

Design and implementation of Psychoacoustics Equalizer for Infotainment

Andrea Azzali¹, Alberto Bellini¹, Angelo Farina², Emanuele Ugolotti³

¹ Dipartimento di Ingegneria dell'Informazione, Università di Parma

² Dipartimento di Ingegneria Industriale, Università di Parma

³ ASK Industries, Reggio Emilia

Abstract

This paper presents a novel dynamic equalization method for car inside. It deals with both design and implementation of a DSP based system for experimental validation of the proposed method. The major task of this method is the increase of sound reproduction quality. This is accomplished relying on the AQT analysis method, which allow an accurate characterization of the car acoustic response, including dynamics. A tool which performs automatic AQT computation was realized, then AQT parameters were used to synthesize an equalization filter. The filter was experimentally validated in a few commercial cars, using a DSP based board.

1. Introduction

Automotive engineering is an attractive field from any point of view, especially when dealing with accessories aimed at increasing comfort. As for Hi-Fi car audio system is being moving to a sound system, even if the specific car compartment composition makes sound reproduction a very hard task.

The particular structure of the cockpit introduces in the response phenomena that cannot be modeled with traditional methods. Some of this effects are noise floor, resonance, echoes that make our task difficult. They degrade the response and make this one so far from "Best Response Curve".

The main target of the equalization procedure is to increase sound comfort and make the response more closed with target curve at the driver position, option which does not directly results in an homogeneous quality in the whole car inside. However this correction produces a pleasant effect only if a perfect acoustic characterization of the environment is performed.

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]. In this way however we accomplish only a first characterization of response from a static point of view. Since both the musical signals and human hearing system are strongly dynamic, the latter approach is probably ineffective [3].

A first step in the direction of a dynamic and psychoacoustic equalization was made in [3], where an approach for the characterization of acoustic attacks and releases was presented. In this paper this work was extended as a synthesis tool and more parameters that allow a better characterization of car inside acoustics were introduced. This approach is defined AQT (Audio Quality test), keeping the same name of the analysis tool, built to obtain a dynamic characterization of a sound system. In this paper an automatic tool was developed in order to obtain AQT parameters quickly, and to use them to synthesize a nice equalization filter shape (section 5).

As last thing we have studied an equalization system that allow us to use this information to obtain an effective improvement of the hearing in car (section 6) and we have tested it in a subjective hearing session (section 7).

2. Acoustic characterization of car inside

In this section a quantitative analysis of sound propagation in cars will be presented. It will show why equalization in car is a critical task. Then an effective correction of sound field in car can be obtained, starting from an accurate knowledge of propagation effects in cockpit.

Due to the geometry and small size of the environment, several effect that do not affect hearing in huge cavity, are determinant in car. Among them the most critical are

- Early reflections
- Standing waves

Both this effect typically cannot be shut down easily.

Early reflections

Early reflections are effects tied with propagation in close environment. We suppose of sitting in a room where an omni-directional loudspeaker is playing an impulse from $t=0$, and a microphone is active as a receiver. The wave

arrive in front of walls and floor, impact them, and then come back, the returning wave is defined as reflection. The direction of the reflection is easily determined because it follows Fresnel's law used in optics. We can distinguish two kinds of waves:

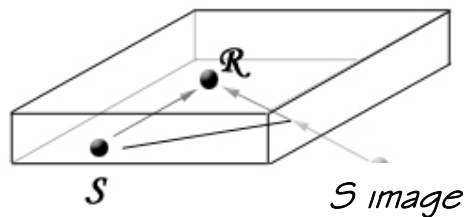
- The waves that arrive at the microphone after only one reflection. These have made a short way and arrive early at the receiver. They represent **Early Reflection**.
- The wave that makes several reflections and arrives at the receiver later. This represents the tail of the impulse response and gives the sound the typical reverb effect. These can be reflected hundreds of times and could be well-represented as a stochastic process. Indeed, they take sense only if we consider them in an energetic way.

The arrival time of early reflections is tied with the size of the environment where we execute the test. In a huge cavity, the early reflections arrive 50-100ms after the direct sound. As will be described later, this is a positive effect for human listening. On the other side, in a small room, the first reflection may arrive 10-50ms after direct sound, when our hearing system is still integrating the direct sound. This reflection has a different effect on hearing quality based on the energy associated. If its energy is below a certain threshold, we have a positive aid to the spatiality of the sound. But the energy associated with this reflection in a car is high, due to its short size and this behaviour affects the hearing quality.

Examples:

- Huge Cavity (20m): First reflection at time $t = 20[m] / 343[m/s] = 58ms$
- Small Cavity (1m): First reflection at time $t = 1[m] / 343[m/s] = 2.9ms$

Figure 1. Sound wave reflection



Standing waves

The latter effect is very critical because it distorts completely the harmonic response of the environment, especially at low frequency. This effect is still heavier if we consider that it is strongly dependent on the position of the receiver. There is a simple way to calculate the frequency resonance in an ideal room, with completely reflective walls. This law can be written as:

$$f = \frac{c}{n * L}$$

where:

c = sound speed.

L = size of cavity

n = order of the resonance.

Our task is to understand how this effect alters the hearing in a room. Consider the figure below:

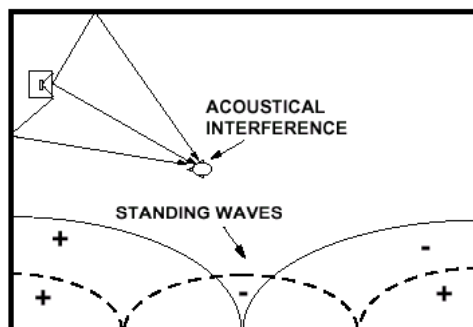


Figure 2 : Standing Waves

The signs “+”/”-“ correspond to a positive/negative value of pressure. The SPL (Sound Pressure Level) is not constant in the room but there are regions where its level is high and others where its level is zero. In the harmonic response, this is represented by the presence of a boost in the first case and a hole in the second one. This results in an unpleasant hearing effect, since at low frequency the real signal is masked by a fastidious *boom*.

3. Human hearing system

As detailed in the previous section in a small cavity as the car cockpit the acoustic energy is concentrated in a short time. So both early and late reflections are closed together. In order to understand how this phenomena degrade the hearing, we need to introduce several concepts about psycho acoustic.

- Haas effect
- Masking
- Spatial sensation

Haas Effect

Early reflection are high energy reflections that arrive to our hearing system few millisecond after the direct front. In this time, hearing system integrate all the information perceived and make a determined sensation. This effect is fundamental for the stereophony because allow to localize with high precision the source, to create only one sound image starting from two channel system. This integration effect is know as Haas effect. Haas established that integration time is equal to 25ms. During this period, our hearing system acquire information about sound localization, images, and intensity. Haas and masking effect, introduced later, allow us to know the influence of early reflection on hearing.

As we have already seen, this effect can be both a positive or a negative aid to sound quality reproduction depending by intensity of reflection, and delay after the arrive of direct front. This dependencies were been studied and the following curves were obtained.

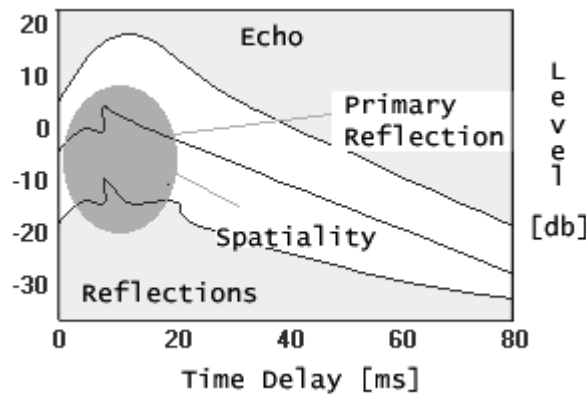


Figure 3 Effect of early reflection

The environment where we listen to the music has a fundamental role on the final reproduction quality. Low energy reflection give a positive aid to sound quality. On the other side high reflection energy distorts the dynamics and harmonic response of the original signal, moreover stereophony is deleted because we have two strong correlated signal, very close to each over. The impact of reflection on sound spatiality can be objectively measured by means of a cross-correlation coefficient, defined IACC. The formula that describe this parameter is reported below:

$$IACC = \max_{\tau} \left(\frac{\int_{t1}^{t2} p_{left}(\tau) \cdot p_{right}(t + \tau) dt}{\int_{t1}^{t2} p_{left}^2(\tau) \cdot p_{right}^2(t) dt} \right) \text{ where } -1ms < \tau < 1ms$$

where:

- $P_{left/right}$: pressures at left /right ear channel
- $t1/t2$: initial and final time of responses
- τ : period where we seek for the maximum.

Considering that in car early reflections occur with a high energy, we can understand why sound correction or equalization is such a difficult task.

As already mentioned another important effect is the “Masking”. It can be described only in terms of the human hearing acquisition process, which will be detailed briefly.

The ear is composed by several parts: outer, middle, inner. The last one can be mainly divided in cochlea, semicircular canals, auditory nerve. Cochlea is formed by two parallel channels, separated by basilar membrane, and filled with two different fluid, which run together. When we play a sound from a loudspeaker, we create a pressure signal that reach the ear and is transformed in middle ear in a mechanical signal. The movements of the stapes are

transmitted through oval window in cochlea, where a different pressure between the two fluid is generated. The variations in pressure are perceived by some haircells and transformed in a electrical signal which is sent to auditory nerve.

The basilar membrane have a fundamental role in acquisition process, because allow us to perceive several distinct tone playing together. Thanks to the basilar membrane we can listen the music. It is like a highly selective bank filter. Low frequencies set in motion the basilar membrane near the helicotrema, where is thick. Going away from this point towards oval window, it is thinner and it has resonance in higher frequency. Therefore two tones at different frequency can be perceived together, because they set in motion two different part of the basilar membrane. This effect is graphically depicted in figure 4.

Now we are able to understand the masking effect and its effect on hearing. There are two kind of masking: Temporal masking, Frequency masking.

Frequency masking

In order to understand frequency masking we propose an example. Assume that a 400Hz tone is playing with a certain amplitude (fig 4). After few instants, a 410Hz tone starts with a smaller amplitude while the first tone keep playing (fig 5). The higher movement of basilar membrane set up by first tone, made the second one inaudible. Only tone at frequency far enough from first one and with a great amplitude can be heard. This represents the frequency masking effect.

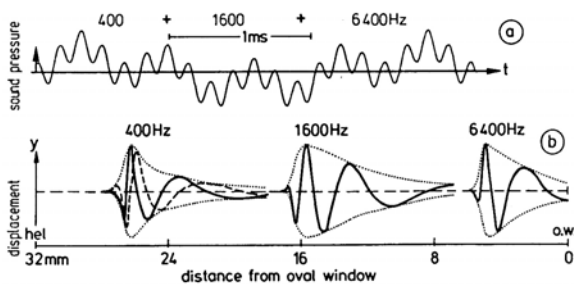


Figure 4. Basilar membrane perception

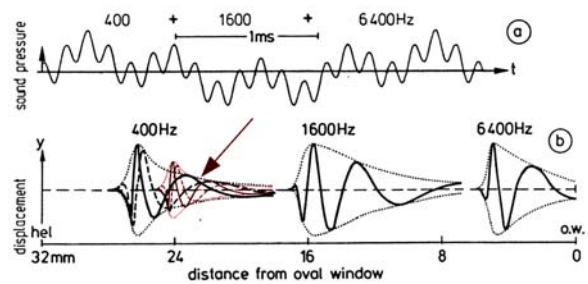


Figure 5 : Frequency Masking

Temporal masking

Temporal masking is not really distinct from frequency masking, but is a another face of the same effect. We propose an example similar to previous one. Suppose that a low frequency tone will be playing at time $t=0$. We refer at this one as “masker tone”. Now we consider a test tone that can be reproduced before, during or after the masker tone. The test tone can be heard only under particular condition. The curve figure the time and amplitude relation that make the test tone audible. We can distinguish three different zone: pre-masking, simultaneous and post masking.

Figure 6 represent this regions.

- Pre-masking: apparently has no physical justification. How can a test tone be masked by a tone which will be played after? The reason of this effect is that our hearing system need a certain time constant before reaching operating conditions. This constant is equal to 25ms when our system integrate the information perceived. If the masker tone falls in this period after the test tone, it contribute to final sensation tied with wave front. If the masker is greater than test tone, its weight on integral is dominant and the test tone is masked and cannot be heard.
- Simultaneous: This is the same effect as frequency masking.
- Post-masking: we have this effect when the test tone falls after the end of masker tone. The reason is that after a tone, the ear need a finite time, before reacquiring full sensitivity. This effect is strong especially at low frequency. In fact as detailed before the basilar membrane can be modelled as a selective bank filters. The filters feature a pole near the origin, a long time constant; therefore they are slow with a small dynamic. At high frequency, the effect is less evident and does not affect the hearing. In fact near helicotrema, where we perceive low frequencies, basilar membrane is thick and it is not under tension. If we think this part as a cord, we understand that the oscillations fade out after a long time and mask the following signals.

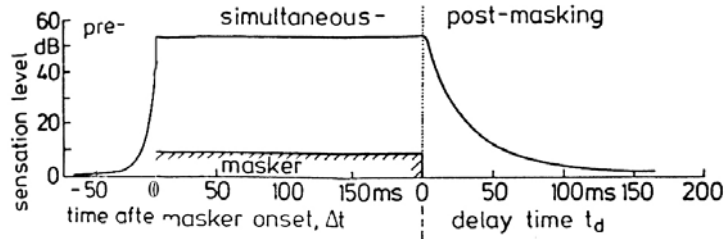


Figure 6: Temporal masking

Fig. 4 shows that the separation between different frequencies near the maximum is clear.

We can see in figure 4 that the separation between different frequencies near the maximum is clear. We can divide the membrane in several selective filters, centred in typical frequencies. We can refer to this scale as *Bark Scale*.

We have introduced physical phenomena that explain human hearing process. Now we have to relate masking effect and trouble in car as we have already done for Haas effect.

In car inside the major problems are produced by reflections and resonances. Highly reflective surfaces (windows) result in highly energetic reflections. This is a negative aid to hearing due to Haas effect, and makes the tail of response longer, causing masking effect. Moreover small sizes of cavity allow burning of resonance in audible frequency. This ones have long tails and are strongly affected from masking effect.

In order to compensate these effect we need at first an accurate measurements of the environment acoustic response from the dynamic point of view. Then we need to characterize and predict the effect of the environment on sound reproduction. This accurate acoustic measurement system is the **AQT Method**, which will be then used to compensate propagation troubles of sound waves..

4. AQT Analysis: Near musical Stimuli and audio system acquisition.

Characterization of the response inside a car is a difficult task because there are a lot of unknowns. The choice of the stimuli, of the microphone, of the position are only few examples. A great step towards the complete characterization of hearing inside the car was made with introduction of AQT Analysis. The acquisition process is complex, and human hearing system is more sensitive to transitory events because of masking, and Haas effect. In summary attacks and releases in sound source are far more relevant than the steady state information. It turns out that traditional methods are limited in order to give a complete characterization of the propagation in car. Moreover musical signal is strongly non-stationary. If we equalize the response of the car in a static way, we cannot obtain a reasonable improvement of quality.

AQT is a pre-existent method introduced by Liberatore [3]. The first version of this method was used as listening test and as graphic indicator. AQT stimuli signals were played in the room under test and the responses were recorded and evaluated in a graphic way. In this paper this method is extended to equalizer synthesis.

The true innovation introduced by AQT method is the stimuli signal. In order to measure the dynamic response of the system the stimuli is a train of burst with variable frequency, fig. 7.

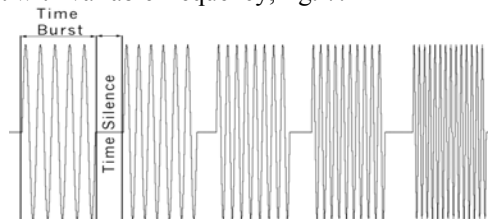


Figure 7: AQT stimuli

This stimuli is more closed to music, characterized by transitory events, and it allows to compute resonance frequencies. The response to AQT is close to the human hearing process, since we compute the values of response during attacks for each frequencies. Haas effect is accounted for keeping the duration of the burst at each frequency longer than human hearing system integration time.

AQT analysis produces two parameters which provide a quantitative evaluation of these effects:

- **Articulation:** estimates the speed of energetic recovery in an environment. Assume that we are playing a burst for a period of 200ms at a given frequency. After this stimulation, due to reflections, a certain time is needed before the energy extinguishes. Then a tail is associated to each frequency, whose length depends on the absorbing property of the materials inside the car. The frequency with a long tail response will be affected a lot by masking effect. On the other side short tail and higher articulation will have better aid on hearing.

- *Dynamic harmonic magnitude response*: represent the effective response perceived by our system. It plots for each frequency the value of the response during attack transient, instead of steady-state values. The block diagram of the proposed AQT method is reported in fig. 8.

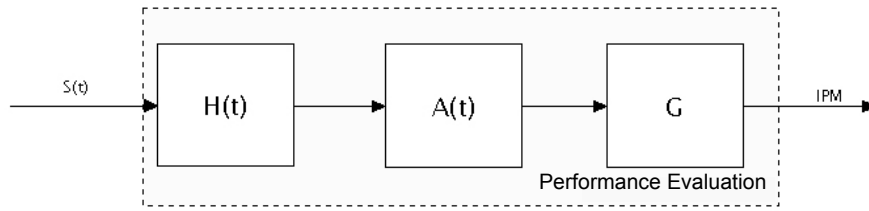


Figure 8: AQT block diagram

Where:

- $S(t)$: AQT Stimuli signal.
- $H(t)$: environment acoustic response to AQT Signal.
- $A(t)$: extraction of quality parameters : Articulation, dynamic harmonic response.
- G : valuation of the weight of these parameters on hearing.
- IPM : Index of measured performance.

The proposed AQT method, starting from articulation and dynamic harmonic response, automatically produces an objective index of acoustic performance of the environment under test, defined as IPM. This index proved to be very reliable and useful. It is well known in fact that subjective tests for audio systems are unavoidable, but very expensive, since they require a lot of man power, and tricky, since untrained listener can provide invalid results. Its validity was proven with comparison with a commonly adopted subjective index, IPA [5]. Since the correlation between IPA and IPM is very high for several different car inside the effectiveness of the AQT is proven. Fig. 9 shows experimental comparisons.

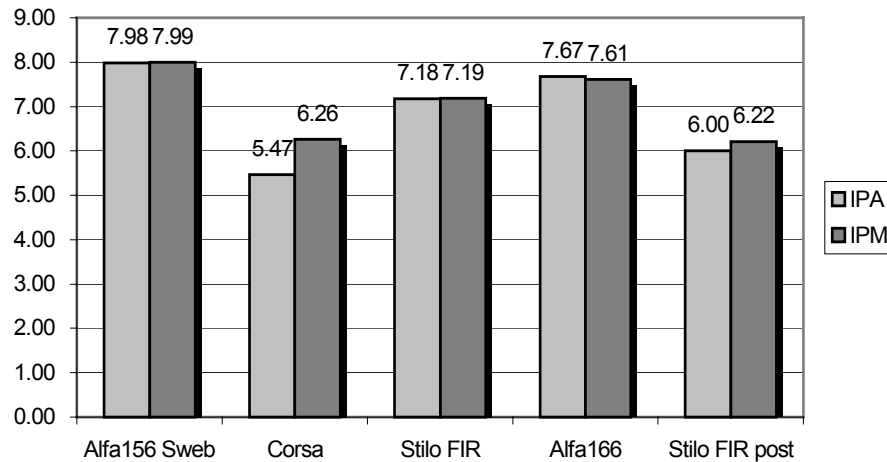


Figure 9: IPA - IPM Correlation

5. AQT Tool

The AQT procedure described was implemented with a Matlab™ graphical tool, which allow a quick and easy computation of AQT parameters, starting from the acoustic response of the environment in the time domain. Moreover a stand alone Visual Studio application is being developed, which will feature remarkable improvements and increase of execution A screen shot of the Matlab tool program is presented in fig. 10.

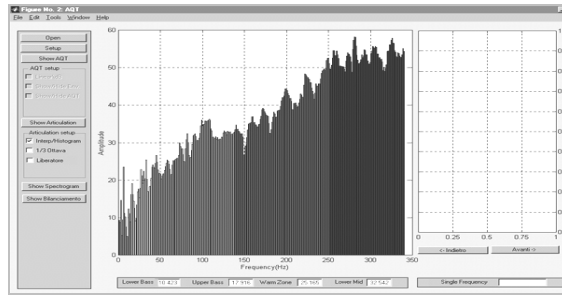


Figure 10: AQTTool

As an example, fig. 11 reports the AQT analysis of an Alfa 156 inside. Left chart shows the dynamic harmonic response in blue, and the usual steady-state response in magenta.

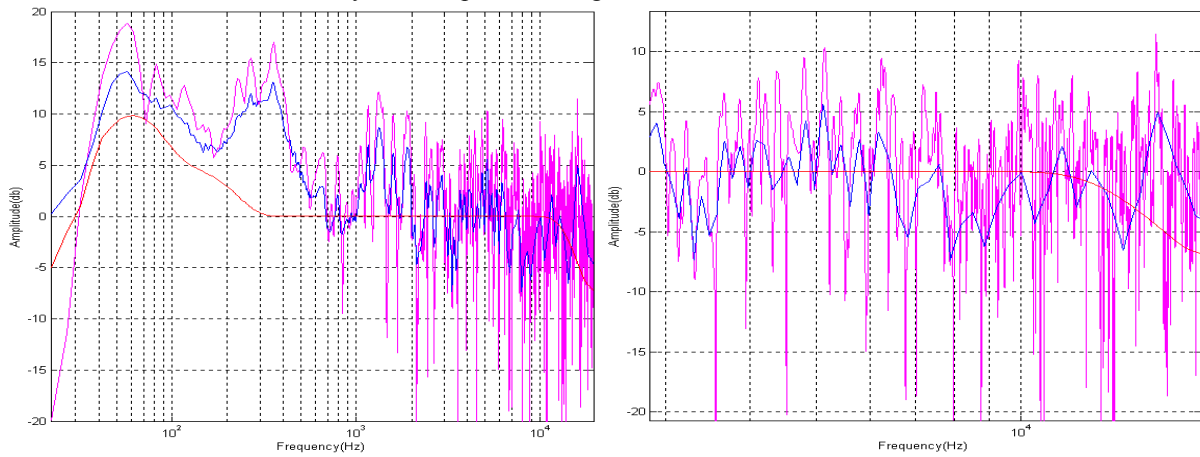


Figure 11: Alfa 156 AQT Dynamic harmonic magnitude response (blue) versus static response (magenta), enlarged around 10 kHz (right).

The right chart is a magnified version of the same measurement session, from which it results that the static response, especially at high frequencies, shows a strange behaviour, which does not appear in the dynamic response since it is an artefacts of the measurement process. Reflection, cancellation, combo filtering and so on produce a large number of thin oscillation in the frequency response. This is usually partially compensated, smoothing the response. However steady-state measurements do not characterize either masking or integration effects. The AQT dynamic response on the other side can be used as a nice starting point for the synthesis of an inverse filter equalization.

Fig. 12 reports the articulation of the acoustic response of the same car inside. Red arrows point out those critical frequencies which feature long tail and thus are affected by masking effect.

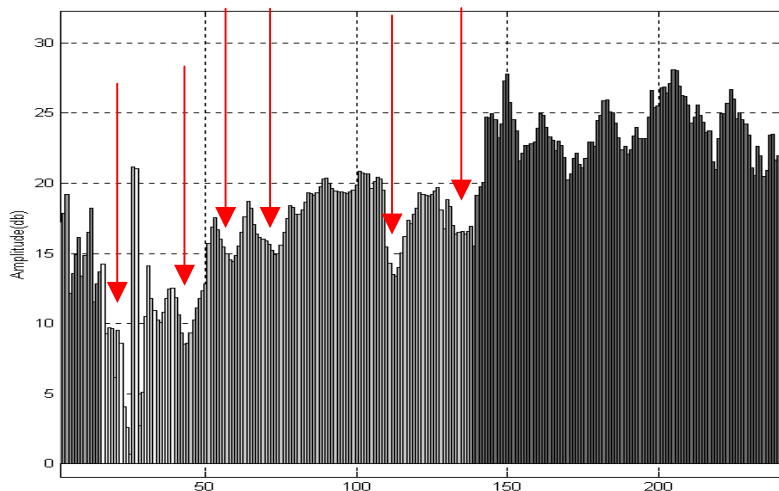


Figure 7: Articulation and Critical frequency

These data will be used in the following for both car inside acoustic characterization, and equalization. The former task is accomplished computing the IPM value corresponding to a car, the latter is accomplished synthesizing a suitable inverse filter to be used before the actual reproduction, in order to increase sound quality.

6. Digital psychoacoustic equalization

A new concept of equalization can be defined. It relies on dynamic frequency response and on articulation to design the equalizer shape. Specifically the inverse filter shape will be based on the dynamic frequency response instead of on the steady-state frequency response. Moreover the inversion will be based on a target frequency response shape, which corresponds to a maximum pleasantness. From the implementation point of view a DSP will be used to implement the psycho-acoustic inverse filter, realized as a standard FIR, fig. 13.

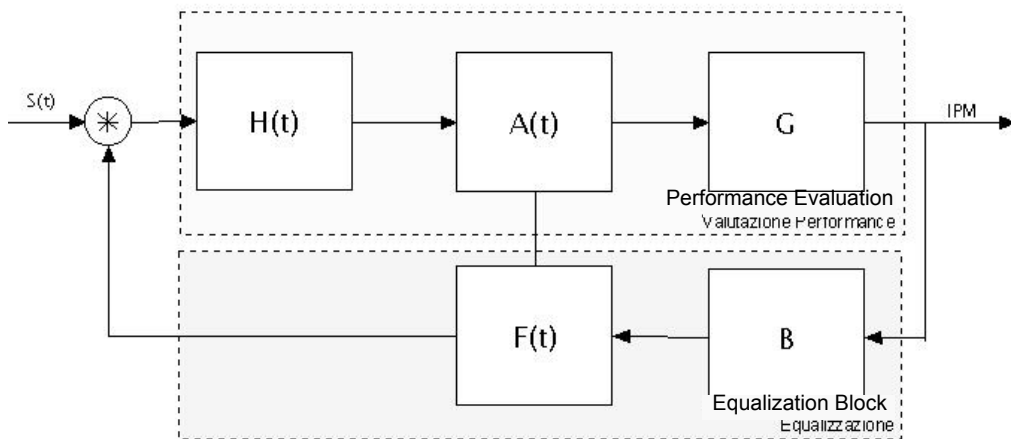


Figure 13: Psycho-acoustic inverse filter.

The direct path depicts the dynamic measurements system, where $s(t)$ is the AQT signal, $h(t)$ is the system impulse response, $A(t)$ is the analysis block, and G is the block of acoustic evaluation. A detailed description of $F(t)$ is reported in fig. 14.

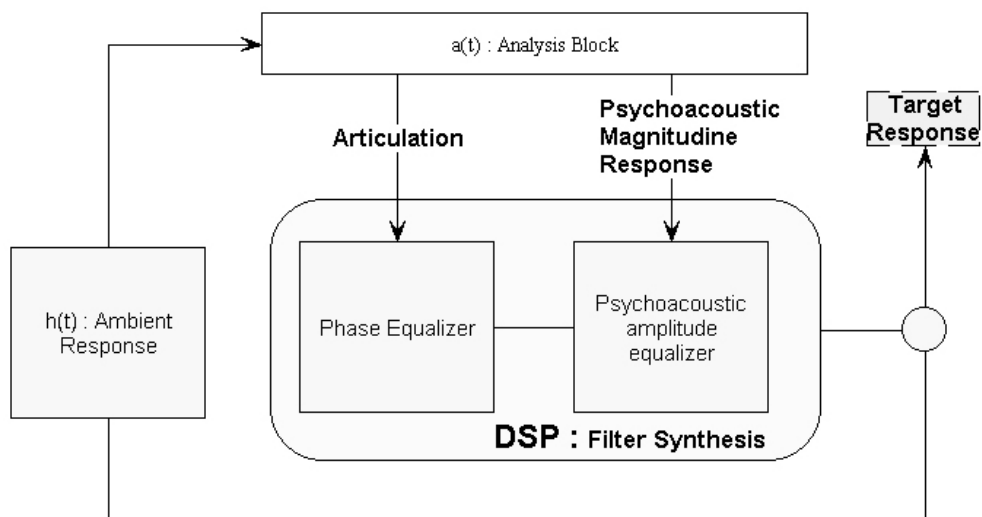


Figure 14 Psycho-acoustic inverse filter core.

Articulation is a quantitative representation of the time domain behaviour of the system at those frequencies, where long tails make the system dynamics low. A compensation of this phenomenon can be applied only in the time domain, and it can be achieved modifying the signal phase. Two main approaches are possible. The former uses a filter bank with variable phase, that cancel the most critical frequencies (e.g. car resonance frequencies). The latter uses non-minimum phase filter, which allows the correction of the response in the time domain. Experiments were performed on commercial cars and listening test confirm the innovative potential of the proposed approach.

7. Experimental results

Experiments were made implementing the filter shown in fig. 14 as a FIR filter in assembly code on a 32 bit floating point DSP-based Ez-kit development board (SHARC 21161N, fig. 15), with 4 analog audio-in sampled at 48 kHz, 24 bit.



Figure 15 SHARC 21161N development board.

Then an Opel Corsa with standard audio system was used for the experiments. It featured a 4 channels audio system with separated tweeter for front channels. Therefore a 4 channels equalizer was designed. The shape of the filters used was obtained with the AQT method detailed in the previous sections. The average level of equalization was limited between +6dB and -6dB with 1 dB resolution, and output levels were tuned in order to obtain similar average energy levels in the 4 channel. The tuning of the levels was made relying on a Audio Precision System 2022. At first the Sound Pressure level SPL response of the car was measured with Aurora [5] MLS, then the Virtual AQT was computed. Figure 16 show the SPL response of the Front left channel transformed in the frequency domain and smoothed at 1/6 of octave. Figures 17 and 18 show SPL measurements of the car FL channel with equalization, using standard inverse FIR technique (fig. 17) and the novel inverse AQT technique (fig. 18). In spite of what objective measurements show, the AQT equalization produces a more pleasant sound, as confirmed by AQT measurements.

In fact IPM vs IPA indexes for the FIAT Stilo, reported in tab. 1, confirm that the inverse AQT equalizing filter is preferable for the listener.

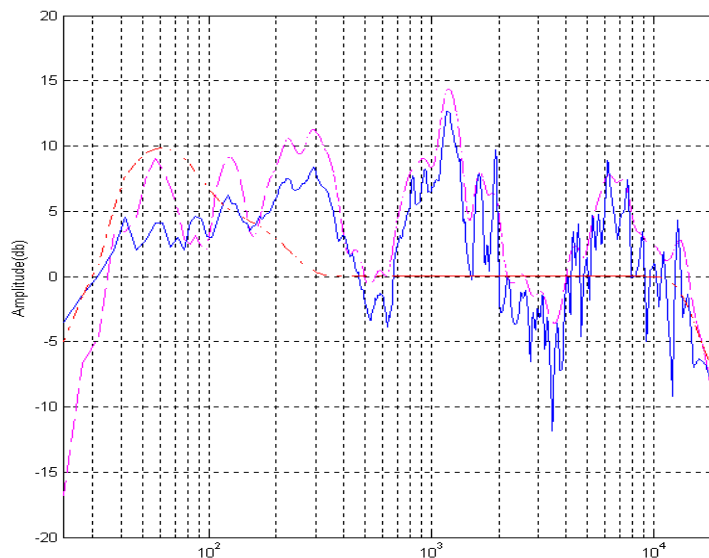


Figure16 Opel Corsa FL Channel : SPL measurements 1/3 Oct smoothed (dashed) vs AQT analysis no smoothed (solid), target curve (dash-dot),

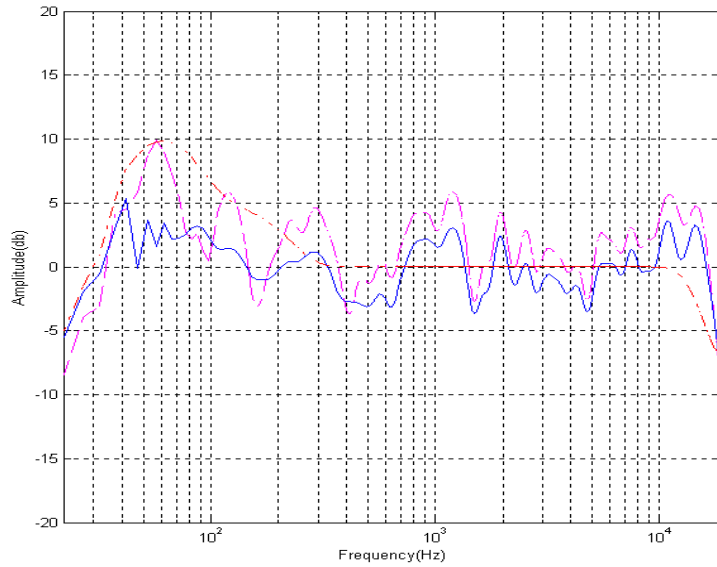


Figure 17 Opel Corsa FL Channel after traditional FIR equalization : SPL measurements 1/3 Oct smoothed (dashed) vs AQT analysis smoothed (solid), target curve (dash-dot).

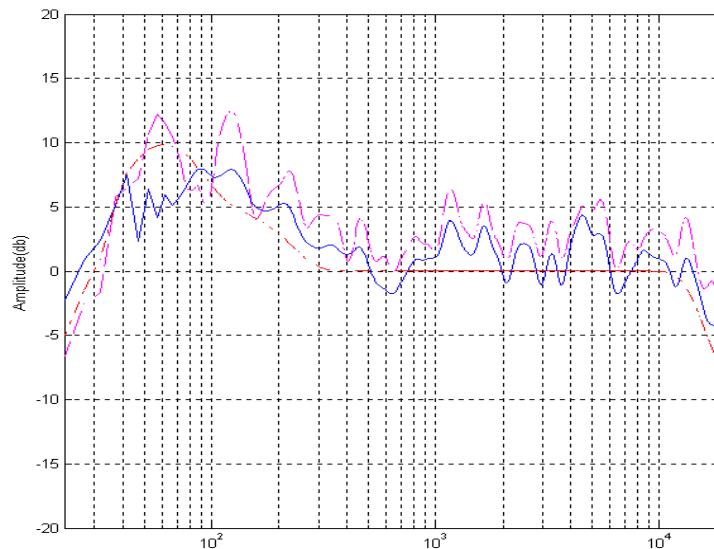


Figure 18 Opel Corsa FL Channel after traditional AQT equalization : SPL measurements 1/3 Oct smoothed (dashed) vs AQT analysis smoothed (solid), target curve (dash-dot).

Tab. 1 IPA – IPM Validation

Parameter	<i>IPA</i>	<i>IPM</i>
Standard audio system	7.18	Not Available
Equalized audio system, with inverse-FIR filters.	7.2	7
Equalized audio system, with inverse-AQT filters.	7.5	7.48

Conclusions

Acoustic equalization of a car inside is a very critical task, because of several effects that occur and that cannot be simply compensated inverting the steady state measurements of SPL inside the car.

In this paper a novel method for acoustic characterization is detailed and tuned for car inside, and its application to the acoustic equalization is presented.

Experiments show that the proposed method, referred to as AQT method, can be used successfully to replace listening tests with an objective evaluation of sound quality, and also as an approach to synthesize equalizing filters, pleasant for the listeners.

References

- [1] G. Cibelli, A. Bellini, E. Ugolotti, A. Farina, C. Morandi “Experimental validation of loudspeaker equalization inside car cockpits”, preprint 4898 AES 106th convention, Munich, May 1999.
- [2] S.T. Neely, J.B. Allen, “Invertibility of a room impulse response”, Journal of the Audio Engineering Society, May 1979, vol. 66, pp. 165-169.
- [3] I. Adami, F. Liberatore, “La messa a punto del sistema Diffusori-Ambiente”, Acustica Applicata srl, Via Roma 79, Galliciano – Lucca - Italy
- [4] A. Farina, G.Cibelli , A. Bellini, “AQT – A New Objective Measurement Of The Acoustical Quality Of Sound Reproduction In Small Compartments”, AES 110th Convention Paper, Amsterdam 2001
- [5] E. Ugolotti, G. Gobbi, A. Farina, “ IPA – A subjective Assessment Method of Sound Quality of Car Sound System” AES 110th Convention Paper, Amsterdam 2001
- [6] A. Farina and F. Righini, “Software implementation of an MLS analyzer, with tools for convolution, auralization and inverse filtering”, AES 103th Convention Paper, New York 1997.