Audio Engineering Society

Convention Paper 8926

Presented at the 135th Convention
2013 October 17–20 New York, NY, USA

Linear phase implementation in loudspeaker systems; measurements, processing methods and application benefits

Rémi Vaucher,

R&D Engineer, NEXO, France (remi.vaucher@nexo.fr)

ABSTRACT

The aim of this paper is to present a new generation of EQ. It provides a way to ensure phase compatibility from 20Hz to 20kHz over a range of different speaker cabinets. This method is based on a mix of FIR filters and IIR filters. The use of FIR filters allows a tuning of the phase independently from magnitude and allows an acoustic linear phase above 500Hz. All targets used to compute FIR coefficient are based upon extensive measurement and subjective listening tests. A template has been set to normalise the crossover frequencies in the low range, enabling compatibility of every sub-bass with the main cabinets.

1. INTRODUCTION

The goal is not to discuss FIR filters, often associated with Linear Phase, but how and why to use them. Implementing a FIR filter is not an achievement in itself; one must be focused on the result obtained and the way to obtain this result. What we have in mind here is the ability to achieve a coherent phase scheme (20Hz-20khz) across a range of cabinets (developed over a 15 year period), as well as a gain structure based on the SPL capability of each cabinet (not discussed in this paper). Contrary to all-purpose digital speaker management systems, where EQ is left to the end users with no specific knowledge of the speaker cabinets, our dedicated NXAMP processed amplifier gives us the opportunity to implement exactly our needs.

2. HISTORY IN SYSTEM EQ

2.1. Analogue Times

The history of EQ has so far been a constant struggle with hardware (and DSP resources since digital times) to achieve a certain result with limited resources. Due to hardware restriction, a limited number of EQs biquad were available a two decade ago. In the case of “NEXO Analogue TDcontrollers” for instance, we were limited to one 3rd or 2nd order High pass filter, a Low pass filter and 4 parametric EQ (among which only 2 were cutting). No time delay. (Analogue all-pass filters alignment is consuming a lot of expensive operational amplifiers…). So with this limited number of EQ (and a practical impossibility to upgrade) it was a challenge to reach a good EQ compromise (and sub-bass alignment) and nearly impossible to have a match with other older or different type of cabinets. Here is an example of the equalisation possible on such systems (Figure 1, electrical response):

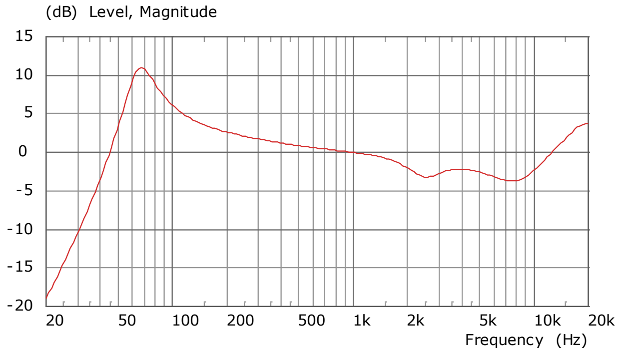

Figure 1 : Analogue Filtering

2.2. Digital Times

Early DSP based controllers appeared in the early 90s thanks to the digital technology development. Unfortunately the units of those early ages had drawbacks which prevented really wide acceptance. (Latency time < 30ms, cost of the units, SNR around 90dB…to name a few).

In 1999 NEXO released the NX241 based on the fixed point 24bit Motorola 56303. At that time the EQ implemented in the NX241 was conservative and reproduced the analogue recursive filtering of their analogue controller counterparts, with the advantage of a software upgrade, a greater number of EQ (still 2nd biquad, Low Pass, High Pass, All Pass, Band Pass, notches, Shelving high and low…) and delay.

Since that date it has been easier to phase align (with all pass filters and pure delays) the lowest frequencies (Figure 2, representing the phase of 10 different sub and bass cabinets), allowing a good compatibility under 200Hz between systems.

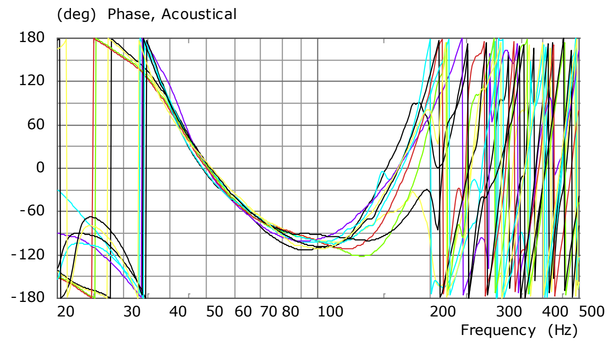

Figure 2 : LF phase alignment

In 2006 we introduced FIR filtering in our setups range. Not for the so-called 'linear phase' behavior but for the ability to modify the phase independently of the magnitude. This was particularly useful for tuning the cardioid directivity of our active cardio sub basses (2 drivers involved with each having its own signal). In that case we wanted to have a specific phase for each frequency that maximizes the ratio of 'power sent to the front' over 'power sent to the back'. This gives a target phase that would not be possible to achieve with standard recursive filtering.

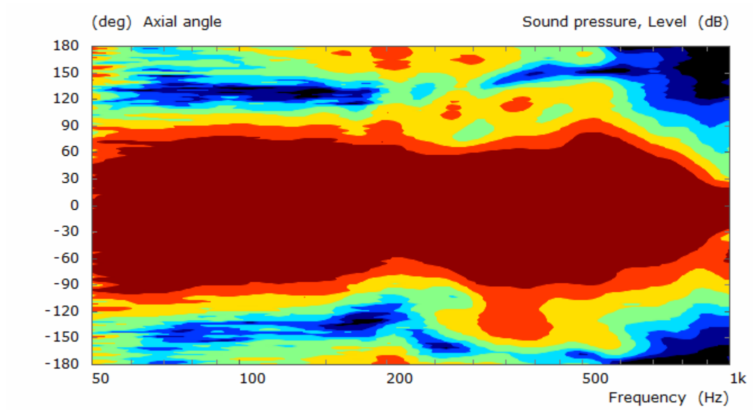

Figure 3 : Cardio coverage plot

Figure 4 (next page) shows the filter applied to the back driver in order to achieve the above cardioid pattern Figure 3 (coverage plot, off axis response in ordinate, frequency in abscissa, color is the acoustical amplitude, blue color represent back rejection).

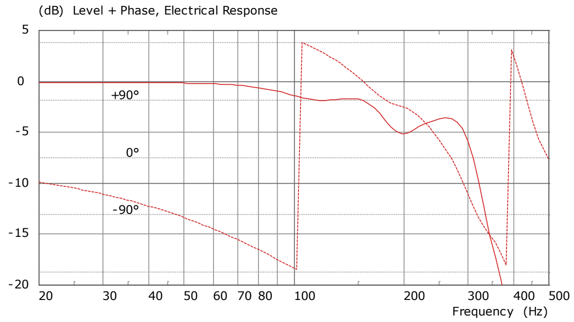

Figure 4 : FIR filter applied to back speaker for cardioid pattern

The FIR is also used in a more obvious way to equalise the system with far better accuracy, with a certain “blindness” that obliges you to precisely define what you want to equalise. The main advantage is also that you can achieve acoustic linear phase that will ensure that every system tuned will be phase compatible from 20Hz to 20kHz. (We are not speaking here of the electrical linear phase obtained with symmetrical FIR filters).

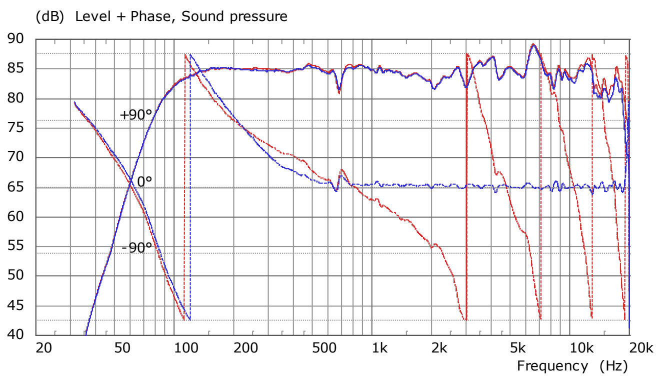

Figure 5 : Recursive IIR versus FIR filtering (acoustical)

Figure 5 above shows the effect of a recursive filter equalisation (in red) versus the effect achieved with a FIR filter in blue (same amplitude response). Notice the phase behaviour above 1kHz.

Figure 6 shows the level of detail possible on the equalisation when equalisation amplitude with FIR filtering. (red is FIR, blue is IIR)

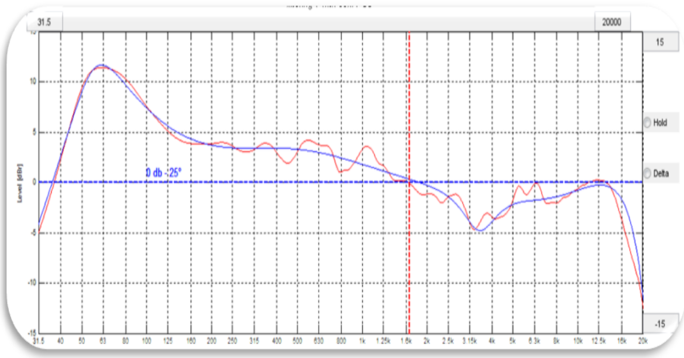

Figure 6 : Recursive IIR versus FIR filtering (electrical)

Now there is a bunch of FIR based Digital audio processors on the market which offer a consequent DSP resource that allows massive computing force. The question is: how to use this new kind of filtering? Some choose to be transparent for the user keeping a conservative user interface and propose filter shapes that weren’t possible with recursive EQ, while others try to implement as well measurement functions to ease implementations based upon impulse responses. This works more or less for architectural EQ or for listening rooms but using FIR filtering to equalize sound systems is trickier…

3. RANGE EQ VS SYSTEM EQ

It is now very easy to tune a system with a pair of good ears and decent measurement tools. The problem arises when you want to be compatible with other systems (aligning for instance horn loaded cabinet, reflex, close box, band pass cabinets…) that have massive differences in their time alignment.

3.1. Perfect phase match

As said before it is fairly easy to maintain phase compatibility in the lower frequencies. The following Figure 7 displays two cabinets whose amplitude responses (red and blue curves) are the same but with different phase responses.

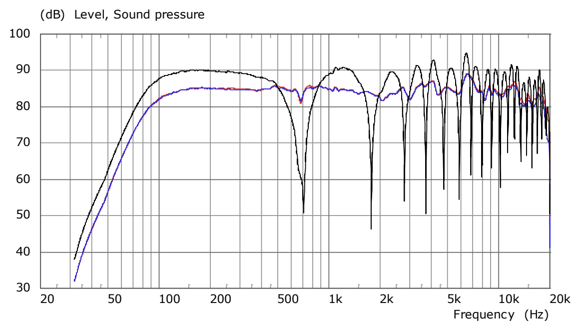

Figure 7 Summation of two cabinets with incompatible phase matching

As a result, the sum of both cabinets is a comb filter (in black). In many cases, this is not an issue as we could only use one cabinet alone with its subbass or when the two cabinets are physically apart one from the other. After all 1ms (cancellation at 500Hz) is only 34cm…The problem arises when using different cabinets side by side or cabinets within a line array that may have different EQ. In that case you don't want the EQ made on the system to affect the phase of the array.

3.2. Uniform Crossover

Another point which is not relevant when you tune a system (main cabinet plus subbass) on its own is the repeatability of the acoustic crossover with the other systems. When you tune an independent system you will take care of the proper alignment and summation of cabinet A and Sub bass B for instance, and for each different case you will choose the type of filter and slope for the best result.

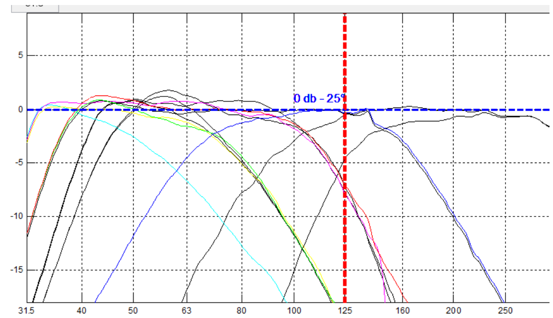

Figure 8 : LF Crossover template

If you choose to match a dozen sub basses with twice this number of different cabinets at several crossover points, you cannot proceed this way and you would be obliged to rationalise the different crossover frequencies and slopes so that all cabinets fit into the same template (Figure 8). By doing so, you can swap every cabinet with every bass and still have the same overall response for each combination (the low end extension and SPL capability excluded of course).

4. SOME THEORY

4.1. FIR vs IIR

It is not the goal of this paper to describe implementation of FIR or IIR as a literature full of examples already exists [1],[2].

What is more interesting is to have a clear understanding of the generation of FIR coefficients.

In the frequency domain, we define the target of the filter (knowing that phase can be defined independently of magnitude). This target may be obtained from an analytic equation, a measurement post treatment, even hand drawing, or a mix of all the above. Restricting the target to the inverse of a single location on axis measurement does not work.

Once the frequency target is defined, we obtain the time domain counterpart thanks to a Fast Fourier Transform and get the impulse response (Figure 9 next page). If the filter wanted is non-causal it will be necessary to apply a pure delay in order to have the maximum of the impulse response in the "positive" time.

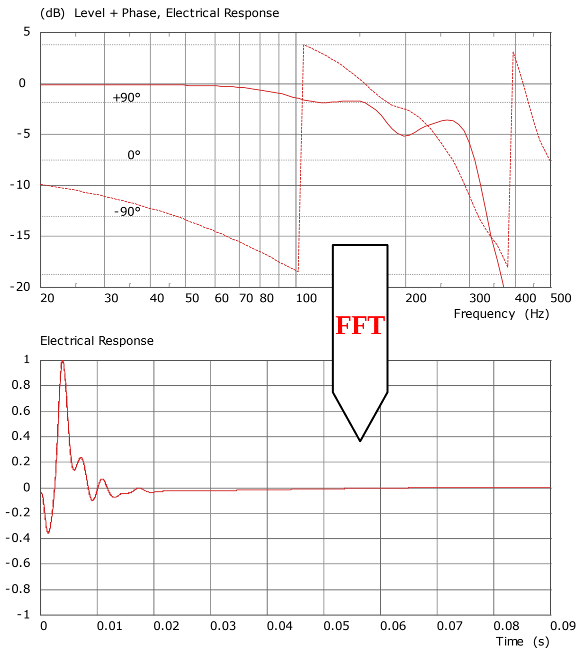

Figure 9 : Time / Frequency domains

Once we get the impulse response, it is sampled at the sampling rate and truncated/windowed (hence the term Finite Impulse Response). The longer numbers of coefficient (taps) the longer LF extension. Note also that a high number of taps is do not mean high latency as it is commonly believed. This latency issue is only determined by the impulse response. Sampling the impulse response of an IIR filter will give exactly the same latency (as well as the magnitude phase behavior – down to a certain frequency-). Only FIR filters that have a symmetrical impulse response, and thus an electric linear phase filter, introduce a phase delay (latency) equal to half their coefficient x 1/sampling rate.

4.2. Linear Phase: Acoustic or Electric?

We need also to stress the fact that the linear phase that we are dealing with is acoustic and not electric. What matters is the acoustic result obtained. In many cases, the speaker systems are minimum phase devices (except for passive crossover systems) and equalising with a minimum phase EQ (like IIR for instance) is also equalising the phase. In minimum phase systems, the phase and amplitude are linked by the Hilbert function, and thus on the same system it won't be possible to keep the same phase while changing the amplitude with minimum phase EQ (recursive, IIR).

A common mistake in the early digital times was to use a symmetrical FIR to obtain a "phase distortion free" filter. Yet the filter itself was distortion free but it would not linearize the acoustical phase and beside that was leading to un-acceptable latency time for live sound (especially monitor application).

It is thus possible to have a partially (let's say above 500Hz) linear phase with a minimum cost of fixed latency (under 4ms). Which is fitting under the 5ms that we consider to be a maximum acceptable in live sound application.[5]

5. MEASUREMENT

As we've said, the FIR filter is obtained from the impulse response of a frequency target. Taking blindly an inverted frequency target response measured in one location will indeed produce a very flat curve (magnitude and phase) at this very location. But the result will just be nice on the paper; this kind of simplistic approach won't work in real life.

The FIR target shall thus be computed (see next chapter) and will require an extensive measurement campaign on the cabinet in order to know where and what to equalise.

5.1. Measurement setups

All measurement are made in the anechoic chamber and laser calibrated to provide the best accuracy possible. The physical precision shall be no more than a fourth of the wave length at 20kHz (4.3mm). All individual phases are then further post processed to guarantee a good complex average.

5.2. Production statistics

The first obvious measurement batch will be to assess the production consistency (Figure 10). Perfection in our industrial world does not exist and the EQ made will have to work on every cabinet and not only on the R&D golden sample.

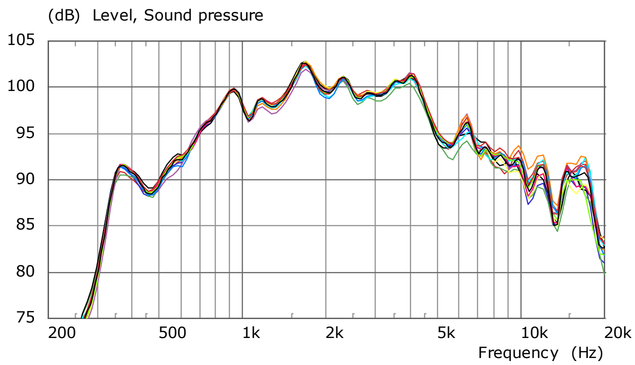

Figure 10 : Production statistics

A standard deviation analysis (Figure 11) will underline the frequency zones where the response curve is likely to vary from one cabinet to the other. In the zone of interest of the FIR filter (>200Hz) we will mostly have:

- Upper HF range of compression drivers (due to the spacing of the diaphragm with the phase plug)
- Crossover zone in case of passive cabinets (due to the summation of the tolerances of the different inductors & capacitors of the filter).

In the figure below for instance we would particularly take care of areas where standard deviation is above -20dB

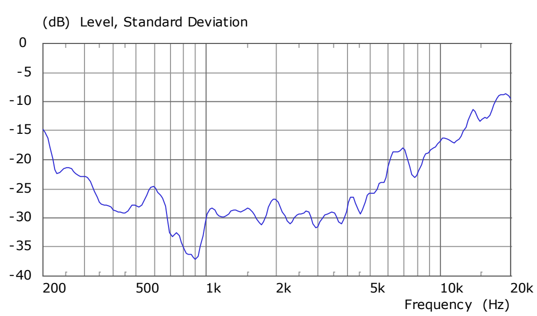

Figure 11 : Production statistics, standard deviation

5.3. Ageing effects

It is also mandatory to take into account the ageing of the components, mostly the change in the stiffness of the suspension that will affect response curve during the lifetime of the cabinet. An equalisation made on brand new speakers may fix response anomalies that are due to disappear, naturally resulting in the adverse effect.

5.4. Coverage averaging

As said before, a filter target made in a single location in front of the cabinet will only EQ properly on-axis response. It is mandatory to measure the coverage of the cabinet in order to have an idea of the off-axis behaviour of the cabinet. Unfortunately the filter is only one dimension when the DUT would require 2 dimensions (even 3 if we consider non-linearities appearing at different SPL). Unless systems are equalised for a single listening location (studio monitor room for instance), we are dealing in live sound and the necessity to have a similar sound all across the listening area.

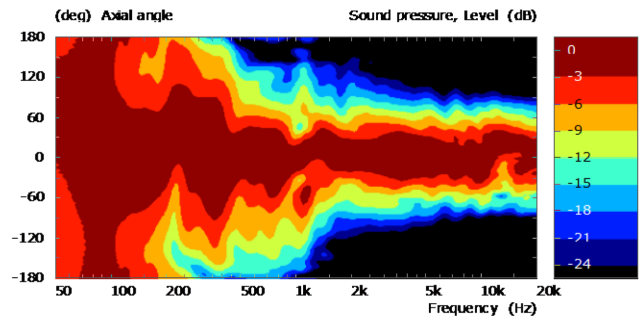

Figure 12 : Polar coverage

On the example Figure 12 (coverage plot normalised to on axis), we can see the difference between the wanted EQ when moving 27° off axis (Figure 13). Should we equalise the cabinet only taking care on axis response, it is obvious that it won't be listenable in an off axis location.

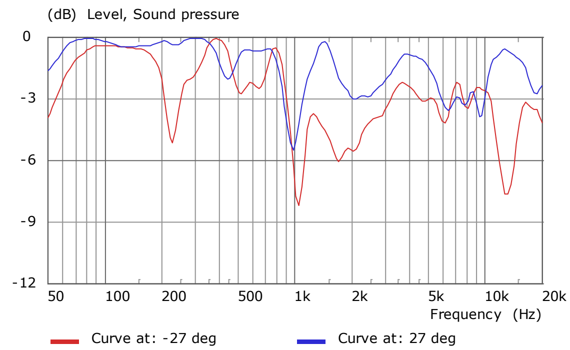

Figure 13 Off Axis responses

It will be again necessary to compromise on the EQ between on-axis and off-axis response. The coverage plot will help us to determine an average curve made on the listening angle. Let's consider that as a special averaging that will smooth out the small diffraction effect. It will also stress the zone where our FIR target won't be the best way to equalise the system (mostly in zones where there are large interferences).

5.5. HF CrossOver

Before implementing the HF crossover, displacement and distortion (mechanical & acoustical) shall be made to assess the working range of the HF driver and the upper frequency range of the driver (cone mode can produce sub-harmonics when driven at very high SPL). Refer to chapter 6.1 page 7.

5.6. Synthesized curve

From the above measurement, we build a synthesised curve that will be our reference curve. This curve is a mix of the averaged contour on the listening angle, weighted with the statistical analysis performed on the production data as well as information about ageing. In some case, even some modelling information is included (mostly for very low frequencies when the measured signal is noisy or suffering from reflections).

6. IMPLEMENTATION

Once we get our set of measurements, we know what shall be equalised and what should not. We need now a good user interface to transform this knowledge into filter coefficients.

6.1. Defining HF Cross Over

Although it is not the objective here to discuss how to set up a crossover, we can stress the benefit of implementing very steep slopes that were not available with recursive (analogue or digital) filters. It is also worthwhile to note that phase alignment is no longer a problem. As the FIR filtering allows us to tune the phase independently, all different crossover shapes will be perfectly aligned.

It will be possible to try out several cut-off frequencies (within the limit found by distortion measurements of the different components) with different slopes. Figure 15 and Figure 14 shows the acoustical curves (magnitude above, phase below) obtained for 3 different cut off frequencies with 2 different slopes.

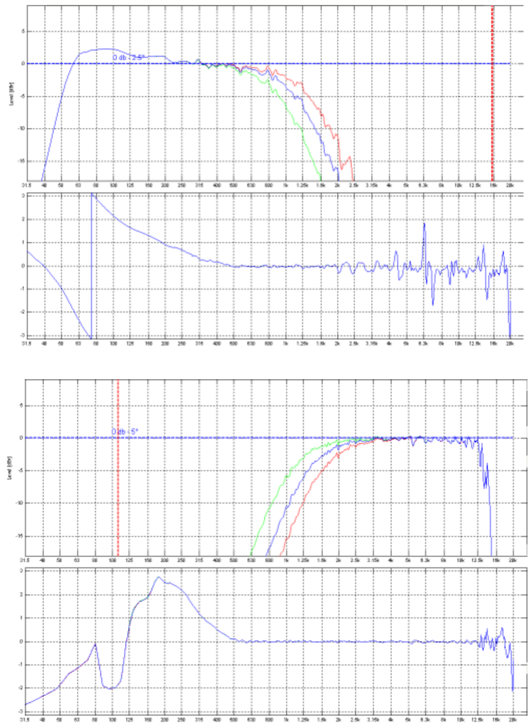

Figure 14: low order crossover, amplitude and phase

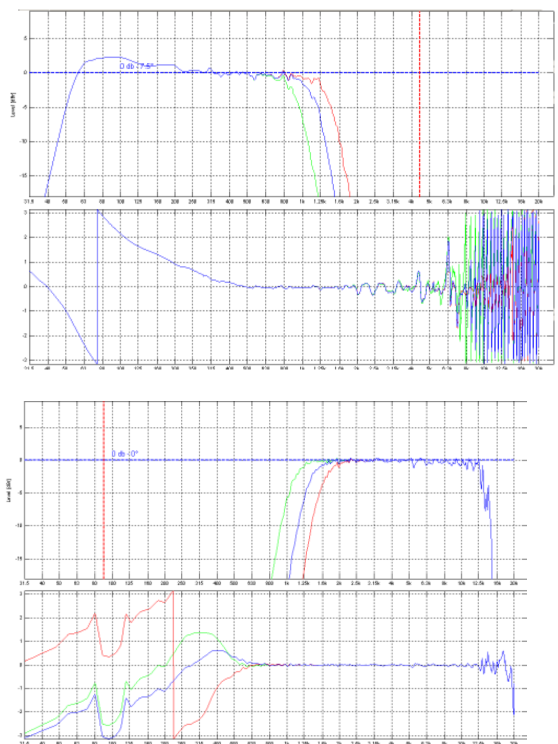

Figure 15 : Steep Crossover, amplitude and phase

Distortion measurements (as seen on Figure 18 and Figure 19) on the recombination of the HF and LF drivers (in addition to coverage measurement as seen in Figure 17) helps to determine the optimum cross over frequency. The figures below stress the difference between "traditional" low order slope (max 8th order) with high order slope (equivalent to 40th to 80th order).

The coverage plot -Figure 16 & Figure 17- (listening angle ordinate, freq abscissa) shows a clearly narrower perturbation zone on the second graph (which is of course the high order slope)

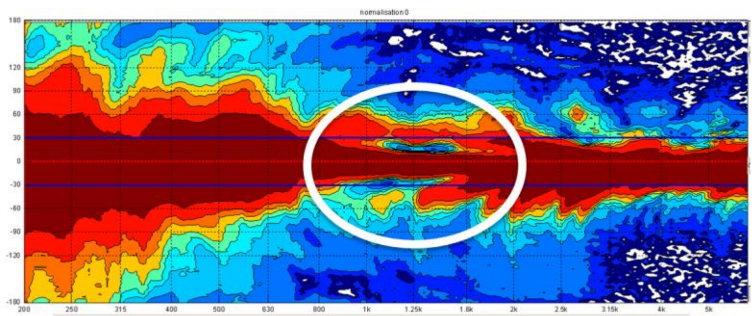

Figure 16 : effect of low order crossover on coverage pattern

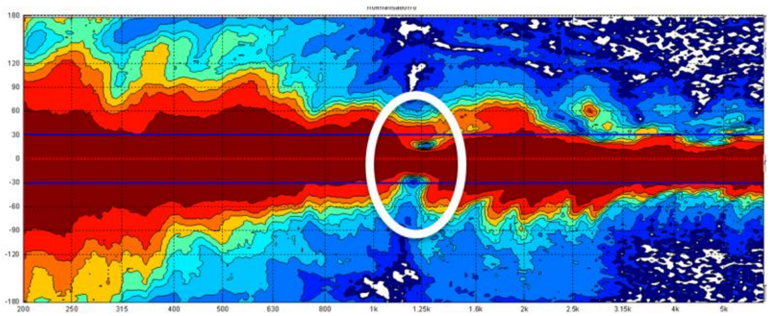

Figure 17 : effect of steep crossover on coverage pattern

As well, the distortion plot (time ordinate, freq abscissa Figure 18 and Figure 19) shows clearly that the low order filtering have a significant rub and buzz, higher harmonic distortion and even subharmonics (due to cone folding at high SPL)

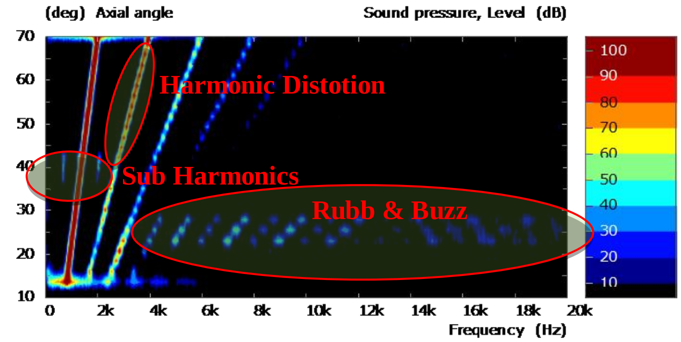

Figure 18 : frequency time, distortion, low order crossover

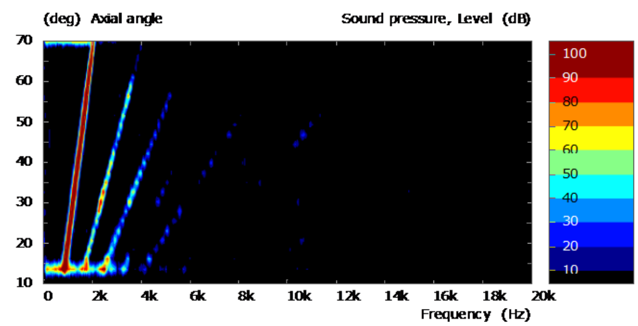

Figure 19 : frequency time, distortion, steep crossover

6.2. Defining LF CrossOver

As we have seen, FIR filters require lot of DSP resources when trying to reach lower frequencies. Straight forward FIR of 1024 taps (@48kHz the impulse response is truncated to 213ms) allows a good control down to about 250Hz (231ms ~= 5 periods @ 250Hz). Under this value it will be difficult to implement a filter without downsampling (with all anti alias and imaging filter required in the down sampling operation).

For this reason all LF Crossovers under 250Hz have been made with standard recursive filtering.

As we said before, the need to have magnitude compatibility on all sub basses of the range has led us to standardise the cut-off frequency as well as the shape. Four frequencies where chosen with a logarithmic step 60Hz 85Hz 120Hz 180Hz.

The shape of the filter was chosen so that the summation of two systems at the same frequency would give a flat response. Acoustical order has been set to 8th@180Hz & 120Hz, 6th@60Hz & 85Hz. Please note we are not referring to electrical orders (such as Linkwitz-Riley 8th order for instance); in our case we are also including the acoustical filter order (different from a reflex enclosure than a band pass for instance).

The result is shown on Figure 20, all sub-basses and main cabinet are now sharing the same low end pattern.

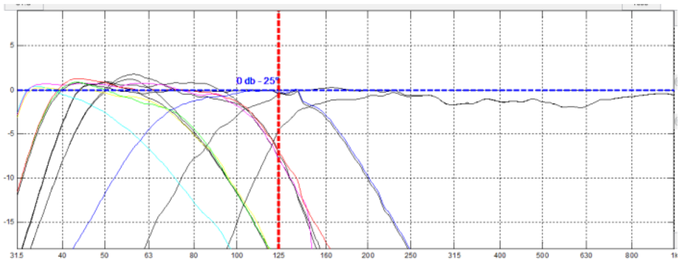

Figure 20 : LF crossover template

It allows us to get the same summation independently of the cabinets (Figure 21). The LF extension and SPL capability depends of course on the cabinet.

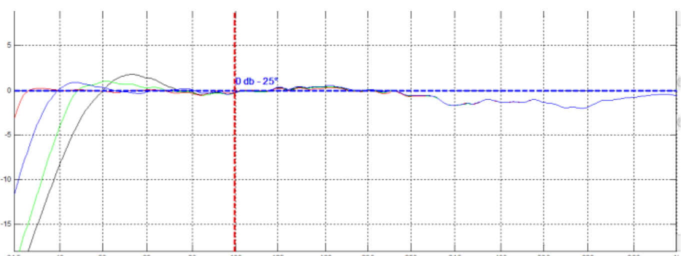

Figure 21 : Main and sub bass summations

The time alignment is done with a mix of pure delay and all pass filters on the subs and all pass filters only on the main cabinet (pure delay is impossible as it would spoil the time alignment up to 20kHz with other systems).

6.3. Defining FIR Target

Here is a description of the (homemade) tool used to generate FIR coefficients.

The reference curve computed in 5.6 is inverted and will act on the basis of the FIR target. According to the standard deviation, we know that the some zones shall not been blindly computed. We are defining three levels:

Zone with a low standard deviation: It is possible to EQ with the maximum of accuracy. In that case the filter target is the inverse of the reference curve.

Zone with a medium standard deviation (such as passive crossover): We will smooth this zone heavily to avoid EQing details which won't be consistent across all cabinets

Zone with a high standard deviation (typically upper HF of the compression HF drivers, or effect of vents on the polar coverage): EQing with FIR won't be possible due to the inconsistency of the reference. In those zones we will use a manual target (Typical analytic parametric EQ or modelling information)

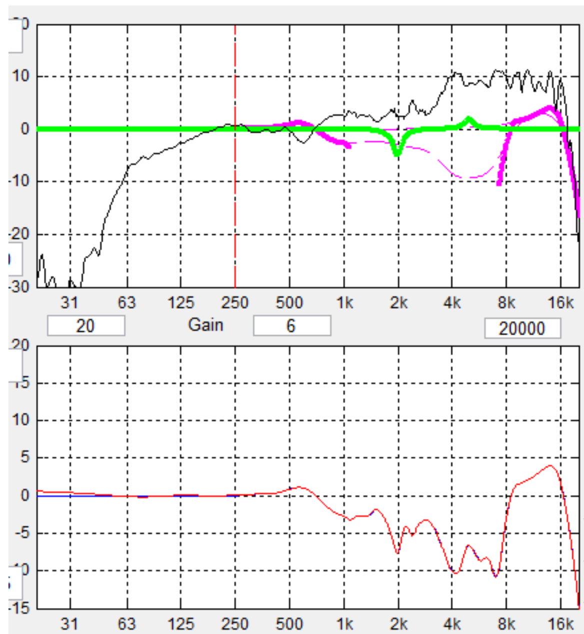

Figure 22 : FIR target

Figure 22 : The filter target is a mix of the inverted reference curve above 250Hz with zone A smoothed (250-1000Hz -magenta), with a low pass filter 15000Hz Q=1.5 above 8000Hz (magenta) and additional user EQ (in green)

Once the cabinet is EQed by the R&D, it is passed through a subjective EQ step in several real life applications (listening room, theatre, outside…). The result of those listening sessions is then re-injected into the routine; only the magnitude is taken into account so the phase remains linear.

6.4. Latency Limits

We have seen earlier that, when computing our filter target response into the time domain to get the impulse response of our FIR filter, it sometimes occurs that the maximum of the impulse response is in the negative time. It is then necessary to delay the overall signal to be able to implement this filter. We consider a maximum of 5ms overall latency to be acceptable in live application[5] (and especially monitors). This value includes also AD/DA converter and audio-network latency.

As a result the pure latency found in our setups is 3.58ms above 500Hz

6.5. Hardware

The DSP platform used to implement our algorithm is the NXAMP, a 4 time 4000Watt amplifier fitted with a DSP board (based upon 2 dual-core Motorola 56371 capable of 720 MIPS).

The control block contains several sub-block that are detailed below Figure 23.

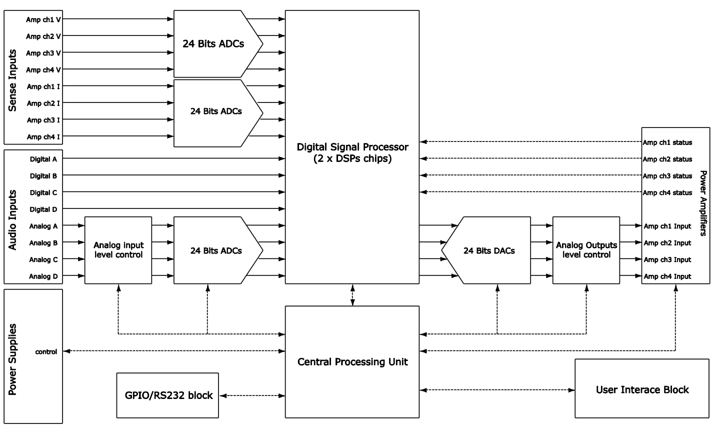

Figure 23 : NXAMP block diagram, Hard

The plain lines show the audio or sense signal (sense are voltage or current signal measured at the output of each amplifier). The dashed lines show the digital communication signal among several blocks.

All signals, audio or sense, use 24 bit converters. The CPU can also set up the analogue input and output gain for each channel, thus ensuring that the dynamic range of the system is always optimized (regarding volume, gain, patch and bridge settings).

Monitoring of the amplifier modules and power supplies (including multiple measurement such as temperature, voltages, current, integrate current …) are done both by the CPU and the DSPs.

The block diagram Figure 24 shows the global signal path inside the DSPs, for one channel (identical for all the channels). The algorithm described in this document takes place in “5”. The rest is mainly standard recursive EQ and modeling of the speakers to provide temperature, displacement protection[4].

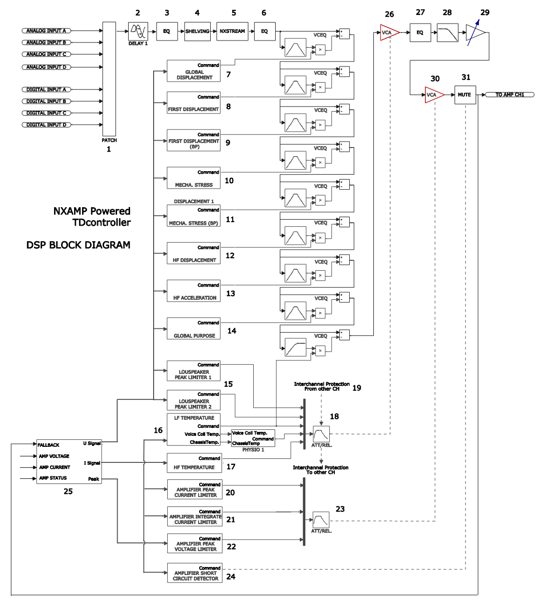

Figure 24 : NXAMP block diagram, Software

7. LISTENING TEST

7.1. FIR vs IIR

You won’t be surprised to hear that this new EQing method is sounding better than the traditional IIR EQing. The stereo image and “precision” were the two characteristics that the listeners described as an improvement. The listening sessions were made on controlled equipment with A/B instantaneous comparison on the same cabinets (both line arrays and stand alone speakers).

What was more of a surprise to us was finding that a FIR setup that was totally mocking up an IIR setup (amplitude and phase) was found as superior to its IIR counterpart during a blind test. It could be explained by the fact that on an IIR filtering the truncation of the 24*48bit data is occurring at each clock cycle. On the FIR, this truncation from 48bit accumulation to 24 bit data just occurs at the end of the computation. A THD measurement shows some difference between IIR and FIR Figure 25.

We can state that the improvement is mainly made by four factors:

- Better accuracy of the magnitude correction
- Linear phase above 500Hz
- Taken into account, production tolerances and coverage pattern
- Less digital artefact (fix point truncation)

7.2. Practical issue

Another improvement is the ability to mix different sub-basses easily. This has been well appreciated in the field, in the installation sector as well as in live sound.

7.3. Phase distortion

There has been considerable discussion in the industry about the audibility of phase variation. It is now globally admitted that a pure delay is not detectable, nor is a global polarity inversion (although some users claim to be able to hear that), but we were concerned with the use of all pass delay in order to align the systems in the low end where the linear phase was not an option for latency issues. We did conduct tests at several frequencies, increasing group delays in all pass filters, and found that a tolerance level, frequency dependant, existed.

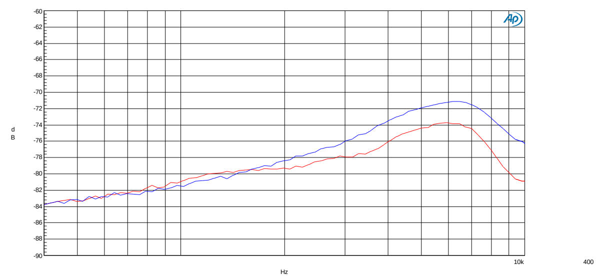

Figure 25 : THD+N, FIR filter versus IIR filter, same transfer function

In the low frequency range, all pass filters with reasonable Q (less than 1) and order (less than 4) are barely audible –if at all. On the other hand, they allow a perfect summation on all the bandwidths which would not be possible with only pure delay.

8. CONCLUSION

A new method of equalisation has been developed, the main advantage of which is to provide full frequency (20-20k) compatibility across an entire range of cabinets and a superior-sounding behaviour. This method relies mainly on the use FIR filters whose targets are optimised with a deep knowledge of the DUTs.

9. REFERENCES

[1] A. Oppenheim, R. Schafer: “Discrete time signal processing”, Prentice Hall, ISBN 0132167719 (1989)

[2] Steven W. Smith “The Scientist and Engineer's Guide to Digital Signal Processing”, second ed. California technical Publishing

[3] Rhonda Wilson, “Filter Topologies”, JAES Volume 41 Issue 9 pp. 667-678; September 1993

[4] NEXO, “NXAMP4x1 & NXAMP4x4 Powered TDcontrollers” User manual, 2012

[5] Carcopino, Deffarges, Ansade “Absolute and relative delays, phase shifts acceptable in modern PA systems”, internal NEXO-AUVITRAN, 2004