Monitors in the digital era

More manufacturers are adding digital monitors to their ranges but what does 'digital' mean in this context? PIERRE THOMAS from Fundamental Acoustic Research documents his company's progress into digital monitoring.

A MONITOR IS DESCRIBED as being digital when there is no analogue signal treatment in the crossover, filters (EQ, shelves) and compression systems. Going digital in the monitor makes a lot of sense as the majority of audio/visual content in the production process is now stored digitally and there is also the major benefit afforded by DSP which is paving the way to new horizons in the world of monitoring.

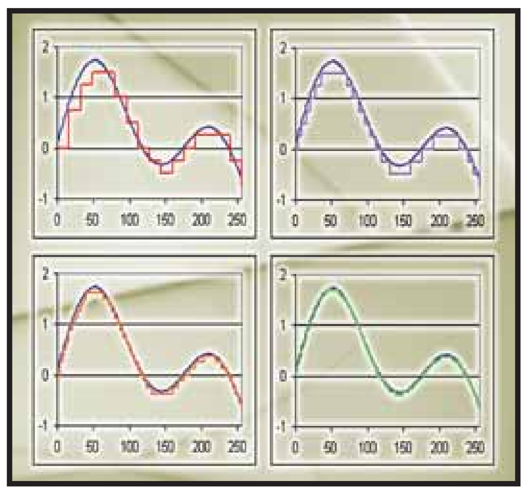
At FAR we used our experience of the market and in analogue systems to write a detailed Functional Requirements Specification for a new digital speaker range and this was taken up by our partners ATD² — a company that specialises in analogue and digital audio electronics. The genesis of a new speaker generation is not trivial, especially when the technological step is such a big one. The product had to be as easy to use as an analogue speaker, it had to be compatible with XLR analogue, AES-EBU and SPDIF standards, and it had to provide the user with more possibilities than analogue models. Additionally, it had to have better sonic quality and draw on FAR's expertise in acoustics.

The resulting digital board was designed to fit in all the models, and the specific speaker tuning effort became a matter of software instead of hardware, which speeds up considerably the final product release. We also wanted the speaker to be easy to configure so we implemented a user interface based on an alphanumerical display and a keypad instead of dipswitches.

Extra functionality was introduced with Curve selection (Academy, Car, TV, Club, etc), a 10-band parametric EQ with multiple presets, variable delay for time alignment in a 5.1 system, a network interface to control all the speakers from the console (via PC/MAC, PDA or a dedicated remote control), and room acoustic matching. For FAR, this was an opportunity to offer a constant and coherent listening experience to the user, whatever the control room used.

Much has been said about the sound quality of digital and the industry has made considerable advances in this area in the last 20 years. There are a number of parameters that affect the performance, such as errors that occur when writing or reading, the sampling frequency jitter, the sampling rate and resolution, the analogue filters (in front of the A-DC and after the D-AC), but there is no black magic.

There is nothing we can do about recorded errors but the sampling frequency jitter is a noise that affects the frequency stability and a very clean master clock design cures this. A digital signal is characterised by its resolution (in bits) and its sampling rate (in kHz). The higher the number of bits, the smaller the quantisation error — the approximation error to the analogue signal.



The top-left figure shows the original signal and the result of a (bad) analogue to digital conversion. In the top-right figure, the sampling frequency has been doubled but the resolution is kept unchanged. The result, even though not perfect, better matches the analogue signal. The sampling rate of the bottom-left figure is the same as the one of the top-right figure but the resolution has been doubled. Again it's an improvement.

These figures show that the sampling rate and the sampling resolution are equally important in conversion quality. At 24 bits, the signal can take about 16 million different values and this is where we are today. Assuming an output signal amplitude of $2V_{RMS}$, the smallest voltage step is only 0.17μV (almost two orders of magnitude below the noise floor of very good audio equipment). The theoretical dynamic range is 144.5dB! Today, the very best D-ACs can achieve about -120dB of dynamic range and THD+N, the major contributor being the noise. Standard D-AC integrated circuits can already achieve better than 100dB of THD+N and a dynamic range around 115dB. This is better than most analogue circuitry.

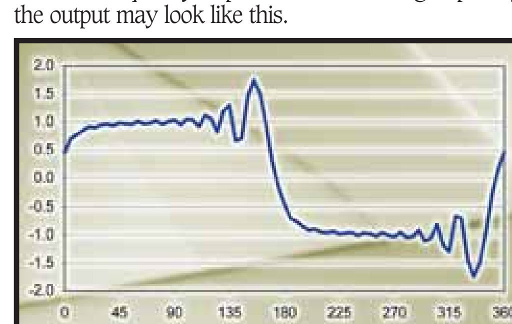
As for the resolution, the sampling rate does define how accurately the analogue signal is measured (sampled) but I want to mention a parameter that is not often discussed, which is group delay, and show how the sampling rate affects it.

Nyquist said that to properly represent an analogue signal by samples, the sampling rate (i.e. the rate at which the signal is measured) must be two times the maximum frequency contained in the signal. Thanks to Fourier, who showed long ago that any kind of periodic signal is made up of a sum of sine and cosine functions, we all think in terms of pure tones when dealing with numbers. Human hearing has been measured, with pure tones, to range from 20Hz to about 20kHz so it was concluded that a sampling rate of 44.1kHz should be good enough for music reproduction.

Music (and any kind of sound) is made of transients and these transients can only be properly played back if the group delay from the microphone to the speaker is kept constant. The group delay is the delay introduced by the complete system at a given frequency. If the group delay is not constant over frequency, the different frequencies contained in a complex signal will not arrive at the same time and the transient will be spread over time.

Here is a square wave modelled by its first 25 harmonics.

If this signal goes through a non-ideal channel that has a flat frequency response but a variable group delay, the output may look like this.



The harmonic content of the output signal is similar to the one of the input signal but obviously the output signal is clearly different from the input signal. In other words, 'Tak' will sound like 'Dag'. Not good at all!

Distortion and frequency response measurements on such a channel will never reveal a problem — single tone measurement cannot pinpoint these issues — and it shows the limitation of thinking in terms of pure tones. It is also probably why some people claim that vinyl has a better dynamic feeling than CD.

CDs recorded at 44.1 kHz employ a low pass brick wall filter to cut anything above 22.05kHz on the analogue to digital conversion. Unfortunately, these filters do not have linear phase behaviour and therefore the group delay is not constant anymore and usually shows an increase from 10kHz. What is recorded digitally is not what was 'heard' by the microphone and it cannot be fixed at playback. Indeed, the situation becomes even worse because there is usually a low pass filter at the output of the digital to analogue convertor to filter out the high frequency energy contained in the converted signal with more deterioration in the group delay.

Modern recording equipment uses 96kHz or even 192kHz sampling rates and the anti-aliasing or reconstruction filters do not introduce any noticeable degradation in the group delay over the audio band. And digital conversion from 96/192kHz to 44.1kHz does not degrade the group delay.

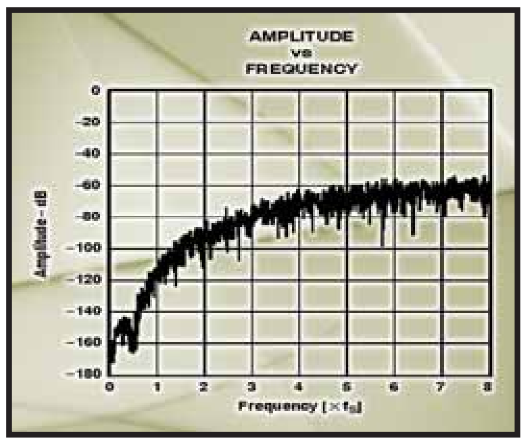
Given the complexity of the treatment chain, it's important to acknowledge the contribution of the sound engineer to a good recording. It's also important to understand that excellent designs using 48kHz/16-bit can sound much better than a poor 96kHz/24-bit implementation.

At FAR, we opted for a 96kHz/24-bit architecture. For the A-DC the analogue signal is first conditioned by a high quality programmable gain amplifier to be able to cope with any kind of signal level. We decided not to use a low pass filter at the A-DC input and assumed that the signal bandwidth would not exceed 48kHz. The sample rate convertor provides a very low jitter sampling clock and we selected a SRC from Texas Instrument because it is the only one that has a deterministic delay, whatever the incoming sampling rate is. It feeds the DSP block with 96 kHz/24-bit data and accepts sampling rates from 32kHz to 108kHz. Additionally, the SRC also filters out the jitter of the incoming digital signal (usually generated by the SPDIF or AES-EBU).

The D-AC comes from Burr-Brown and operates at 96kHz and a second order Bessel low pass filter (Fc at 50kHz) follows it. This filter does not degrade the group delay until the cutting frequency. The topology of the filter section is fully symmetric even though the output is unbalanced and this ensures extremely low noise and distortion.

In its datasheet, Burr-Brown recommends the use of a NE5532 and an OPA134 for the output operational amplifiers. During our listening tests, we realised that it was clearly not the best sounding solution. We finally took other parts from TI and ADI and it is interesting to note that the solution that sounds best is not the one that measures best. With the NE5532/OPA134, the THD+N was about 0.0007% (-103dB) at 0dBFs/1kHz while our eventual choice gave 0.0011% (-99.2dB). The THD with the NE5532/OPA134 was not flat over frequency, it started to rise at 4kHz, while with our eventual choice it was flat up to 30kHz.

A little bit of theory to try to explain why the NE5532/OPA134 did not perform well. In delta-sigma convertors, there is a noise shaper that cleans the audio band and rejects the noise above it. It increases very quickly with frequency and this is why it is always recommended to add a low pass filter after the convertor.



In the presence of high frequency noise, the semiconductors are subject to a phenomenon called 'input stage RFI rectification'. While amplifying small signals in the audio band, these devices can rectify unwanted high frequency signals. This results in DC error, which appears at the output in addition to the wanted signal.

Semiconductors are not all equally sensitive to this phenomenon as it also depends on their bias current. Generally, bipolar transistors are much more sensitive than FET transistors. The NE5532 is a bipolar operational amplifier. That's why I would recommend the use of FET operational amplifiers when dealing with delta-sigma convertors.

We decided to go for fixed point processing for the DSP. The DSP has an internal data path of 48 bits with 76 bit accumulators so rounding errors will never be a problem. It also permits a higher dynamic in the processing and there is no need for internal scaling before the signal exits the DSP.

We used Biquad IIR filters as the best compromise in terms of filtering capacity with respect to the required computational power. It is true that IIR filters are not linear phase filters but they can be minimum phase filters and they affect the group delay exactly as an analogue filter would. The Biquad IIR filters do not introduce more than two samples of computational delay.

As we target our speakers for use in broadcast studios, we considered a very large delay, like the sort introduced by FIR filters, to be unacceptable as it would have caused lip sync issues.

On top of this, having an ultra flat frequency response (±0.1dB) as can be provided by a FIR correction, does require an almost perfect acoustic design prior to equalisation. Indeed, it may be dangerous for the drivers to be compensated for for any faulty resonance. The driver will be heavily stressed and the sound quality will seriously degrade.

With IIR filters, we only compensate for local bumps in the frequency response. These bumps are usually generated by the electromechanical filter behaviour of the 'driver + enclosure' system. These filters are always minimum phase so it is easy and natural to compensate them by another minimum phase filter. No stress and a good sound.

In our system, the delay caused by the digital section (DSP, A-DC, D-AC, SRC) is about 70 samples, which means less than 0.73mS.

As a result of the processing, each loudspeaker can be optimised individually taking into account its position in the room. FAR measurement software provides all the information needed to calibrate the speakers individually and to alert the user to major acoustic problems. In such cases we will be able to suggest an acoustic solution.

All systems can be configured with the display panel at the rear of the speaker but it is much easier to connect the speaker to a PC and increase the accuracy of the settings and to be able to download corrections tailored to the user.

The quality of our loudspeakers, boxes and manufacturing process plus the digital expertise of ATD² has created what we believe is a total electro-acoustical solution for audio professionals. ■