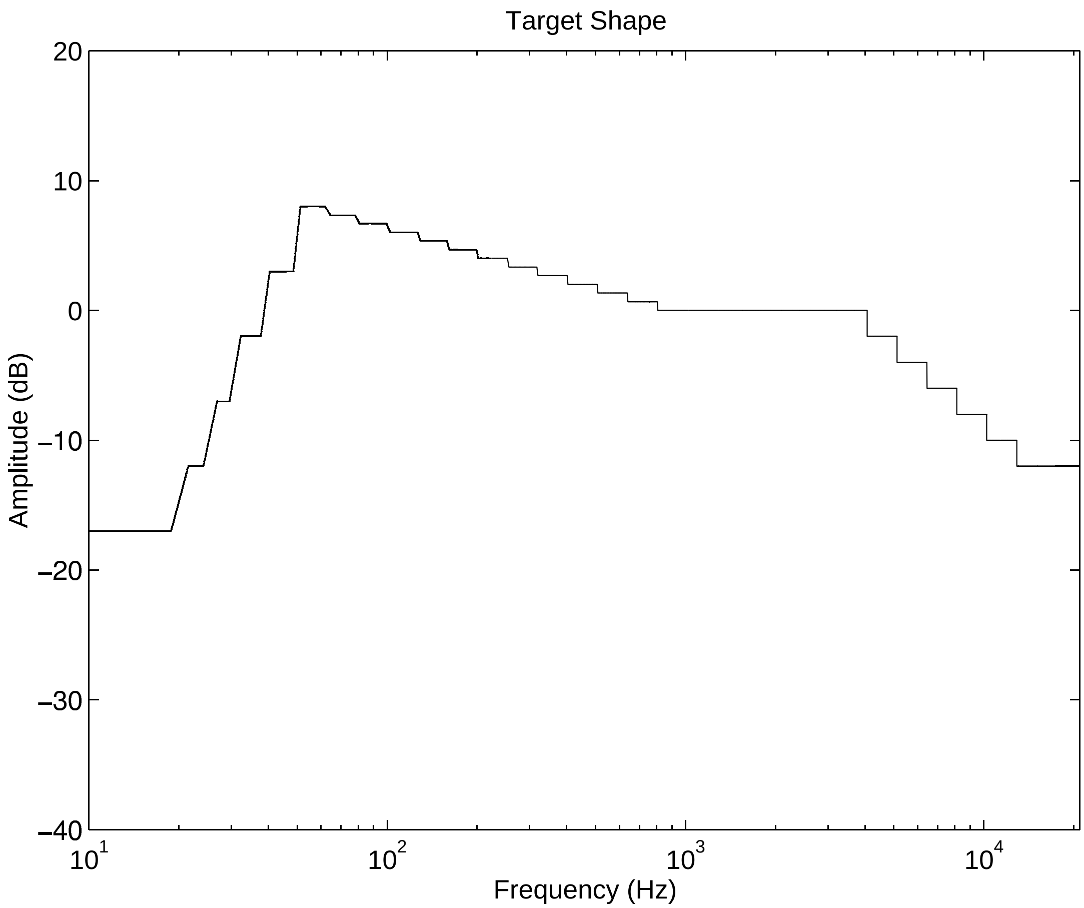
Target Shape
20
10
0
−10
−20
−30
−40
Amplitude (dB)
10^1
10^2
10^3
10^4
Frequency (Hz)

Figure 7:

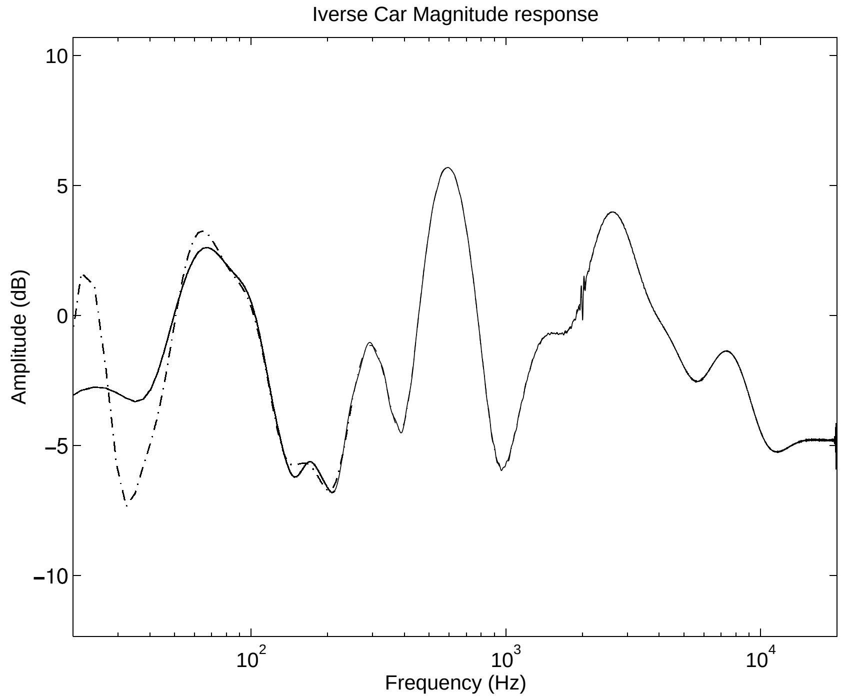

Figure 8: Harmonic reponse of the targe equalizer (Amplitude).

sounds. In order to smooth the frequency response the transformation $f_1 = 20log_2(f_0/10)$ was applied. This allows to eliminate higher harmonics contents through FFT transformation. The further application of the inverse of the relationship above defined reports to the “normal” frequency range. For the same purpose a median filter was applied [4]. These operations result in a new frequency response with reduced complexity, while the perceptual contents is almost unchanged.

Moreover an ideal equalizing system should have a dynamic range of about $30 \div 40dB$. This means that the insertion of such a device before the amplifier could overcome the maximum allowable dynamics of the others devices, with a high possibility of arising distortion phenomena. To avoid this drawbacks it is necessary to limit the operating range for istance to $\pm 6dB$, so that the correction is pleasant, smooth and natural. The shape of the filter which results after all the discussed elaborations is reported in figs. 8 and 9, and the number of its Taps is of about 1800.

4 Architecture of the digital equalizer system

In the previous section measurements data have been elaborated in order to obtain an equalizer system, whose duration in Taps is reducted to 1800. Nevertheless the available hardware forces the maximum number of taps to 800. To accomplish this furher reduction other techniques must be investigated. As an option fast hardware architecture could be employed to reduce the computational cost of the filter implementation, for instance overlap-save convolution techniques, but this is impractical because of the excessive amount of memory required with respect to the commonly available. The best trade-off was the adoption of a multi-rate architecture. In fact in a digital system no theoretical constraint exists, so that a single sampling frequency must be used. Moreover the selective averaging operation on $H_{inv}(f)$ suggests that the low-pass and

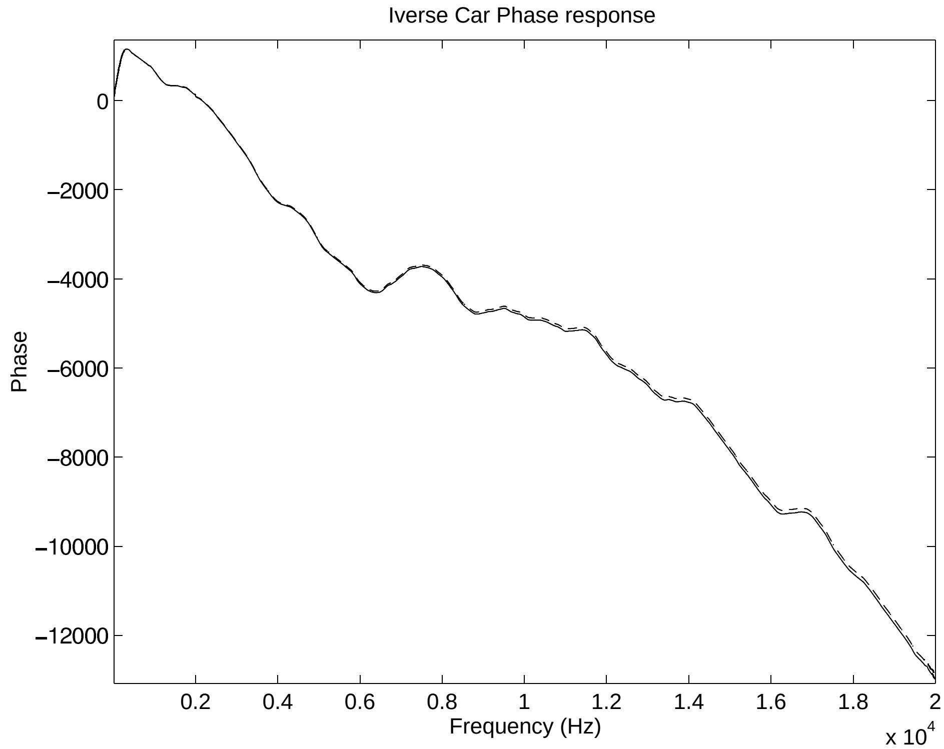

Figure 9: Harmonic reponse of the targe equalizer (Phase).

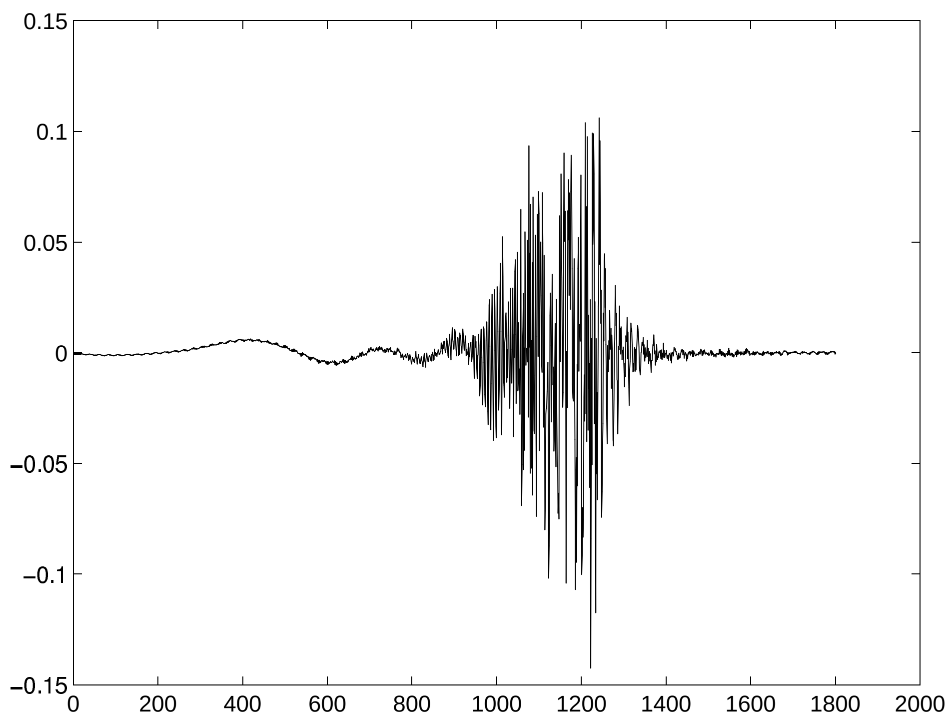

Figure 10: Impulse response of the target equalizer.

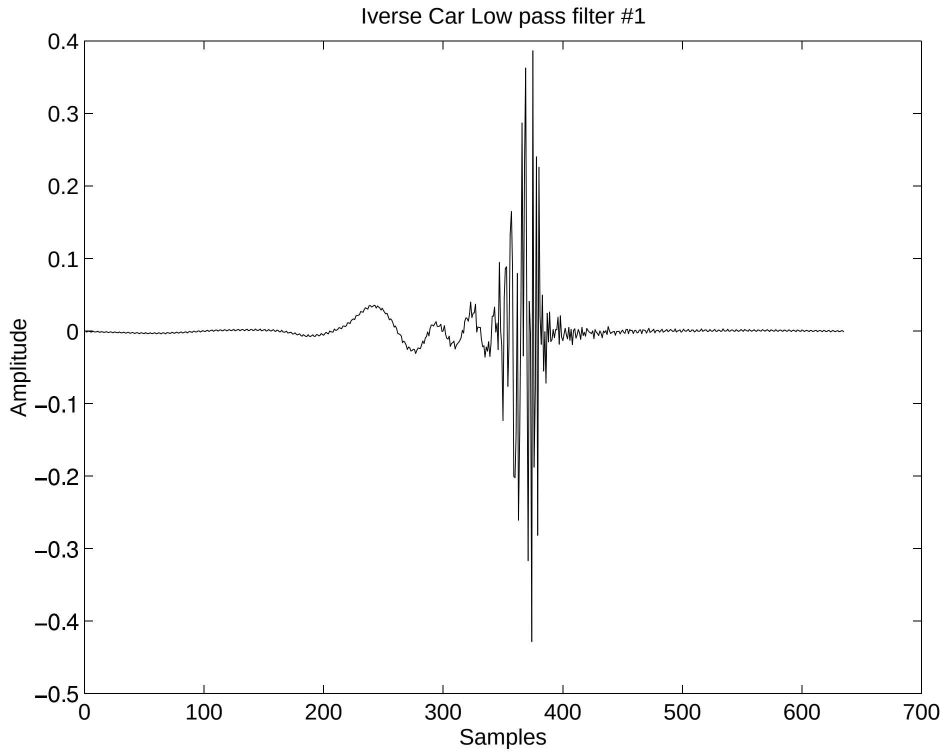

Figure 11: Impulse response of $h_{ilp}(n)$

the high-pass parts of the signal should be operated in different ways, in order to be consistent with human earing system conformation. Therefore a generic audio signal whose frequency band is in the range $[20Hz, 20KHz]$ can be split in two components, the former called $x_{lp}(t)$ whose frequency band is $20Hz\frac{F_s}{2N_{dec}}$, and the latter, partly overlapped, is called $x_{hp}(t)$ whose frequency band is $\frac{0.8F_s}{2N_{dec}}20000Hz$. If we assume that it is possible to sample $x_{lp}(t)$ at $\frac{F_s}{N_{dec}}$, and $x_{hp}(t)$ at F_s we obtain the following advantages

- The complexity of the filter which operates on $x_{hp}(n)$ is much lower than $H_{inv}(f)$, since its frequency range is much lower.

- The signal $x_{lp}(n)$ is operated at a sampling frequency N_{dec} times lower than F_s, and thus with a highly reduced cost.

The filters so obtained are defined $h_{ilp}(n)$ and $h_{ihp}(n)$, and their length is of $N_l = 436$ and $N_h = 451$ Taps respectively. Moreover the sampling rate of the low-pass part is $\frac{F_s}{N_{dec}}$, and thus the effective cost of the processing is of $\frac{N_l}{N_{dec}}$ Tap for each sample. The global cost is therefore $\frac{N_l}{N_{dec}} + N_h$, which in our examples is about 660 Tap for each sample. In figs. 12, 13, 11, 15, 16 and 14 are reported the harmonic responses and the corresponding impulse responses of the above defined filters.

All the charts presented are defined unless of a constant phase, which results in a delay constant with frequency. However in order to obtain consistent processed data it is necessary to identify exactly the delay introduced by each processing stage. To this aim since all the FIR filters with phase arbitraria are synthesized in order to obtain the maximum energy of the impulse response within the desired window. This delay is therefore automatically computed by the algorithm of FIR synthesis. In the following the delays introduced by the two equalizing

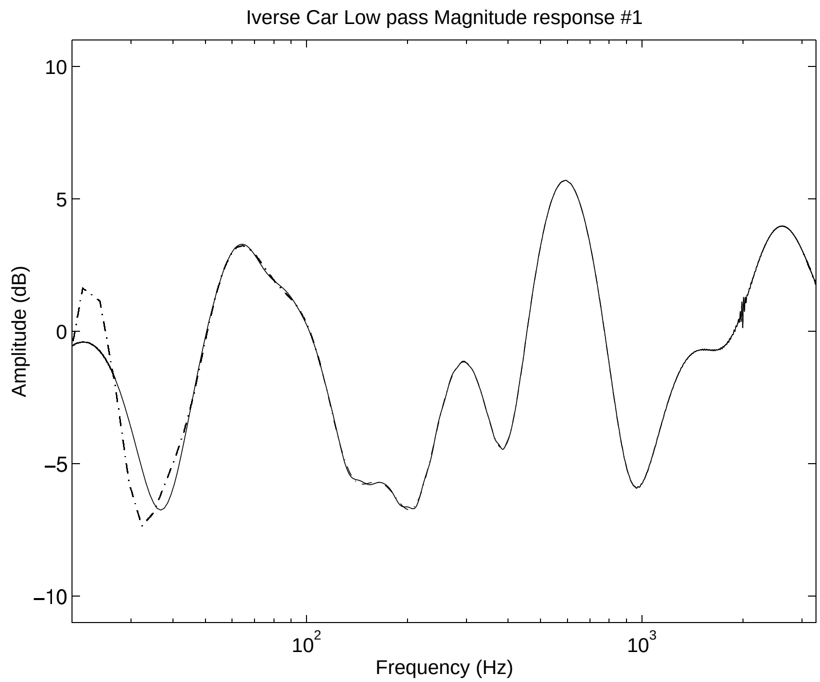

Figure 12: Amplitude of the system transfer function of $h_{ilp}(n)$

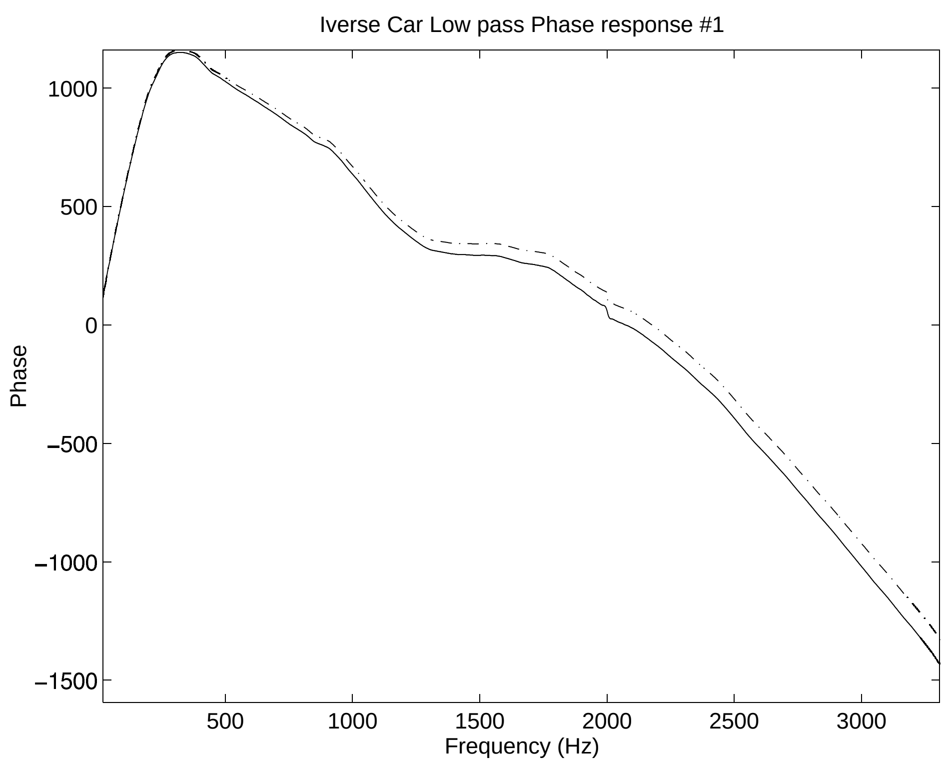

Figure 13: Phase of the system transfer function of $h_{ilp}(n)$

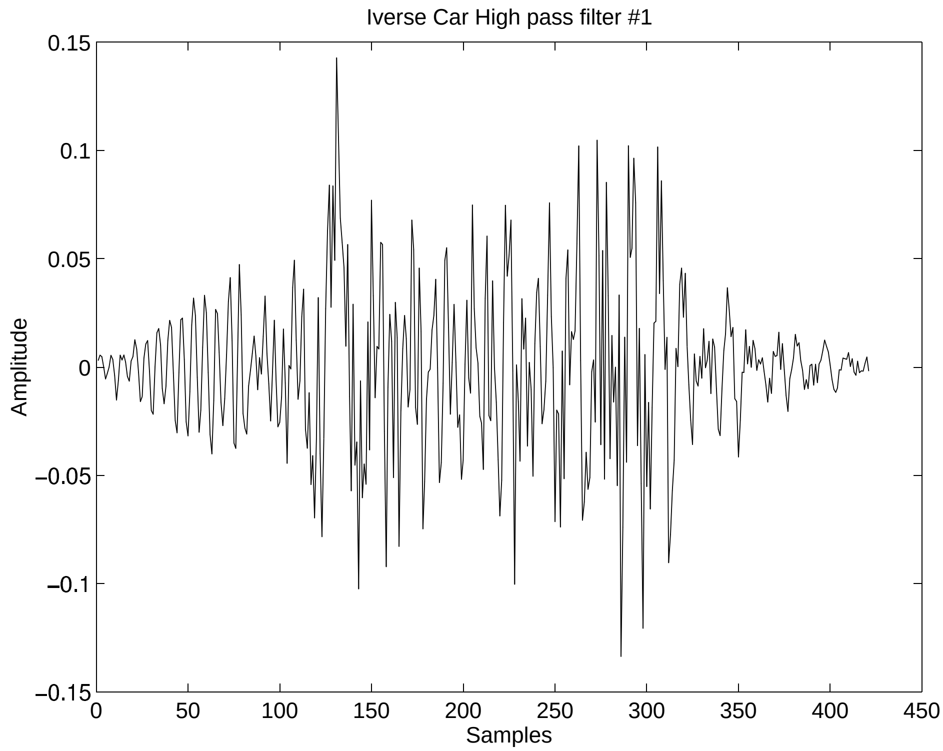

Figure 14: Impulse response of $h_{ihp}(n)$

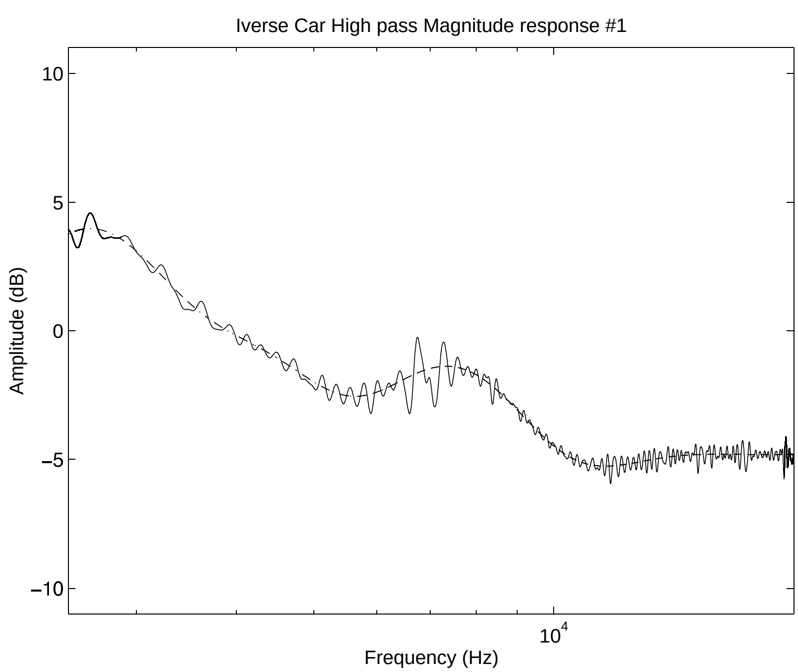

Figure 15: Amplitude of the system transfer function of $h_{ihp}(n)$

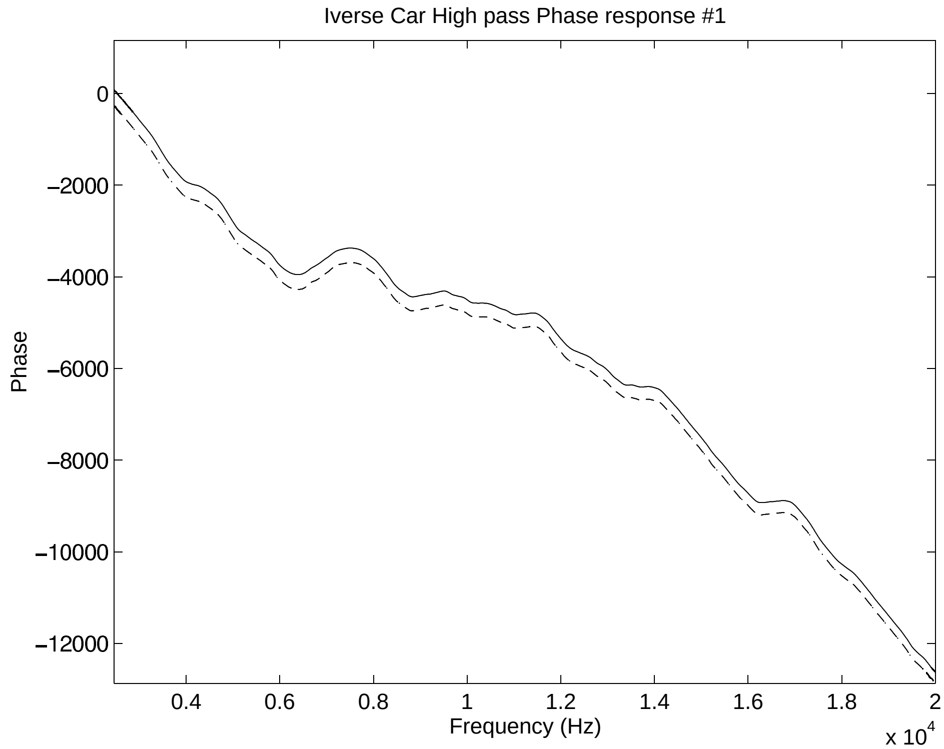

Figure 16: Phase of the system transfer function of $h_{ihp}(n)$

filters h_{ilp} and h_{ihp} are defined as $Ndelay_{lp}$ and $Ndelay_{hp}$ respectively The choice adopted was that of FIR filters with linear phase, therefore the delay introduced by each of them is equal to $(N-1)/2$ Tap, where N the filter length.

The proposed architecture features only amplitude equalization. This option is a fine trade-off between computational cost and subjective results. In fact, so doing we can achieve a fine equalization especially at low frequencies and the multi-rate architecture allows to implement the filters on a single AD 21061 EZ-LITE for both stereo channels. Moreover the correction of the phase would be still possible, simply by changing the tool of Taps synthesis, but proves to produce a slight increase of sound quality due to the limited delay resolution capability of human earing system.

4.1 Multi-rate architecture

As described in the previous section it is necessary to split the signal in the two components $x_{lp}(t)$, and $x_{hp}(t)$, with band respectively $20\frac{F_s}{2N_{dec}}Hz$ and $\frac{0.8F_s}{2N_{dec}}20000Hz$.

This can be achieved through a low-pass filter (LPF) with a cut-off frequency equal to $\frac{0.8F_s}{2N_{dec}}$ and with its complementary high-pass filter (HPF) which can be implemented at no cost as in fig. 17. In order to achieve to phased signal it is necessary to introduce a delay in the high-pass flow, and its delay is equal to $\frac{N_{LPF}-1}{2}$, i.e. the delay of the LPF filter.

Then the signal $x_{lpF_s}(n)$ must be decimated with a factor N_{dec}, i.e. the sequence $x_{lp}(n)$ is built, so that $x_{lp}(n) = x_{lpF_s}(N_{dec}n)$. The signal $x_{lpF_s}(n)$ is affected by time-aliasing only if in it frequency spectrum appear components beyond $\frac{F_s}{2N_{dec}}$. In order to avoid aliasing is required that the LPF filter cuts more than $80dB$ beyond Nyquist frequency of the decimated signal.

The reconstrution phase of the signal is composed by an interpolation of both low-pass and

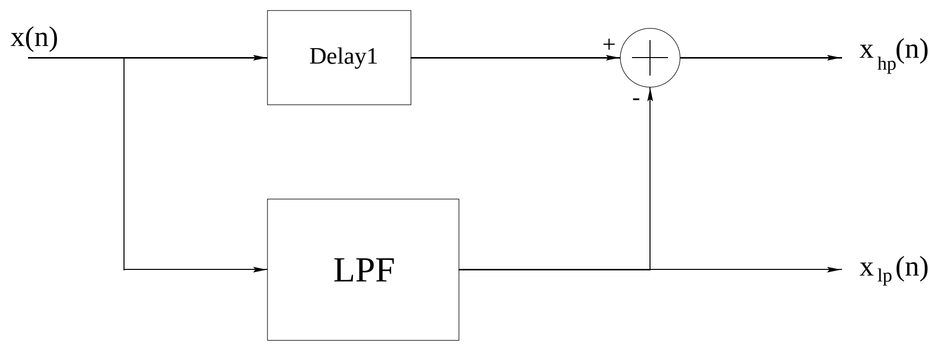

Figure 17: Block diagram of an digital crossover system.

high-pass signal elaborated by the $h_x(n)$ filters, called $y_{lp}(n)$ and $y_{hp}(n)$ respectively. If we define a sequence sampled at F_s so that

$$z_{lp}(n) = \begin{cases} y_{lp}(m) & \text{when } n = mN_dec \\ 0 & \text{otherwise} \end{cases}$$

The frequency spectrum of $z_{lp}(n)$ is given by

$$\begin{aligned} Z_{lp}(e^{j\omega}) &= \sum_{n=-\infty}^{\infty} z_{lp}(n)e^{-j\omega n} = \sum_{m=-\infty}^{\infty} y_{lp}(m)z_{lp}(n)e^{-j\omega N_{dec}m} = \\ &= Y_{lp}(e^{j\omega N_dec}) \end{aligned}$$

which can be interpreted as the periodic repetition of $Y_{lp}(e^{j\omega})$ every $\frac{F_s}{N_dec}$.

In order to mix the two signal it is sufficient to apply a low-pass filter and compute the suitable delays for each path. This low-pass filter is called LPF_2 and must comply with the following constraints

- It must be a flat filter in the band of $y_{lp}(n)$, i.e. in that frequency range where LPF fase less than $20dB$;
- The fading beyond $\frac{F_s}{2N_dec}$ must be higher than $80dB$.
- The gain of LPF_2 must be N_{dec} in order to the interpolation.

Since the $z_lp(n)$ sequence is significative only each N_{dec} samples. Therefore the computation of LPF_2 can be simplified accounting only for significative terms. Namely

$$z_{lp}(n) \star LPF_2 = \begin{cases} y_{lp}(n) \star LPF_2(nN_{dec}) & \text{where } n = mN_dec \\ y_{lp}(n) \star LPF_2(nN_{dec}+1) & \text{where } n = mN_{dec}+1 \\ \\ y_{lp}(n) \star LPF_2(nN_{dec}+N_{dec}-1) & \text{where } n = mN_{dec}+N_{dec}-1 \end{cases}$$

Where $\star$ stands for the convolution operator, and each convolution can be implemented with filters wholse length is $\frac{1}{N_{dec}}$ than expected. The architecture depicted in fig. 18 is thus obtained.

Figure 18: Block diagram of the multi-rate equalizer architecture.

The anti-aliasing filters consistent with the above defined constraints are implemented with two Kaiser windows, whose length is 181 for LPF and 341 for LPF_2. In figs. 19, 20, 21 , 22, 23, 24, 25 , 26, 27, 28 are reported their time and frequency behavior. Every filter is synthesized with even simmetry and with an odd number of taps. Therefore their phase is perfectly linear as a function of frequency, and the delay introduced by each filter is exactly $\frac{N_{tap}-1}{2}$.

5 Conclusion

The equalizer structure described in the previous sections was implemented in assembly on a 21061 EZ-LITE DSP board, and the software tools of Matlab [5] were used to synthesize the suitable filter taps. A Fiat Lancia Delta was characterized and used as a target with the insertion of the audio equalizer. The direct measurements of the only amplitude equalizer is reported in fig. 29 e 30 and its amplitude and phase characteristics are in fig. 31 and 32 respectively. We can stem from them that the phase is perfectly linear unless a slight distortion introduced by A/D and D/A converters.

Listening tests as well as distortion measures confirm that the equalization task can be achieved, and that a subjective improvement of sound quality can be obtained. The main impacts of equalization are the nice balancing of sounds coming from different instruments, and the increase of low frequencies, an option which is necessary to compensate running car noise, as resulted from experience. Moreover the virtual source of sound is shifted from the doors, where loudspeaker are, so that the listener perceives a frontal audio source, and a more natural and pleasant sound. A 70 taps delay introduced in the left channel is also very useful to improve the sound image for a listener in the driver position and does not affect the others. The equalization system seems robust with respect to the listener position and different measures in different

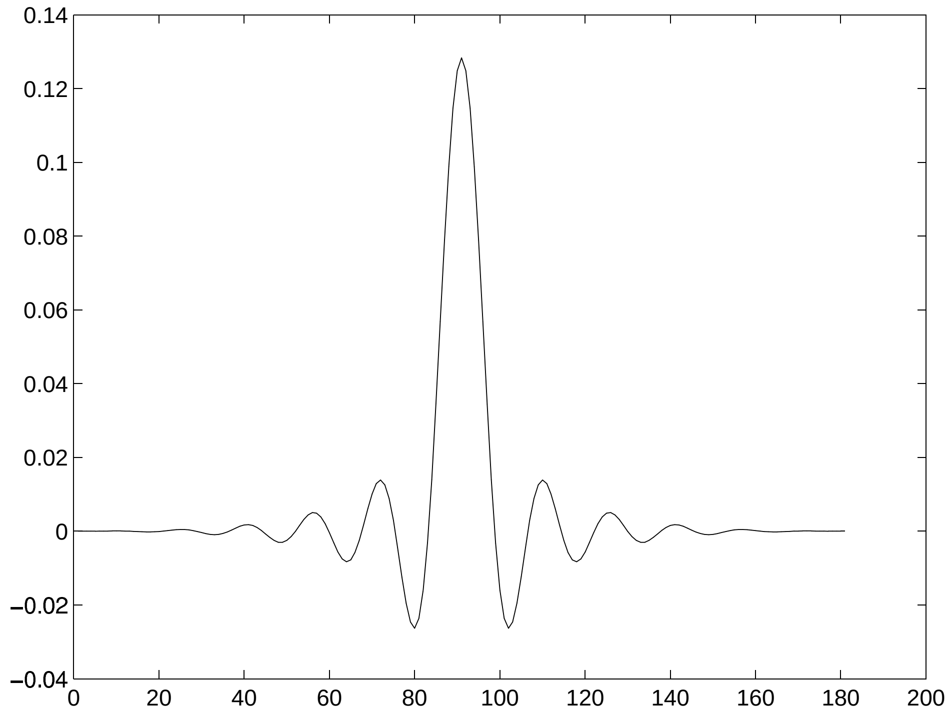

Figure 19: $lpf(n)$

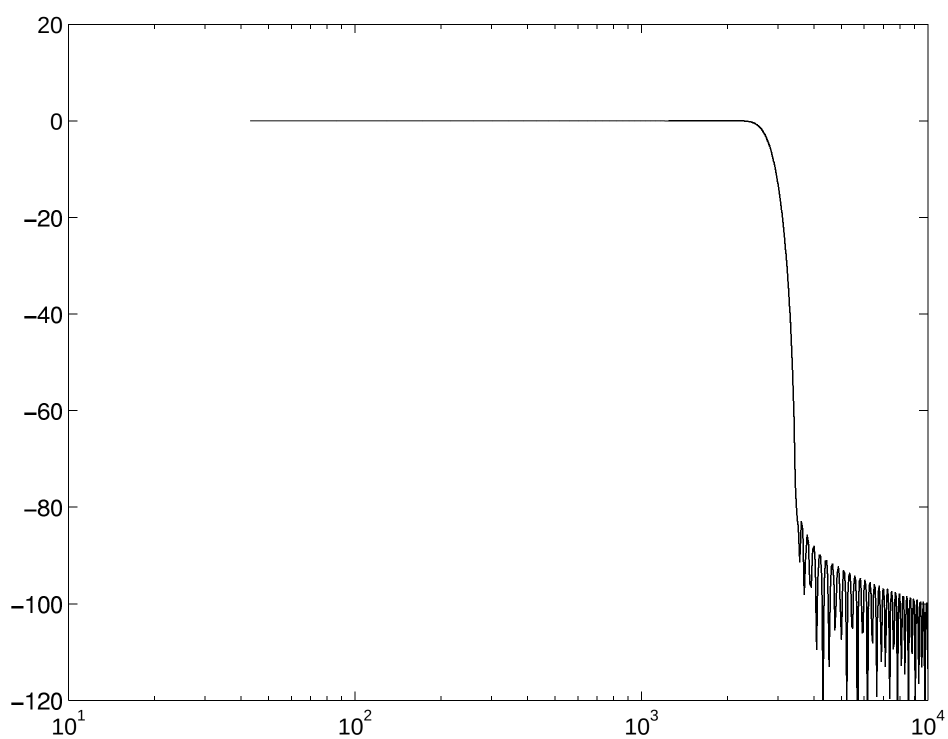

Figure 20: $LPF(f)$

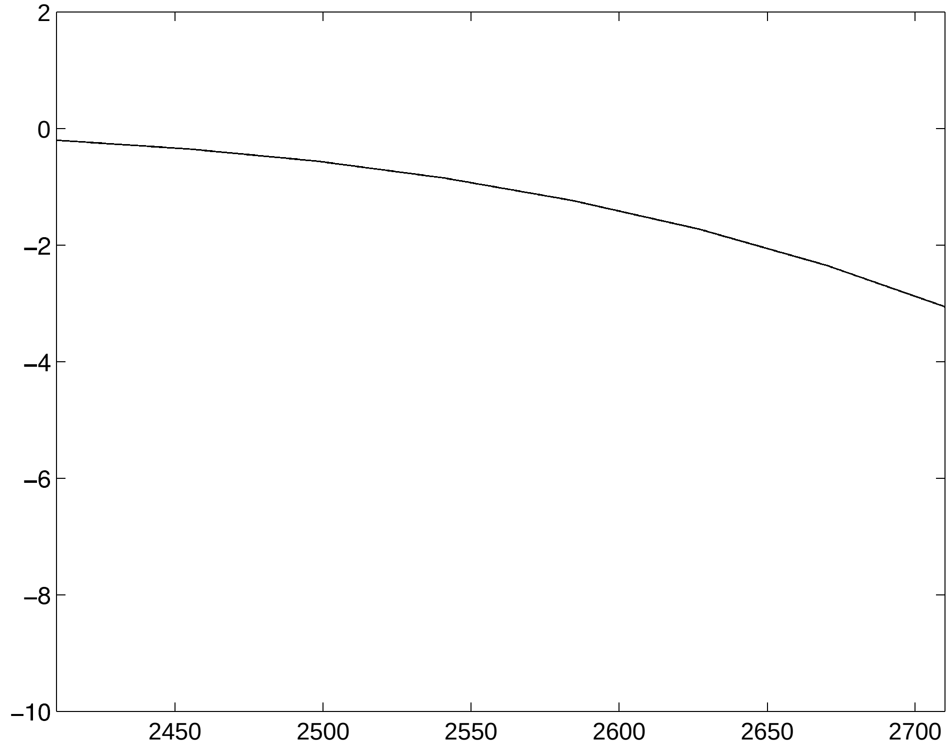

Figure 21: $LPF(f)$

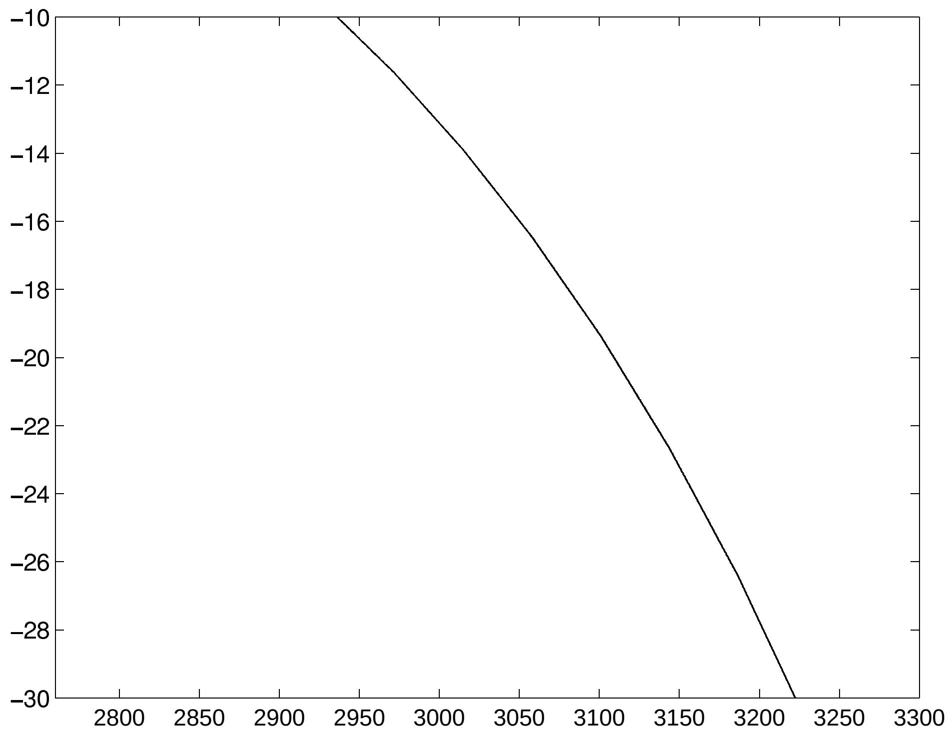

Figure 22: $LPF(f)$

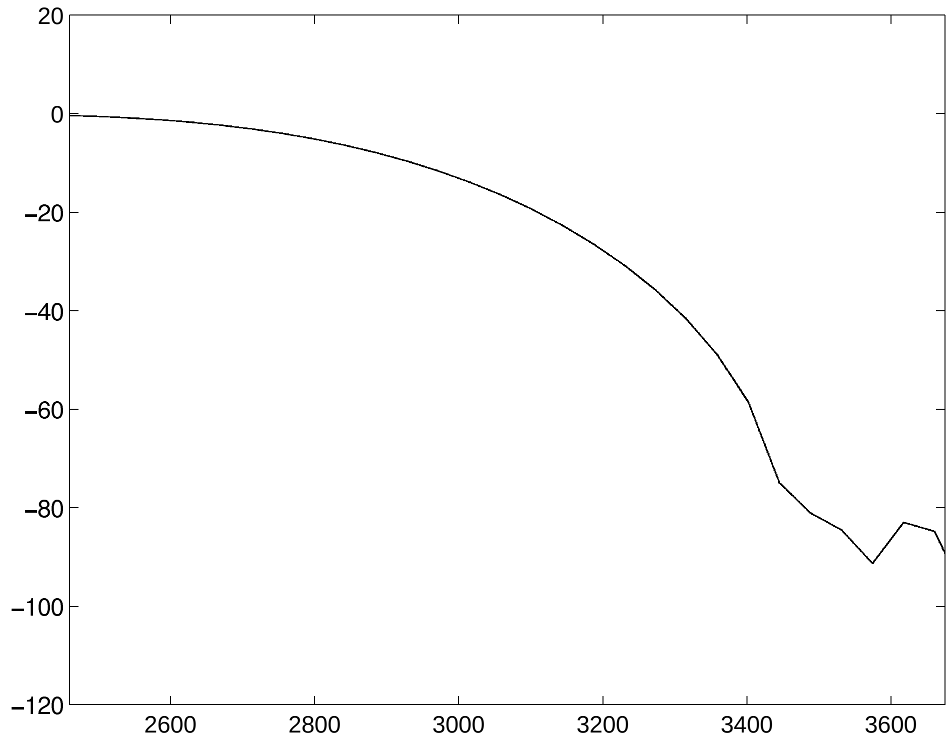

Figure 23: $LPF(f)$

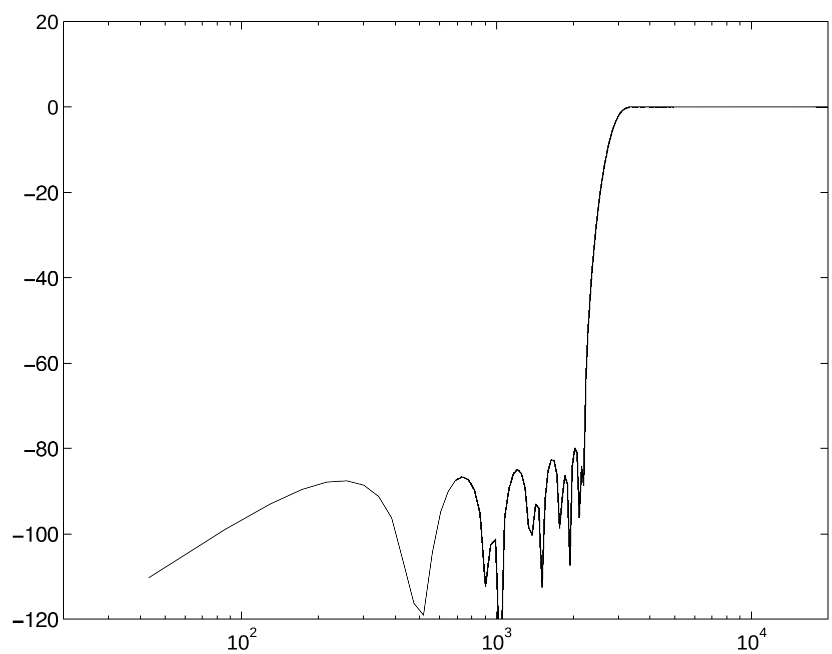

Figure 24: . $HPF(f)$

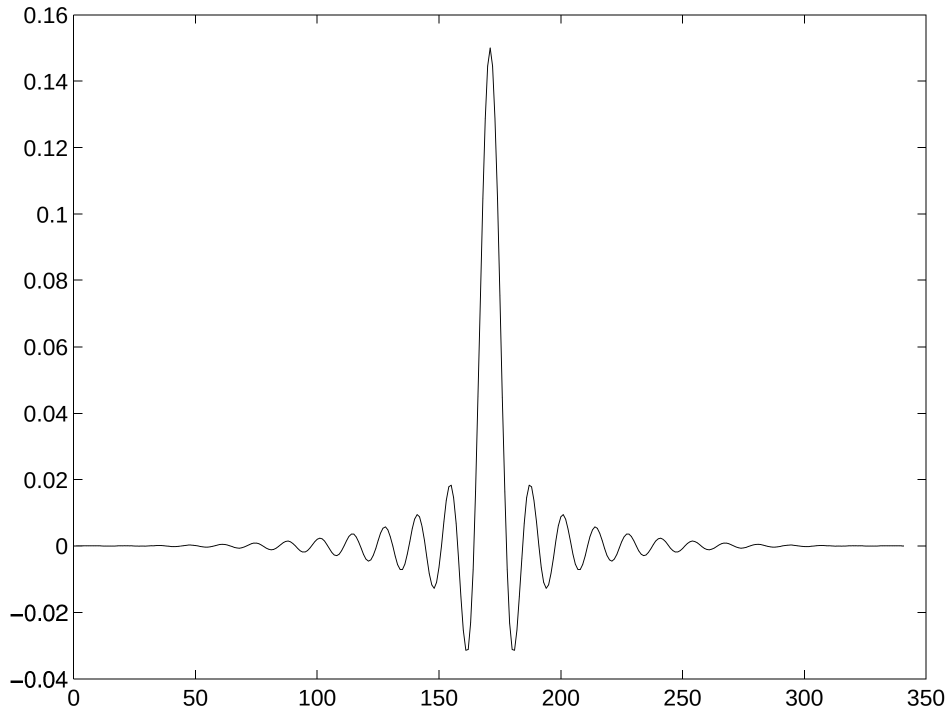

Figure 25: $lpf_2(n)$

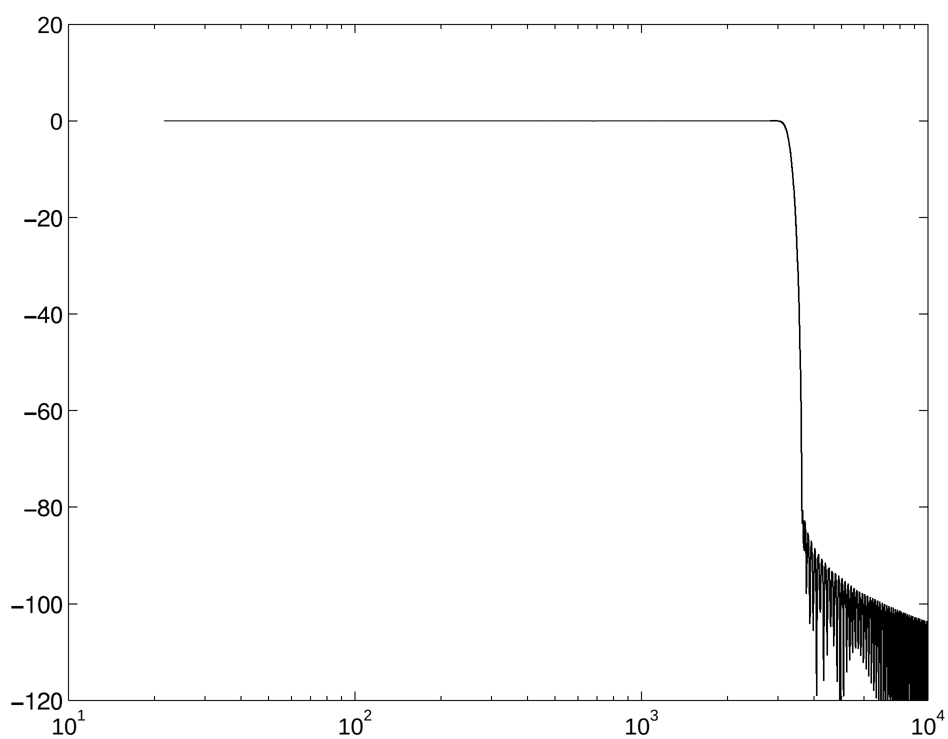

Figure 26: $LPF_2(f)$

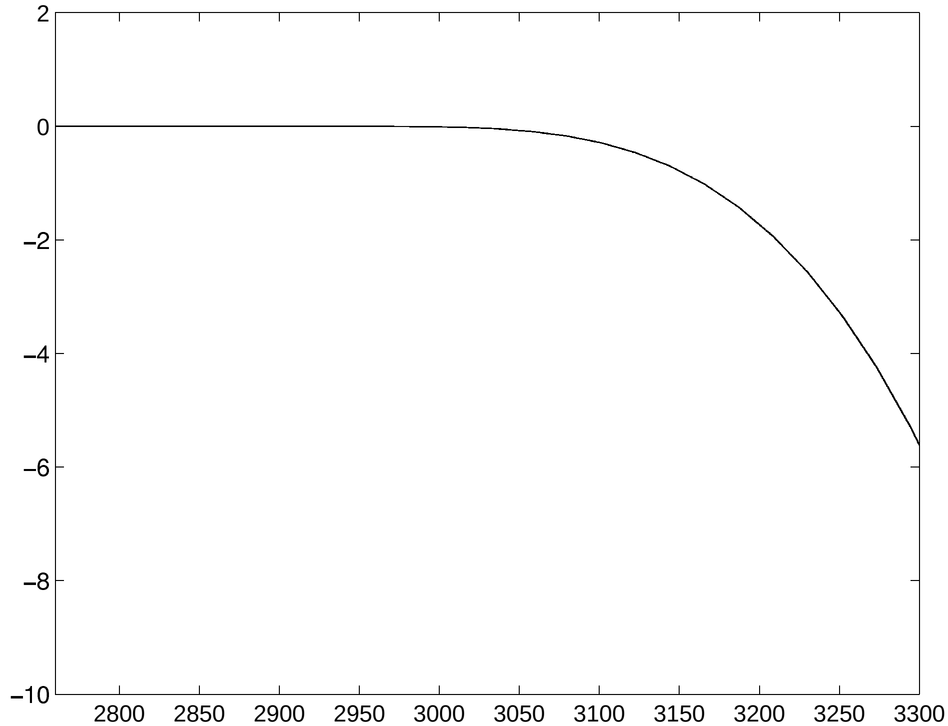

Figure 27: $LPF_2(f)$

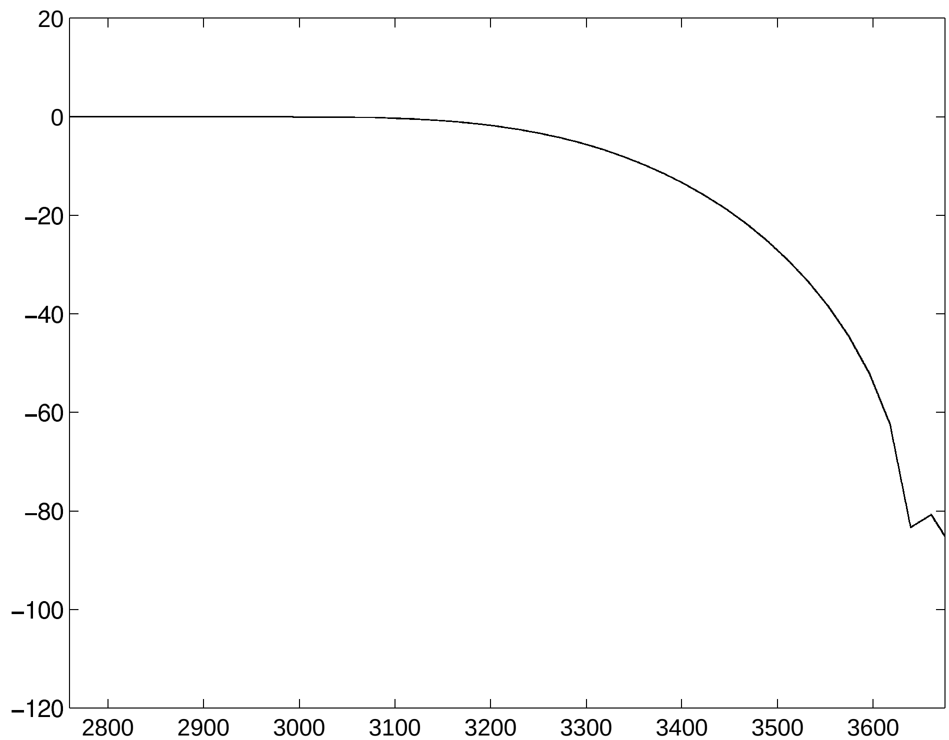

Figure 28: $LPF_2(f)$

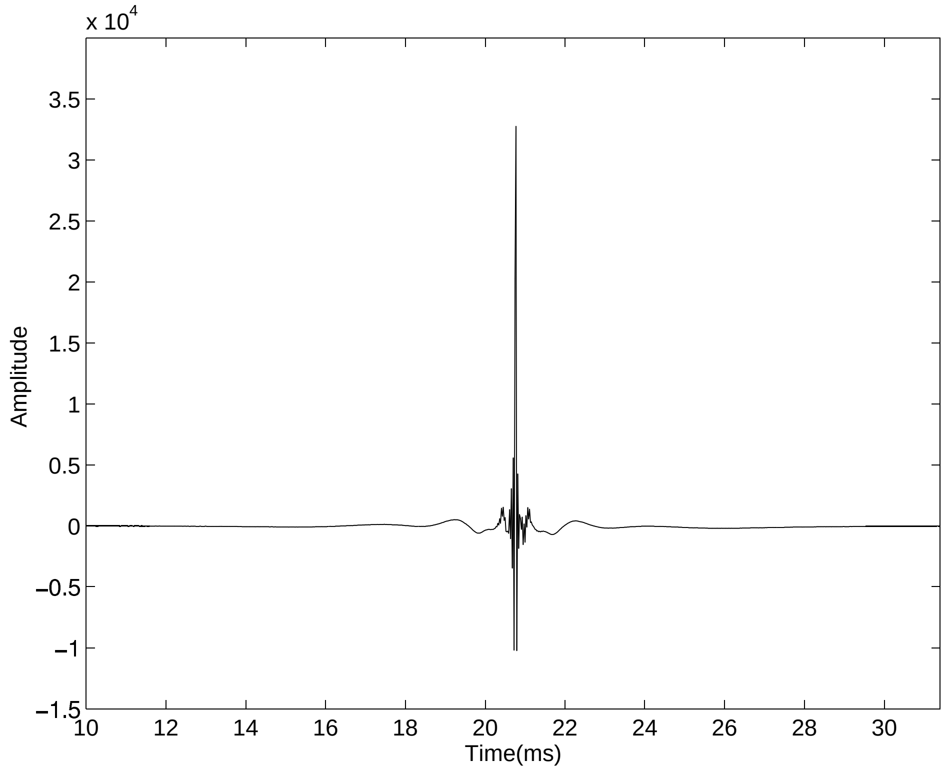

Figure 29: Impulse reponse of the multi-rate equalizer.

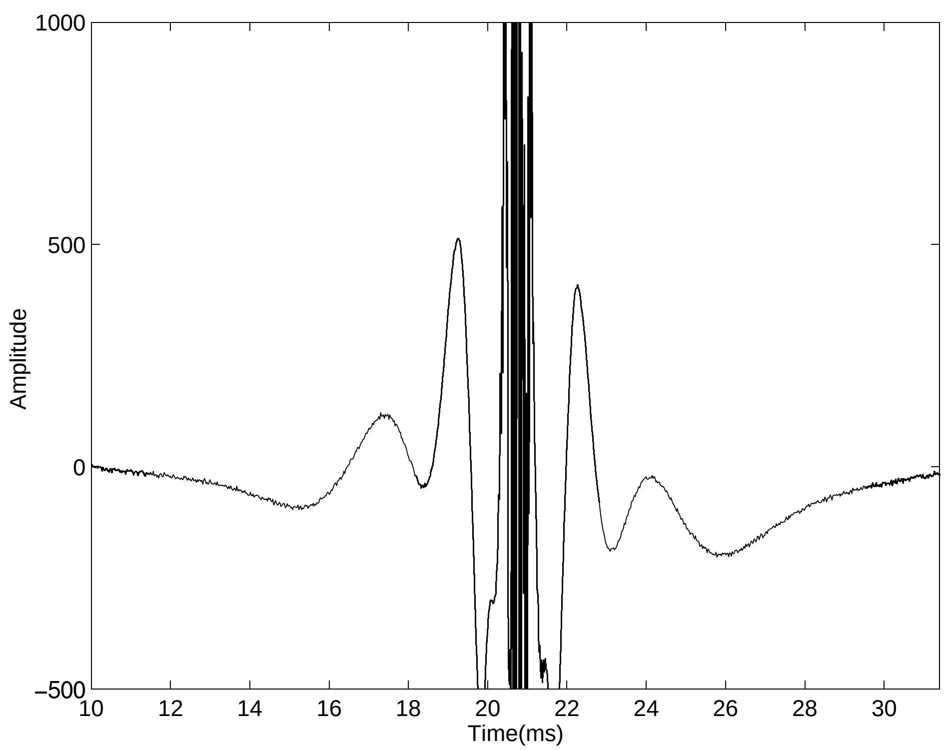

Figure 30: Impulse reponse of the multi-rate equalizer (detail).

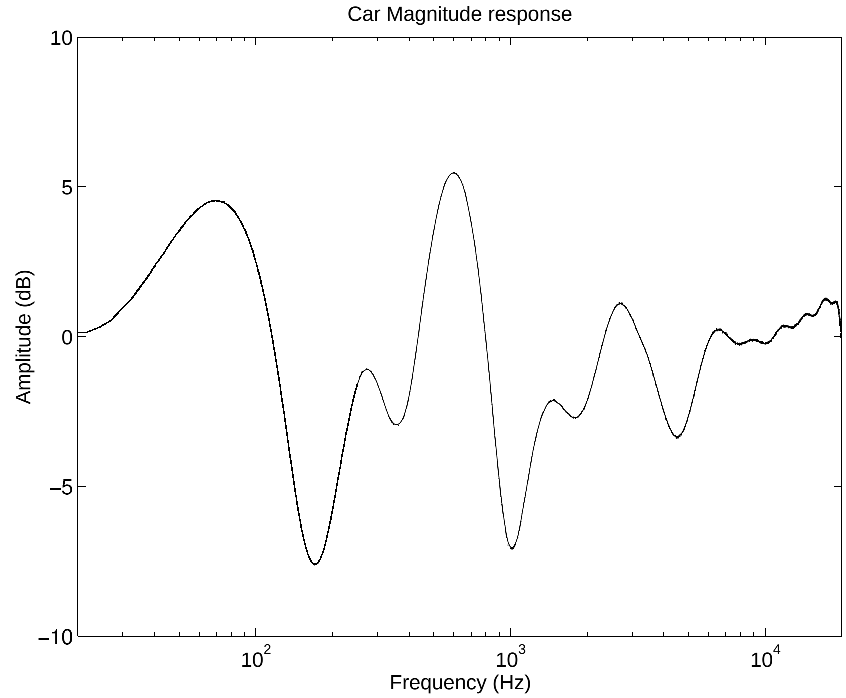

Figure 31: Frequency response of the multi-rate equalizer (Amplitude).

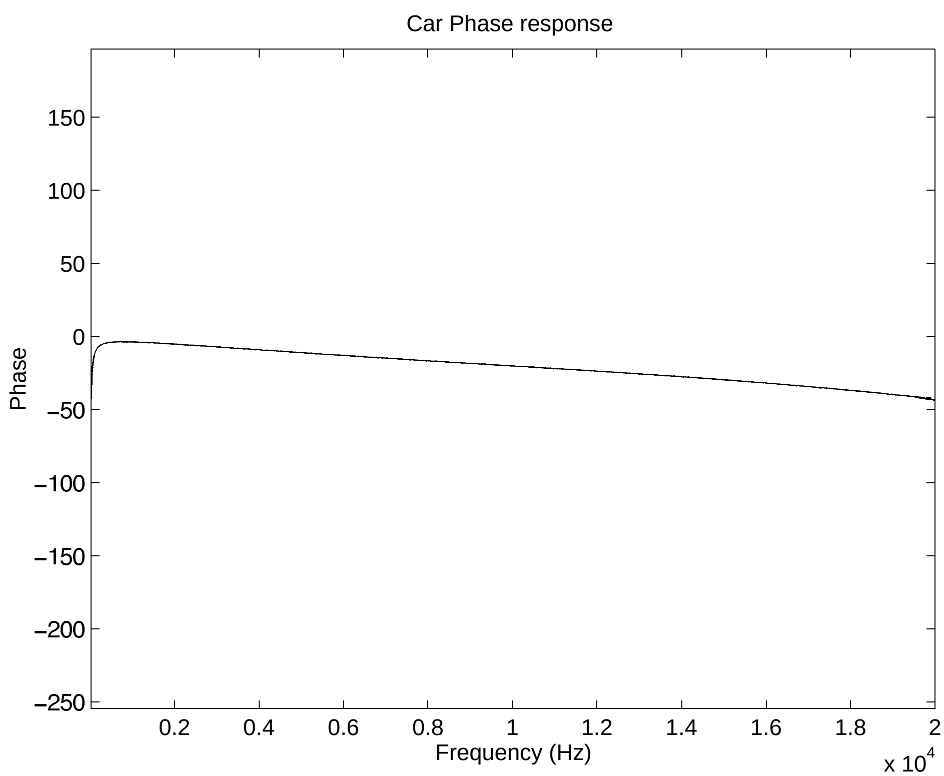

Figure 32: Frequency response of the multi-rate equalizer (Amplitude).

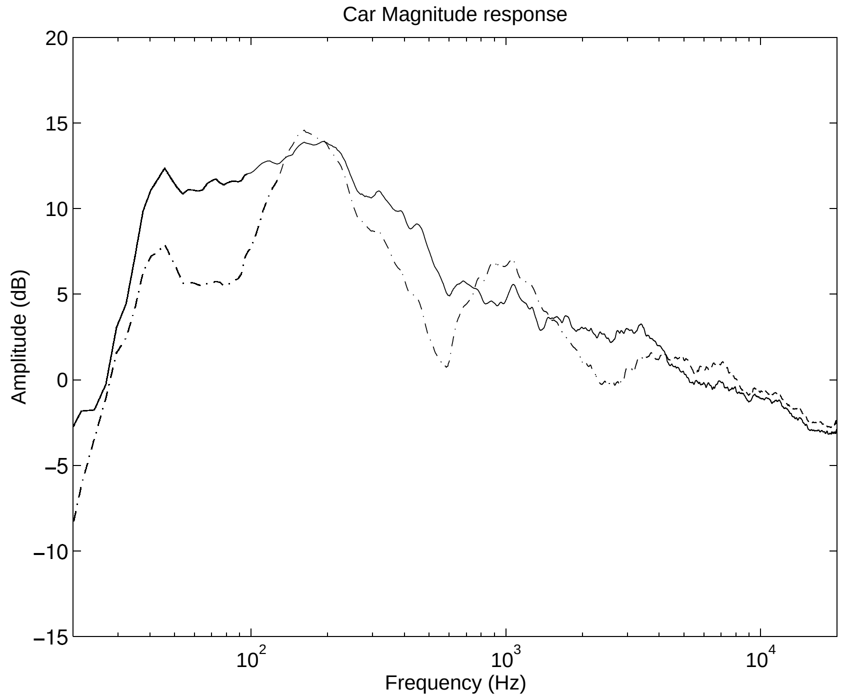

Figure 33: SPL of the left channel measured with (solid line) and without equalizer (dashed line).

points produce the same shape in amplitude but different phase. Due to the limitation of the used microphone the target shape was compensated adding $4 \div 6dB$ beyond $4KHz$. However we can still consider that riverberating microphones affect the effectiveness of the equalization. In conclusion we can observe that

- A digital system acting on third of octave can be effective within a car cockipt as an equalization system.
- The proposed system can be implemented on commercial low cost DSP based boards.
- The converters available introduced a critical noise and dynamic range reduction. A better option would be that of inserting the proposed equalizer structure in the car radio before amplifier and volume control with neglectable conversion noise (i.e. less than 80dB).
- A delay introduced in the left channel improves sound image for the driver.
- The simple phase equalization seems not necessary.
- The equalizer system proves to be robust and effective at each position of car passengers.
- A more effective system, considering phase correction too, should rely on binaural measurements, on more powerful DSPs, and should account for crosstalks effects (stereo dipole) and will object of future developments.

In figs. 34 amd 34 are reported the SPL curves measured within a standard Lancia Delta cockpit, after the proposed equalization.

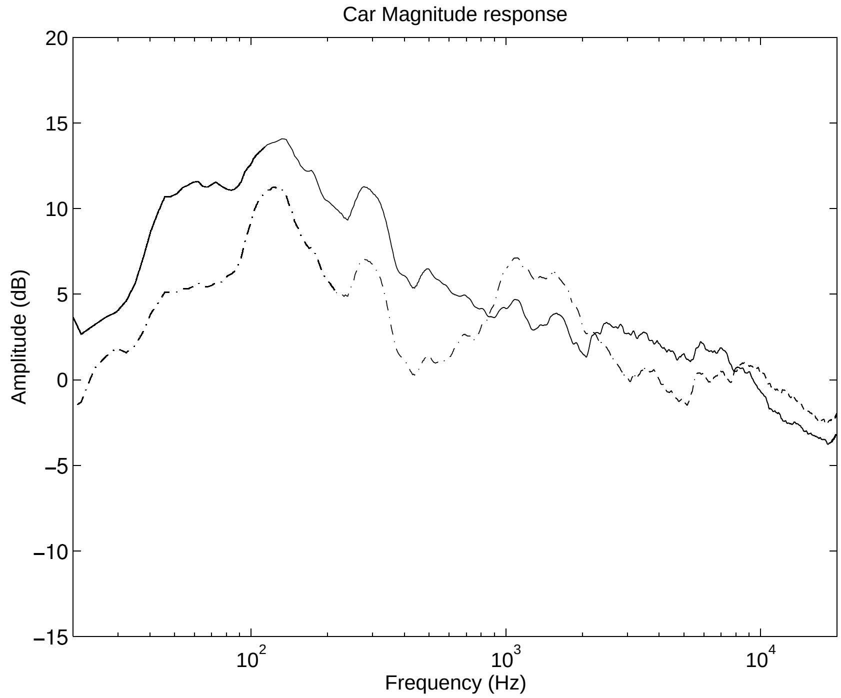

Figure 34: SPL of the right channel measured with (solid line) and without equalizer (dashed line).

Acknowledgments

The authors would like to thank Ing. Gobbi and the Technical Staff of ASK Industries for their encouragments and suggestions.

References

[1] W. Klippel, "Compensation for nonlinear distortion of horn loudspeakers by digital signal processing," *Journal of the Audio Engineering Society*, vol. 44, no. 11, pp. 964–972, November 1996.

[2] A. Bellini, G. Cibelli, E. Ugolotti, A. Farina, and C. Morandi, "Non-linear digital audio processor for dedicated loudspeaker systems," *IEEE Transactions On Consumer Electronics*, vol. 44, pp. 1024–1031, August 1998.

[3] A. Farina and F. Righini, "Software implementation of an mls analyzer, with tools for convolution, auralization and inverse filtering," in *Pre-prints of the 103rd AES Convention*, New York, September 1997, pp. 26–29.

[4] S.K.Mitra and J.F.Kaiser, *Handbook of digital signal processing*, McGraw-Hill, New–York, 1993.

[5] The Mathworks, Inc., *Signal Processing Toolbox*, 1998.