Audio Engineering Society
Convention Paper 5835
Presented at the 114th Convention
2003 March 22–25 Amsterdam, The Netherlands

AQTtool an automatic tool for design and synthesis of psychoacoustic Equalizers

Andrea Azzali[1], Alberto Bellini[1], Eraldo Carpanoni[1], Marco Romagnoli[2], and Angelo Farina[1]

[1]University of Parma, 43100 Parma, ITALY

[2]ASK Industries, 42100 Reggio Emilia, ITALY

ABSTRACT

Steady-state characterization of acoustic environment is not enough for an efficient compensation of resonance and distortions. On the other side the common availability of digital systems has spread the use of acoustic equalizer at any level of sound reproduction.

Therefore a new concept of equalization need to 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. This is obtained relying on AQT methods, which use a variable frequency burst as a stimuli.

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. Moreover an automatic software tool was realized, which allows to perform AQT measurement with a user friendly GUI, and allows the user to synthesize a nice equalizer, fixing a few degrees of freedom.

The novel dynamic equalization method was especially tailored for a car inside. The equalizer 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. 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 acoustic phenomena that traditional methods cannot model. Among them noise floor, resonance, echoes make sound reproduction a difficult task. Specifically they keep the response far from the "Best Response Curve".

The main target of the equalization procedure is to increase sound comfort and make the response closer to the Best Response 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 fist step in the direction of a dynamic and psychoacoustic equalization was already made together with an approach for the characterization of acoustic attacks and releases.

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. Moreover an automatic software tool was realized, which allows to perform AQT measurement with a user friendly GUI, and allows the user to synthesize a nice equalizer, fixing a few degrees of freedom.

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. Starting from an accurate knowledge of propagation effects in cockpit an effective correction of sound field in car can be obtained.

Due to the geometry and small size of the environment, several effects not relevant in huge cavities play a key role in cars. Among them the most critical are:

- early reflections;
- standing waves.

2.1. Early reflections

Early reflections are tied with propagation in a 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 returning wave from walls and floor is defined as reflection. The direction of the reflection is easily obtained from Fressnel law as in optic. We can distinguish two kind of waves:

- The waves that arrive to the microphone after only one reflection. They make a short path and arrive early at the receiver. They represent **Early Reflection**.
- The waves that make several reflection and arrive to the receiver later. They represent the tail of the impulse response and produce the typical reverb effect. They are reflected hundred of times and can be effectively modeled from the energetic point of view. They are well represented by a stochastic process.

The arrival time of early reflections is tied with the size of the environment. In huge cavities early reflections arrive 50-100ms after the direct sound. As described later, this is positive effect for human listening. On the other hand in a small room, the first reflection may arrive 10-50ms after direct sound, when our hearing system is still integrating the direct sound. Depending on the energy associated early reflections result in different effects. If their energy is below a certain threshold, they contribute to increase the spatiality of the sound. If their energy is high, they decrease the sound quality. As an example in huge cavities (20 m) the first reflection occurs at time *t= 20/343=58ms*. In *s*mall Cavities (1 m) the first reflection occurs at time *t= 1/343=2.9ms.*

Early reflections are very critical because they distort completely the harmonic response especially at low frequency. Moreover it is strongly dependent on the position of the receiver.

2.2. Standing waves

Standing waves affect the harmonic response of a system. Specifically standing waves are created in correspondence of cavity modes, i.e. in correspondence of resonance frequencies.

There is a simple way to compute the resonance frequencies of an ideal room, i.e. a room with completely reflective walls:

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

where:

c = sound speed.

L = size of the room

n = order of the resonance.

Our task is to understand how this effect alter the hearing in a room. In fig. 1: 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 reach where its level is zero. In the harmonic response this is represented in the presence of a boost in first case and in 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*.

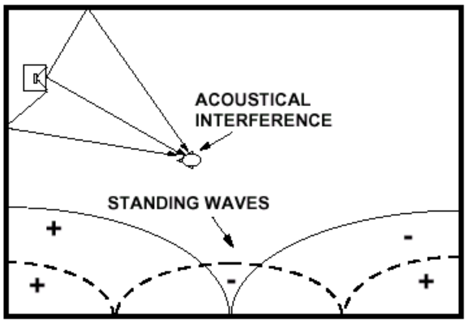

Figure 1 Standing Waves.

In summary sound reproduction in car is very critical since standing waves appears and early reflections occur with very high associated energy. Both effects contribute to decrease heavily sound reproduction quality and are difficult to compensate.

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 these phenomena degrade the hearing, we need to introduce several concepts about psycho acoustic.

- Haas effect;
- Masking (Frequency and temporal);

3.1. Haas Effect

Early reflections are high energy reflections that arrive to our hearing system few millisecond after the direct front. In this time, hearing system integrates all the information perceived and builds its sound sensation. As an example in stereophony it allows to localize sharply the source, and to create one sound image starting from a two channels 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, provide a physical interpretation of early reflections.

This effects can be both a positive or a negative aid to sound quality reproduction, depending on the intensity of reflections, and on the delay. This dependencies were studied thoroughly and the curves reported in fig. 2 were obtained.

The environment plays a fundamental role on the final reproduction quality. Low energy reflection give a positive aid to sound quality. On the other side high energy reflections induce negative phenomena. They reduce the dynamics and distort the harmonic response of the original signal. Moreover stereophony is deleted since the two channels are strongly correlated.

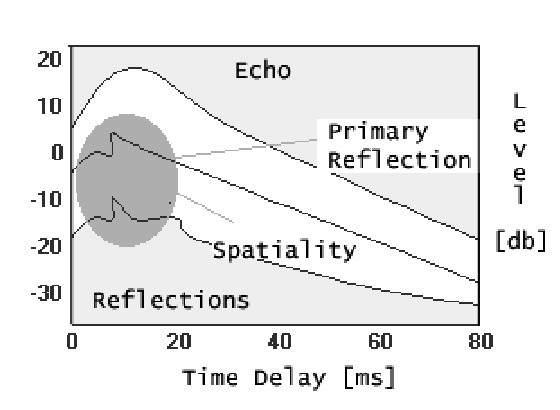

Figure 2 Effects of early reflections.

The impact of reflection on sound spatiality can be objectively measured by means of a cross-correlation coefficient, defined IACC:

$$IACC = \max_{\tau} \left(\frac{\int_{t1}^{t2} p_{left}(\tau) \cdot p_{rigth}(t+\tau)dt}{\int_{t1}^{t2} p_{left}^2(\tau) \cdot p_{rigth}^2(t)dt} \right)$$

where -1 ms < τ < 1ms;

$P_{left}/_{right}$: is the pressure 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.

3.2. Sound masking

Sound masking 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 the middle ear in a mechanical signal by ossicles named malleus, incus and stapes (figure 3) . The movements of the stapes are transmitted through oval window in cochlea, where a different pressure between the two fluids is generated. The variations in pressure are perceived by some haircells and transformed in a electrical signal which is sent to the auditory nerve.

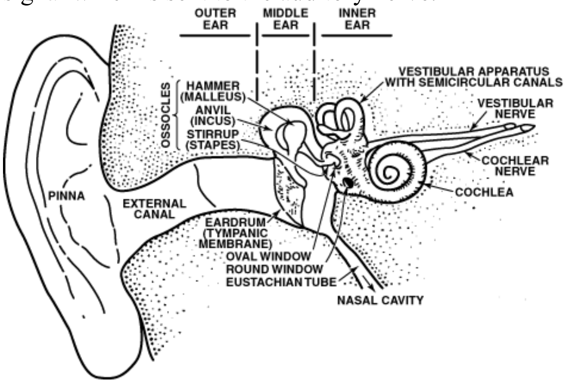

Figure 3 Human hearing system.

The basilar membrane plays a fundamental role in acquisition process, because it allows us to perceive several distinct tones playing together. Thanks to the basilar membrane we can listen to 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.

Specifically there are two kind of masking: Temporal masking, Frequency masking.

3.2.1. Frequency masking

Let us consider an example to describe frequency masking. 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 keeps playing (fig. 5). The higher movement of the basilar membrane set up by first tone, made the second one inaudible. Only tones at frequency far enough from first one and with a great amplitude can be heard. This represents the frequency masking effect.

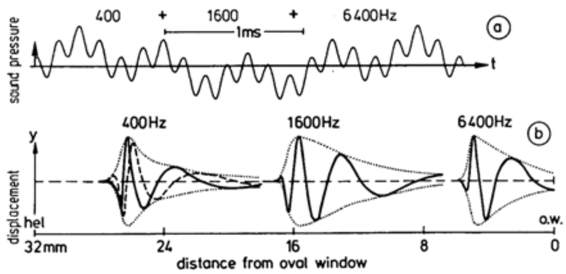

Figure 4 Basilar membrane perception.

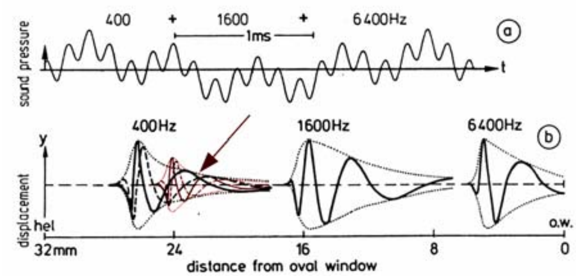

Figure 5 Frequency Masking.

3.2.2. Temporal masking

Temporal masking is not really distinct from frequency masking; it is a another face of the same effect. Let us consider an example similar to the previous one. Suppose that a low frequency tone, referred to as “masker tone” is playing at time t=0. Now we introduce a test tone that can be reproduced before, during or after the masker tone. The test tone can be heard only under particular condition. Fig. 6 shows the time and amplitude relation that make the test tone audible. We can distinguish three different zone: pre-masking, simultaneous and post masking.

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 needs a certain amount of time to restore operating conditions. This time is 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.

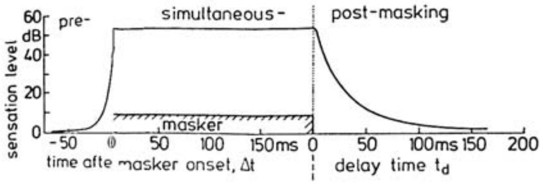

Figure 6 Temporal masking.

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 for listening tests and as a graphic chart. AQT stimuli signals were played in the room under test and the responses were recorded and drawn in a chart. 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.

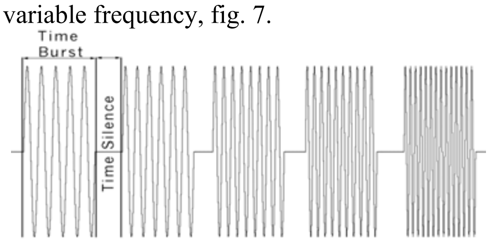

Figure 7 AQT stimuli.

This stimuli is more closed to music, characterized by transitory events, and it allows to compute resonance frequencies. Then 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:

- <u>*Articulation*</u>: 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, 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.

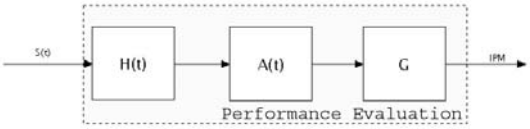

Figure 8 AQT block diagram.

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.

5. AQT Tool

The AQT procedure described was implemented with a graphical tool, which allows a quick and easy computation of AQT parameters, starting from the acoustic response of the environment in the time domain. It is a stand alone Visual Studio application, which features remarkable improvements and a fast execution A screenshot of the tool program is presented in fig. 10.

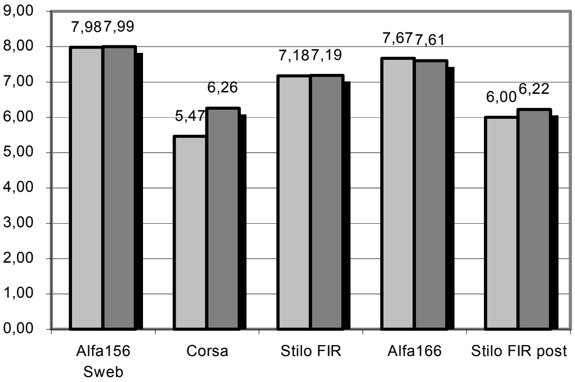

Figure 9 IPA - IPM Correlation: IPA (light grey) versus IPM (heavy grey).

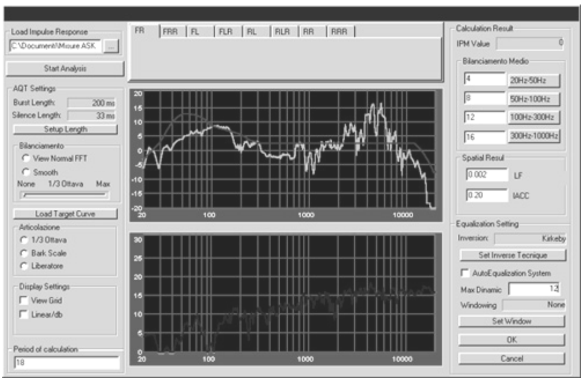

Figure 10 AQTTool screenshot.

AQTtool is also available as a CoolEdit plug-ins, that can be conveniently used in combination with other plug-ins for car inside characterization and compensation [6]. As an example, fig. 11 reports the AQT analysis of a FIAT Stilo Front Left channel. It shows a comparative analysis of equalized versus non equalized AQT response. Fig. 12 and 13 shows the AQT dynamic response versus the target curve of equalized and non equalized car audio system respectively.

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.

Fig. 14 reports the articulation of the acoustic response of the same car inside. Arrows point out those critical frequencies which feature long tail and thus are affected by masking effect and low articulation.

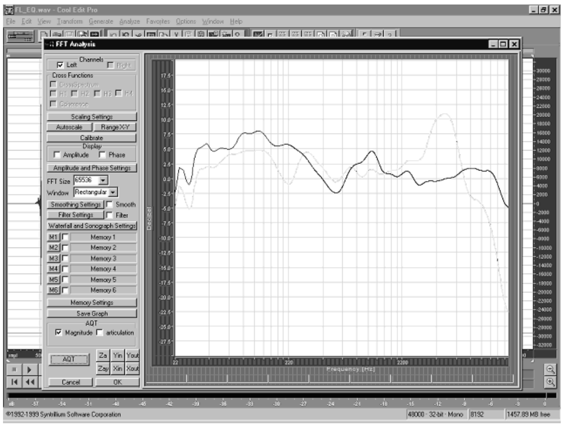

Figure 11 Stilo AQT Dynamic harmonic magnitude response non equalized (grey) versus equalized AQT dynamic response (black) of FL channel.

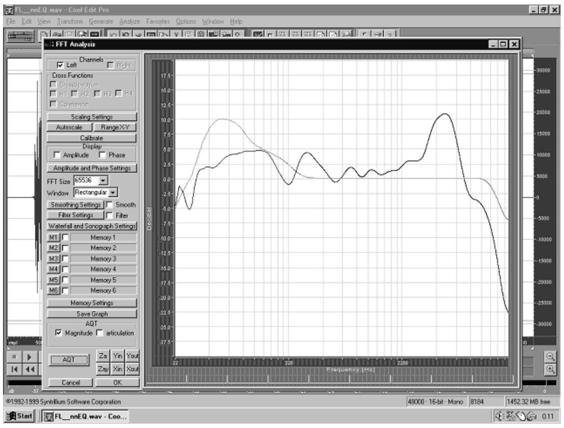

Figure 12 Stilo AQT Dynamic harmonic magnitude response non equalized (black) versus target curve (grey) of FL channel.

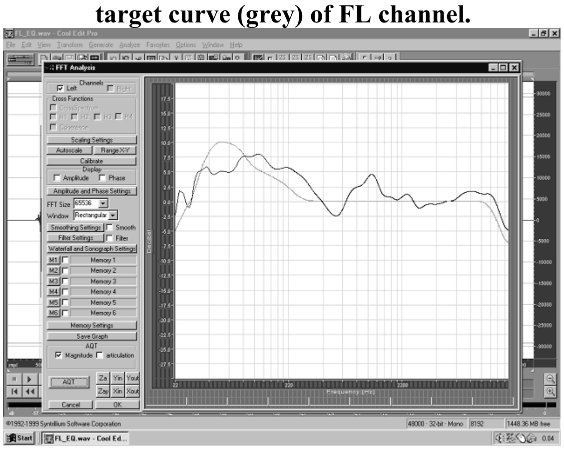

Figure 13 Stilo AQT Dynamic harmonic magnitude response equalized (black) versus target curve (grey) of FL channel.

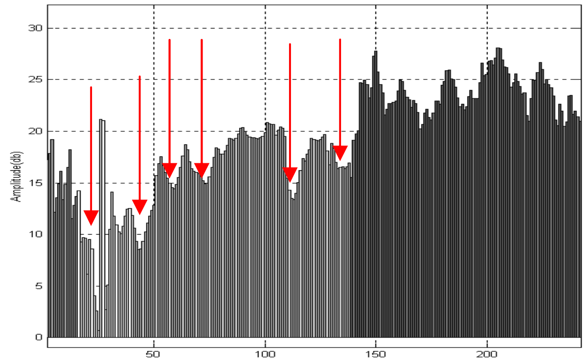

Figure 14 Articulation and Critical frequencies.

6. DIGITAL PSYCHOACOUSTIC EQUALIZATION

A new concept of equalization can be defined. It relies on AQT 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. 15. The latter architecture will be referred to as AQT filter in the following.

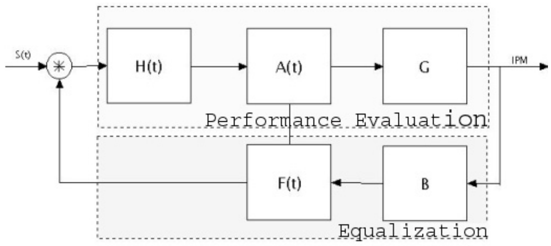

Figure 15: Psycho-acoustic inverse filter.

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. 16.

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.

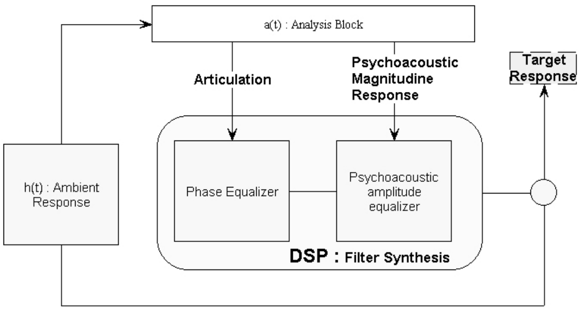

Figure 16 Psycho-acoustic inverse filter core.

Future developments will make available single programs embedded in AQTtool for measurements, articulation computation, AQT filter synthesis, spatialization and auralization. Main purpose is to obtain automatically an equalization system able without any input. The tool at first will characterize car acoustic response, then it will produce be a series of filter coefficients that can be loaded in any general purpose DSP board for sound processing.

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 architecture shown in fig. 14 as a FIR filter in assembly code on a 32 bit floating point DSP-based board (SHARC 21161N, fig. 15), with analog audio-in sampled at 48 kHz, 24 bit. The car is equipped with a four channel audio system, and each channel is operated independently.

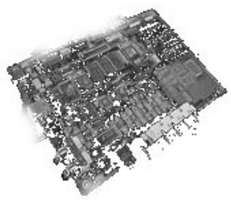

Figure 17 SHARC 21161N development board.

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.

	IPA	*IPM*
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*

Tab. 1 IPA – IPM Validation

8. 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. A stand alone program AQTtool was developed to compute and design AQT filters with a user friendly GUI.

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, Gallicano – 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 103 th Convention Paper, New York 1997.