Proceedings of The Institute of Acoustics

ANALOGUE FILTERING IN DIGITAL AUDIO SYSTEMS

K.V.Castor-Perry

Kemo Ltd.,12 Goodwood Parade, Elmers End, Beckenham,Kent.

What is 'digital audio'?

Digital audio is a general term bestowed on digital systems designed for the transmission and processing of high quality sound. Digital systems are now widely used for the processing and storage of programme material especially since the arrival of Compact Disc. Since the majority of the sources which are combined to produce a soundtrack or record are analogue, they must be converted into a stream of digital words; it is here that the requirement for analogue filtering first appears. In addition, after processing and storage, the digital information must be converted back to analogue in order to interface with current reproduction devices. Here too there is a requirement for analogue filtering.

Digitising an analogue signal.

The audio signal is represented in the analogue regime as a continuous function whose value is relevant at every instant. The digital regime in contrast is 'quantised' in both dimensions. This is not a fundamental restriction but stems from engineering limitations. As the resolution of the measurement of voltage increases, the number of possible states in the digital word has to go up, which requires more 'bits' per word. Also as the resolution increases, more readings per second have to be made. Having made the choice of word length (quantisation of amplitude) and sample rate (quantisation of sequence) the digital system is now taking defined-precision 'snapshots' of the signal, at regular intervals. Do these snapshots define the signal uniquely? The answer, unfortunately, is NO. It can be shown that the set of values generated by sampling a sinewave at regular intervals can be fitted not only by that sinewave but by an infinity of other sinewaves; these 'impostors' are called aliases (for obvious reasons); their existence and frequencies can be deduced from the following identity:

$$\sin(2*pi*f*t) = \sin(2*pi*(f+m*Fs/2)*t) \ : \ t=n/Fs$$

where f is signal frequency and Fs is the sampling frequency

This implies that there exists an infinity of frequencies which will produce the same set of snapshots as the wanted frequency. It can be shown that they will all come out at the other end at the lowest, originally wanted frequency, which is in the desired band and cannot be filtered out. By putting m=1 in the above identity, we can set an upper bound to the allowed frequency spectrum; frequencies above half the sampling frequency must not be allowed in. They must be filtered out, and the filter designed to do this is called an anti-aliasing filter.

ANALOGUE FILTERING IN DIGITAL AUDIO SYSTEMS (cont)

Analogising a digital signal

Reconstructing the digital snapshots is like projecting a cine film. Each is displayed for a moment, and the succession of shots has an effect similar to viewing the original continuous process. However, just as a projector causes flicker on the picture, the regular way in which the snapshots are presented makes itself felt in the output signal. The interaction between the frequencies in the signal and the repetition rate of the samples shows up as reflections of the desired frequency spectrum around discrete multiples of the sampling frequency. Technically this is because the output analogue signal is the convolution of the desired signal and a sample pulse which has a continuous spectrum extending to infinity. It is generally deemed necessary to eliminate the high frequency signals; they are not in themselves audible but can lead to difficulties in the subsequent reproduction equipment. It is also difficuly to measure the system THD on conventional equipment if they are not removed down to the required distortion threshold. The filter that performs this function is termed an anti-imaging or reconstruction filter.

Requirements for filters in digital audio

The function of an anti-aliasing filter is to strip off those parts of the incoming signal which could 'alias' down to the audio frequency band. Viewed from this standpoint alone, the ideal filter for this job would be the so-called 'brick wall' filter. This passes all frequencies up to a certain limit, and rejects completely all those frequencies above it. If this filter were possible, the system bandwidth could be extended right up to half the sampling frequency.

In practice we must compromise and set a bandwidth of less than half the sample rate. This allows a transition region for the filter between passing and blocking modes. The ratio between the passband and stopband frequencies in a lowpass filter is called the discrimination; the higher it is, the more difficult is the filter. Assuming the sample rate and the word length to be known, the following choices must be made:

1 Maximum frequency of interest, i.e.system bandwidth.
2 How much the through gain of the system can vary in this range.
3 By how much the unwanted band needs to be attenuated.
4 If this attenuation should be constant, or are there any relaxations or further restrictions to it.

This lays down the basic amplitude specification for the required filter. In addition the signal handling capabilities must be addressed:

5 Maximum signal level required; constant or frequency-dependant.
6 Maximum acceptable non-linearity, ditto as above.
7 Maximum level of intrinsic noise, and is the spectrum important

ANALOGUE FILTERING IN DIGITAL AUDIO SYSTEMS (cont)

Once the anti-alias filter has been specified, the spectrum of the signal is known. The spectrum of the 'images' in the system can now be calculated, and the whole of the above procedure repeated with new input for steps 3 and 4 above. As a general rule, the antiimaging filter is easier than the antialiasing filter. This is because the lowest frequency image is equal to the sampling frequency minus the highest signal frequency; since that in turn is less than half the sample frequency it follows that the worst case image frequency is higher than this, and thus the discrimination is lower than that of the anti-alias filter.

Typical values

As representative of a typical professional digital audio setup, the following numerical values will be assumed for illustrative purposes:

> Sample rate 48kHz; desired bandwidth 20kHz (discrimination .8333)
> Word length 16 bits; alias rejection 72dB (.025%)
> System noise 2LSB pk-pk, 1LSB rms; image rejection 66dB(.05%)

Choice of filter functions to meet these requirements

One of the conventional goals of audio has always been flatness of amplitude response. The filter of choice will be therefore a member of the class of filters which are approximations to the brickwall filter. Most of the popular filter functions fall into this category, and thence into three subcategories:

- i Maximally flat in the passband
- ii Equiripple in the passband
- iii Conforming to some other error norm in the passband

In addition a distinction is made between filters with and without transmission zeroes in the stopband; the former are called rational function filters, the latter, all-pole filters. The following table shows the most popular members of these categories.

Passband is:	all-pole:	rational function:
Maximally flat	Butterworth	Inverse Chebychev
Equiripple	Chebychev	Cauer (or Elliptic)
other norm:	Legendre	least-square monotonic

(in both given 'other norm' cases the function is obtained by maximising the second derivative of the filter function)

The high discrimination required of the filters in a digital audio system means in general that the filters required will be rational function ones. The transmission zeroes in the stopband 'pin' it down, with a great potential increase in sharpness. However the stopband does 'bounce' up between the zeroes, resulting in a defined finite attenuation at most points in the stopband.

This is in contrast with all-pole filters, whose stopband response (in theory at least) falls off continuously with a slope determined by the number of filter poles.

It is always possible to meet the attenuation requirements with any of the above filters, but the degree of the function, that is, the number of poles of the polynomial expressing the shape of the response, can become excessive. This in turn exacerbates certain engineering problems.

Usually chosen for the task is the rational function equiripple filter; this filter has the highest discrimination of them all. More importantly though is that exhaustive information is available for a special case of this filter, the Cauer filter (1,2). Cauer filters of up to 15 poles can be designed from these tables with no more than a pocket calculator. This has endeared them to a whole generation of engineers.

It is not necessary in every case however to utilise the maximum number of stopband zeroes if a smaller number will do. This choice is facilitated by using a more general method of filter specification with the assistance of a small computer (3,4).

It was mentioned earlier that the reconstituted output from such a system is the convolution of the desired signal and the continuous spectrum of the sample pulse. If the sample pulse is very narrow (in the limit, a delta function) so that the snapshot is only shown briefly, the output spectrum in the wanted band has the correct shape. If the pulse is of finite width then a roll-off of the frequency spectrum ensues, following a sin (x)/x function. This may be taken into account in the filter function, as mentioned later.

In addition, there is often a requirement to consider the phase or delay performance of the system. The variation of the delay suffered by different signal components as they pass through the filter gets worse as the discrimination of the filter increases, for all of the standard filter designs. Additional sections which contribute to the delay but not the amplitude response of the system may be required. These are usually found by computer optimisation methods (5,6) rather than from a catalogue.

The designed may choose to use computer optimisation methods from scratch, rather than catalogue techniques; in this case the response may be fitted using some other criterion such as the r.m.s. error of the response from the desired value; such methods are flexible but tend not to converge to such aesthetically satisfying solutions as the general parameter method. They can be extended to give a simultaneous constrained optimisation of phase and amplitude, thus minimising (in a total power or equiripple sense) the difference between the filter being designed and a delayed version of the ideal filter.

ANALOGUE FILTERING IN DIGITAL AUDIO SYSTEMS (cont)

Approaches to the filter design problem

1. Filters without response correction

The anti-aliasing filter will probably be specified as having a flat frequency response. The filter will in many cases be a conventional Cauer design derived from one of the classical texts. These texts were written with the assumption that the filter would be realised as a passive LC network, usually operating between equal resistances for optimum sensitivity. This is still a valid way of making such filters today, and such passive filters excel in several respects, one being generation of noise (8). As there are no amplifiers in a passive filter, the noise generated is due to the resistive elements only. However it is fair to say that the quality of amplifiers available today has made the passive LC filter a poor choice for most of today's cost and size sensitive applications.

All other current approaches utilise capacitors, amplifiers and resistors to generate the required filter function. Attempts to realise the function fall generally into one of three categories:

- i Generating a network of components whose properties in some way mimic that of a passive LC network;
- ii Breaking the function down into quadratic portions, and realising a series connected string of these;
- iii Generating a complex network of interconnected quadratic sections with multiple feedback loops.

Much work has been done on the first category. The consensus here is that the active filter obtained by applying the Bruton transformation to a minimum-capacitor LC structure is the topology of choice (7). This is a very well-researched filter topology (8,9,10) which retains to a large degree the insensitivity to component variation of the passive ladder network from which it is derived. However it is one of the aims of this paper to point out that the topology is not without its problems.

The basis of the Bruton transformation is that all components in the network concerned are given an extra order of frequency dependance. Inductors become resistors, capacitors become so-called FDNRs and resistors become capacitors.

A buffer amplifier looking at the output of a passive LC filter sees as its source impedance the parallel value of load resistor and the component of source resistance transmitted through the filter. At low frequencies this is purely resistive. After application of the Bruton transform this impedance becomes capacitative, and hence not able to provide a bias current for the amplifier. This means that a resistor has to be tacked across the load capacitor to provide bias current for the amplifier. If the high-pass pole formed with the total capacitance leads to an unacceptable roll-off in the response, another resistor has to be placed in parallel with the source capacitance. This has two effects:

ANALOGUE FILTERING IN DIGITAL AUDIO SYSTEMS (cont)

It modifies the transfer function, causing a phase 'woodle' at the transition frequency defined by the capacitor and resistor. This can be corrected with difficulty (11).

It causes the impedance presented to the amplifier (and any other amplifier monitoring this point) to rise with decreasing frequency until it reaches a value set by the added resistor.

The first effect makes specifying the delay linearity of the filter rather difficult. The second effect has two major consequences:

1. The noise of the system increases at the low frequency end, because of the rising value of the real part of the impedance presented to the outside world. This effect is exacerbated if the amplifiers monitoring this point have a significant noise current contribution.

2. The non-linearity of the system increases at the low frequency end. Due to the modulation of input bias current with common-mode input signals, a distortion voltage is developed on the source impedance of any amplifier monitoring this point, which is not removed by any form of feedback. In one particular practical case this form of distortion was so bad that a 0.05% distortion criterion could not be met at 5V, 20Hz. At 1kHz the distortion was 20 times better!

The choice of impedance level and amplifier type is critical for correct performance of this circuit. Previous work has suggested that the 5534 style of amplifier works well in this circuit. This device is relatively free of the second effect shown above, but if impedance levels are chosen to give realistic freedom from current overload the low frequency noise contribution is dominant.

This is why a filter made with 5534's is only about 4-6dB quieter than a filter made with TL071's, although the latter devices have 13dB more voltage noise. In the band below 1kHz the 5534 filter is up to 6dB noiser than the filter made with BiFET devices, due to their very low noise current. However, the TL071 series (and all commodity BiFET devices) suffer from bad input current modulation, making it less easy to acheive acceptable linearity.

Another problem with the FDNR circuit is the internal dynamic range. The outputs of the amplifiers inside the filter can have quite sharp peaks near the cut-off frequency, and there is little that can be done about this. It is possible to optimise the ratio of the components in the individual D elements, but as a general rule, the signal level going into the filter has to be held at about 8dB below output level in order that full signal levels can be obtained up to the cut-off frequency without internal distortion.

The second discipline of realisation, the cascade method, is the old-fashioned way of generating high-order filters. Much work has been done to avoid having to use cascade techniques, hence the predominance of ladder derived filters. The third technique is an extension of the cascade technique where controlled amounts of feedback are applied between the blocks. Conclusions derived for the cascade techniques apply broadly to the multiple loop designs too.

There is a great flexibility in the cascade approach to filter realisation which comes from the multitude of ways of constructing second order sections, and the many ways of ordering them in any particular design. It cannot be denied that the sensitivity to component errors is higher than for the ladder-derived filters; in practice this does not present a major problem. The two problems mentioned for the FDNR filter do not occur in cascade designs; at the frequencies involved in digital audio, the impedances in the filter blocks are usually commensurate with the m.f. noise impedance of the amplifier. Also, for any combination of pole frequency and Q a filter block can be found which gives the optimum signal to noise ratio. By using generalised design techniques to investigate the effect of using different numbers of poles and zeroes, an optimum solution can be found with a higher performance than the standard Cauer design.

2. Filters with response correction

It is, as has been said, often appropriate to incorporate some frequency response shaping into some filters in digital audio systems, for instance to remove the sinx/x 'droop'. In some systems it is appropriate to apply these corrections in the form of separate circuit blocks. However , incorporating the desired response into the filter has two advantages:

- i it does not generally increase the complexity and hence saves on overall system component count
- ii because of the high order nature of the filter function, the approximation to the desired function can be high order hence accurate.

The desire to modify the filter response to accomodate the various losses leads to a great shift in emphasis in the choice of topology. Unless the desired response correction is mild enough to be acheived by appending a few first order shelf sections to the basic filter (8), the FDNR design is not suitable. This is because the inherent low sensitivity of the power-matched equiripple passband filter prohibits the gross changes in response required.

The topology of choice in this case is the cascade filter. Using either a generalised curve-fitting program, or a closed technique (2) in special cases, a polynomial can be generated which fulfils the specification. Since any polynomial of the form which will be produced can be realised as a cascade of filter sections, the realiseability of the device is guaranteed. This is not always the case with the more complicated disciplines.

It has already been mentioned that the flexibility of cascade design comes from the ease with which the topology and filter function may be simultaneously adjusted to best meet the requirements. Experiments have been carried out to develop a reliable means of pinpointing the combination which lead to the lowest overall noise. It is a discrete optimisation problem of considerable size, and the work is still in progress. Initial results seem to indicate that rational function filters with less than the maximum number of stopband zeroes perform best. Additionally, the technique of multiple critical poles (3) can reduce the noise quite dramatically by eliminating the magnification due to the high Q's of the critical pole in conventional filters. Another route being investigated is the use of catalogue FDNR filters with complementary equalisation at input and output to optimise internal dynamics.

Examples of performance

1. Antialias for a 48kHz system. A ninth order catalogue Cauer filter was chosen for this job and built as a FDNR ladder. A two-pole allpass section was added to give phase linearity up to 15kHz, and the internal signal level set to 8dB below output level. With optimised component ratios in the FDNR's there is still a slight degree of excess internal gain at 20kHz, the end of the passband. Fitted with bipolar devices the output noise is about 45μVrms, of which much is below 2kHz. With the FDNR's built with BiFETs, the noise is about 55μV, rising to 65μV if the output buffer is also BiFET.

2. Antiimage for the above system. The filter started life as an eighth order Cauer with the passband pushed out to 21kHz a simple search procedure was used to relocate the Q's to fit the measured data. In the course of this one zero was thrown out. The final filter was also given a two-pole allpass, configured this time for the best linearity through the whole system. The filter was built as a cascade, using Friend (4) sections all through except for the high Q section, which as a variant on the basic TT Biquad (5). With optimum pole ordering, the output noise is around 35μVrms. When it is considered that the filter has a 4.5dB boost at 20kHz, it should be apparent that the cascade technique is more than adequate for low noise purposes.

Audibility of filters

There are two sides to this problem which can be looked at separately:

- i Is the filter FUNCTION intrisically audible, and if so, is it degrading the sound or just making it different,
- ii Is the filter CIRCUIT audible; do different circuits which realise the same function sound different?

If a certain filter function is deemed to be necessary for the operation of a digital audio system, there is little the filter designed can do to avoid the function from giving its signature to the system response. The non-linearity of the phase response can be corrected for to an extent, although noise and distortion are likely to suffer. It should be remembered that the average vinyl playing system will be at least eigth order at the top; with a VHF tuner, even more.

The second factor is more pernicious, but easier to get to grips with. Especially when the highest audio quality is being aimed at, the filter must be designed with the same attention to detail as a state-of-the-art amplifier system. A multitude of effects such as correct earth management, supply sensitivity, thermal distortion, device parameter modulation by the signal, must be taken into consideration.

It is now accepted that some early domestic digital audio devices had sound quality problems due to inadequate analogue engineering; it is the authors opinion that the funger has all too often been pointed at the filter function when in fact the realisation was at fault. In this context the author also beleives that the problems inherent in the FDNR type of filter are partly responsible for the criticism in some quarters of a 'disconnected' upper bass quality to some digital material.

Conclusions

In this paper I have attempted to illustrate the role that analogue filtering performs in digital audio systems. I have also shown some of the techniques available to the filter designer to generate the types of filters needed, and to point out some of the pitfalls for the unwary. Figures for what is currently acheived in commercial practice have been given, along with an indication of the direction in which this particular discipline is moving. The topic of digital filtering has not been covered; although to a degree it overlaps with analogue filtering, it is more properly viewed in a systems context with the rest of the d.s.p. in the digital audio environment. The problems involved, and the techniques used to solve them, are rather different from those presently under discussion here.

References:

1. Handbook of filter synthesis — A.Zverev (J.Wiley)
2. Filter design tables & graphs — E.Christiansen, E.Eisenmann (J.Wiley)
3. The filter approximation problem — R.W.Daniels ProcIEEE ISCT 5/73
4. Approximation methods for electronic filter design — R.W.Daniels (McGraw Hill)
5. The analysis and synthesis of electrical filters — DeV.S.Humphreys (Prentice Hall)
6. A digital computer program for the design of phase correctors — J.V.Fall IRE TransCT 9/61

ANALOGUE FILTERING IN DIGITAL AUDIO SYSTEMS (cont)

7. Network transfer functions using the concept of F.D.N.R.'s
L.T.Bruton IEEE TransCT 8/69
8. High quality analog filters for digital audio
R.Lagadec, D.Weiss, R.Greutmann 67th. Conv AES
9. Comparison of overload and noise characteristics of different F.D.N.R. circuit realistations P.Skritek IEEE 1982
10. Vergleich verschiedener Tiefpassfilter-Realisierung fuer hochwertige PCM-Audiosystem P.Skritek Frequenz 38(1984)1
11. Exact synthesis of active LP FDNR filters
C. Nightingale , J.M.Rollet Elect.Lett 2/74
12. Simple generation of high-order filters with sinx/x correction
K.V.Castor-Perry (unpublished)
13. Multiple pole and zero approximation of general attenuation specs.
A.Premoli IEEE TransCAS 12/80
14. STAR: An active biquadratic filter section
J.J.Friend , C.A.Harris, D.Hilberman IEEE TransCAS 2/75
15. The Biquad Parts I & II L.C.Thomas IEEE TransCT 5/71