Audio Engineering Society

Convention Paper 5613

Presented at the 112th Convention
2002 May 10–13 Munich, Germany

Importance of Precision on Performance for Digital Audio Filters

David Zaucha

Texas Instruments, Dallas, Texas, USA

*Correspondence should be addressed to (*dzaucha@ti.com*)*

ABSTRACT

Digital audio systems are unlike conventional analog systems in which signals can be any value between a minimum to maximum and occur continuously in time. Digital audio systems use finite precision in representing signals and coefficients and in performing arithmetic operations. Consequently, system performance is determined by the precision that is used throughout the system. This paper discusses the factors that influence the performance of Infinite Impulse Response filters in high performance audio applications using fixed point arithmetic.

INTRODUCTION

Unlike conventional analog signal processing in which signals can be any value between a minimum to maximum and occur continuously in time - digital signal processing uses a finite precision in representing signals and coefficients and finite precision arithmetic in determining the response. To a large degree, the system performance is determined by the number of bits used to represent the signal, filter coefficients and perform the operations.

Naturally this brings up the question "***How many bits are enough*?**"

Digital filters are used in a variety of digital audio signal processing functions. A few of the more common uses of digital filters are shown in Table 1.

Application	Filter
Tone Controls	Bass and treble shelf filters
Graphic and Parametric Equalizers	Band gain and cut filters
3 D spatial effects and Head Related Transfer Functions	Phase shift, Band-gain and band-cut filters
Multi-channel Decoders	Sub woofer and Bass Management Filters. Small / Large loudspeaker filters

Table 1 Filter Applications

This paper gives an overview of the relationship of precision to performance for second order Infinite Impulse Response (IIR) filters. To illustrate the trade-offs, two digital filtering examples will be described: a parametric equalizer and a loudspeaker crossover.

IIR DIGITAL FILTERS

The primary construct that is used to describe digital filters is the difference equation. The difference equation describes the output in terms of past and present inputs and previous outputs. The present input and output is specified as a function of n. n-1 indicates the previous output. The difference equation to describe a second order filter is shown in Equation 1:

$$y(n) = b_0\, x(n) + b_1 x(n-1) + b_2 x(n-2) - a_1 y(n-1) - a_2\, y(n-2)$$

Equation 1 Second Order Difference Equation

In this equation, we see that the nth sample of the output y is a weighted sum of the inputs x (n), x (n-1), and x (n-2) plus the weighted sum of previous outputs y (n-1) and y (n-2). The weighting factors, or filter coefficients, for the system input x, are b_o, b_1, and b_2. The coefficients for the system output, y, are $-a_1$ and $-a_2$.

Another way of describing the system response is the system transfer function. The system transfer function is produced by taking the Z transform of the difference equation and reordering terms to form Equation 2.

$$H(z) = \frac{Y(z)}{X(z)} = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}}$$

Equation 2 Transfer Function

There are a number of ways that the difference equation can be implemented in a digital system. In this discussion we will focus upon the second order Direct Form I filter architecture which is shown in Figure 1. As will be discussed later, this implementation architecture has often been used, because it is less prone than many other structures to producing numeric overflow and noise from zero-input limit cycles. The Z^{-1} blocks shown in the figure are single sample delays. The multiplication of the y and x terms with the **a** and **b** coefficients are shown by the coefficient labels over the arrows.

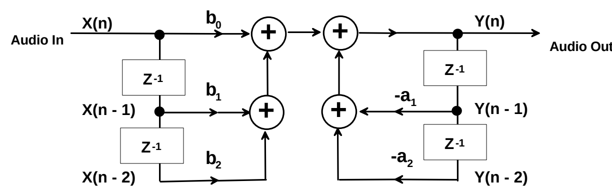


Figure 1 Direct Form I Second Order IIR Filter

Performance Factors

In digital processing, precision refers to the number of bits, or numeric resolution, that is used to represent the input signal, coefficients, intermediate calculations, and the resulting output. As will be discussed in more detail in the subsequent sections, precision drives filter performance in four ways:

– Coefficient Precision and Quantization Errors
– Round-off and Quantization Errors in Filter Calculations
– Overflow, Underflow, and Scaling
– The computation method

The consequences of increasing or decreasing the precision affect the filter performance in two ways:
– Response Accuracy
– Signal to Noise Ratio

Coefficient Quantization

Quantization is a process where a continuous value (a real number) is translated into a fixed precision representation. The difference between the original value and the quantized value is the quantization error.

Figure 2 illustrates the numerical effects of quantization upon a value. In this example a value on the x-axis is quantized, using magnitude truncation, into one of 8 levels that are shown on the y-axis. Using additional precision to more accurately represent the value will reduce the quantization error.

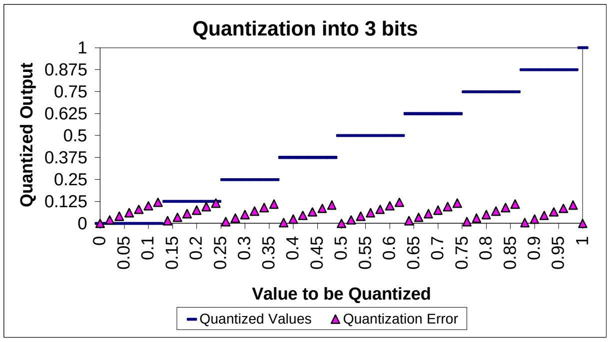


Figure 2 Quantization Effects

The effect of the quantization error is to perturb the filter coefficients from their ideal values. This produces a filter response that is different from the desired filter in amplitude, phase and/or frequency. The magnitude of this difference is related to quantization error and the sensitivity of the filter. The difference between the desired response and the response produced by coefficient quantization is the response error of the filter.

The following example illustrates the filter response deviation characteristics that are produced by coefficient quantization. The example is a 48 kHz sample rate parametric equalization filter that has a gain of 6 dB at 100 Hz with a Q of 6.7. The amplitude response of the filter is shown in Figure 3.

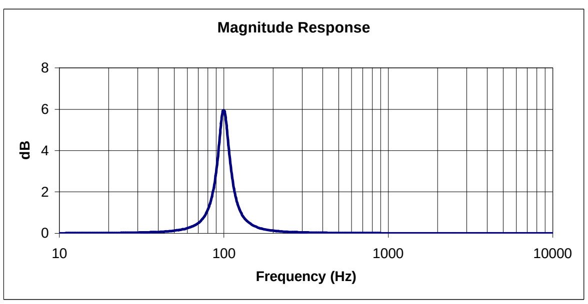


Figure 3 Example filter transfer function

Figure 4 illustrates the signed amplitude error that is produced using 24 bit quantization. The coefficients are encoded using a twos complement binary format with 4 bits assigned for the integer and 20 bits assigned for the fractional values. The positive to negative swing of the amplitude error indicates that a portion of the error is attributable to a shift in the frequency response.

As shown in Figure 5, the absolute value error amplitude can increase **substantially** as the number of bits used to represent the coefficients decreases. The amplitude response errors that are shown in the figure are not unusual cases. These are representative of the response errors produced by a number of filters with center frequencies near 100 Hz and Q's of 6. However, there can be considerable variability in the size of the error for other combinations of gain, center frequency and filter Q.

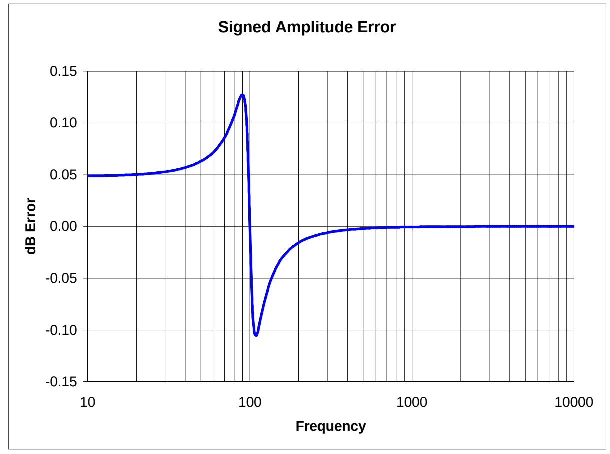


Figure 4 EQ filter at 100 Hz with 48 kHz Sample Rate

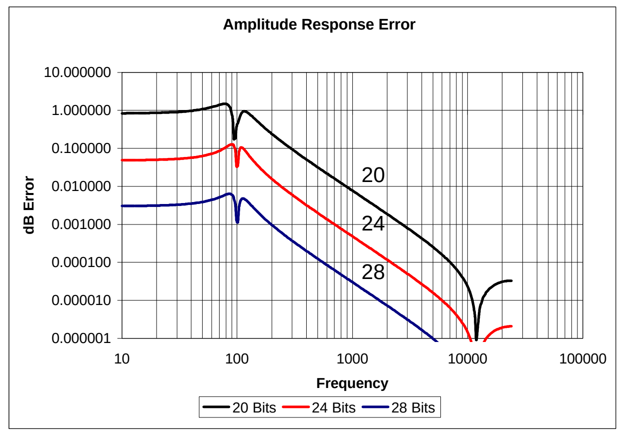


Figure 5 Response error for 48 kHz sample rate filter

Response errors that are produced by coefficient quantization can vary with respect to several factors. In general, the filter response error will increase when:

– Filter center frequency decreases
– Q of the filter increases
– Sample rate increases
– Filter gain increases

Figure 6 shows the substantial increase in the response error when the filter frequency decreases to 50 Hz. In this example, the magnitude of the filter response error produced by 20 bit coefficient quantization **is greater** than the 6 dB gain of the filter. As illustrated in these examples, 20 bit coefficients at 48 kHz sample rates can produce substantial filter response errors at low frequencies. Similarly, as the sample rate is increased, the response error due to quantization can increase substantially.

Figures 7 and 8 illustrate the limitations of using 24 bits to represent the coefficients at 96 and 192 kHz sample rates. In the 192 kHz example, the filter response error produced by 24 bit coefficient quantization **is greater** the than the 6 dB gain of the filter.

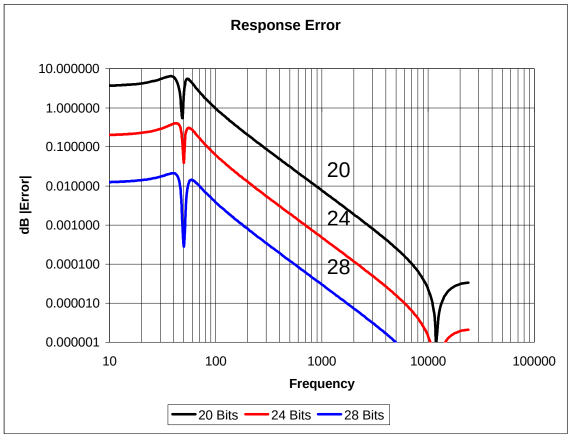


Figure 6 EQ filter at 50 Hz with 48 kHz Sample Rate

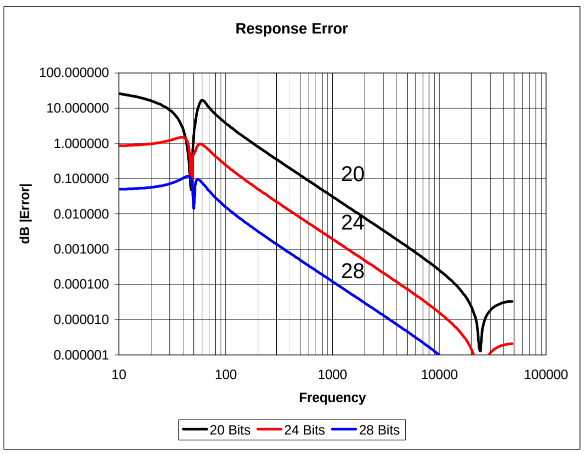


Figure 7 EQ filter at 50 Hz with 96 kHz Sample Rate

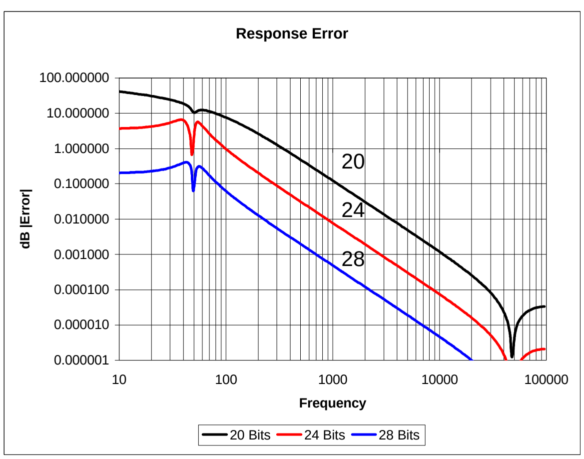


Figure 8 EQ filter at 50 Hz with 192 kHz Sample Rate

From these examples we can see that coefficient precision can significantly affect the filter accuracy. To achieve filter response errors that are less than 10% of the desired filter response at 96 and 196 kHz sample rates, coefficient sizes of 28 bits or larger are required.

Round-off and Quantization Errors in Filter Calculations

Quantization also impacts system performance when it occurs in the filter response calculation. Quantization occurs in the filter response due to the physical constraints of the processing architecture. An example of this is illustrated in the second order Direct Form 1 filter architecture shown in Figure 9.

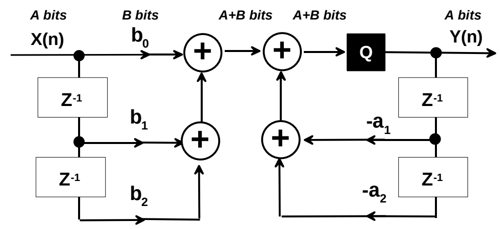


Figure 9 Implementation of a Digital Filter

In this figure, the input signal and the output signal are represented in A bits. The ***a*** and ***b*** coefficients are all represented in B bits. Registers of size B bits produce the input and output delayed signals. The multiplication of the input and output signals with the b and a coefficients is performed by a A bit by B bit multiplier that produces an A+B bit result. An A+B bit wide adder is then used to sum the products of the multiplier and compute the unquantized filter response.

In this example, quantization occurs at the output of the adder, the highlighted "Q" block, where the A+B bit result is quantized into A bits. This quantization step produces a quantization error.

The quantization error is the noise source of the digital filter. The noise characteristics are determined by the quantization error characteristics, error magnitude and type of quantization that is performed (round-off, twos complement, or signed magnitude truncation).

The noise output, f (n), from the filter is a function of the noise source, e (n) and the ***a*** coefficients of the filter as shown in Equation 3 and in Figure 10.

$$f(n) = e(n) - a_1 f(n-1) - a_2 f(n-2)$$

Equation 3 Filter Noise

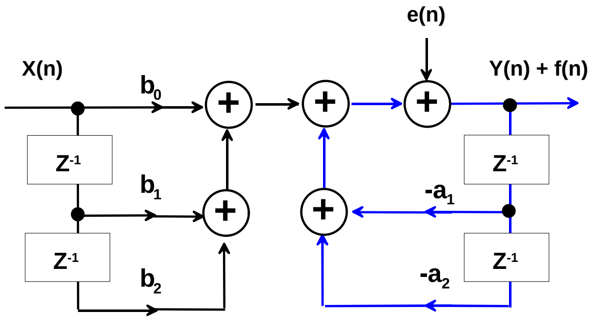


Figure 10 Digital Filter Noise

To preserve as much as possible of input signal SNR, we would like to have the noise amplitude below the smallest signal amplitude of interest, as shown in Figure 11.

Because the filter noise is proportional to the quantization error and the quantization error is reduced by increasing the calculation precision, the filter noise is reduced by increasing the calculation precision. This is achieved by increasing the precision of the signal representation above the number of bits that are absolutely necessary to represent the input signal. These additional bits are called noise bits.

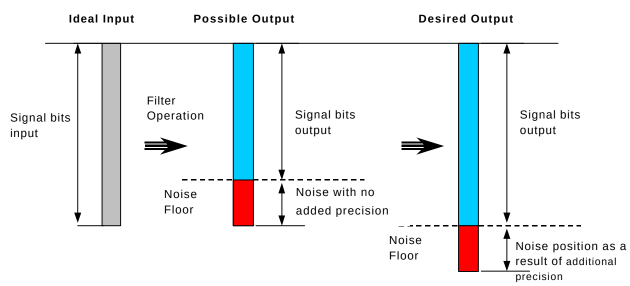


Figure 11 Impact of Additional Precision for Noise

How many noise bits are enough?

To illustrate the characteristics of truncation noise in IIR filters, a parametric equalization filter is evaluated at several frequencies. Figure 12 shows the Signal Transfer Function of a parametric equalization filter with a gain of 12 dB and a Q of 6.667 for a full-scale sine wave across the spectrum. To show the impact of frequency, the filter is evaluated at 1000 Hz, 100 Hz, and 30 Hz.

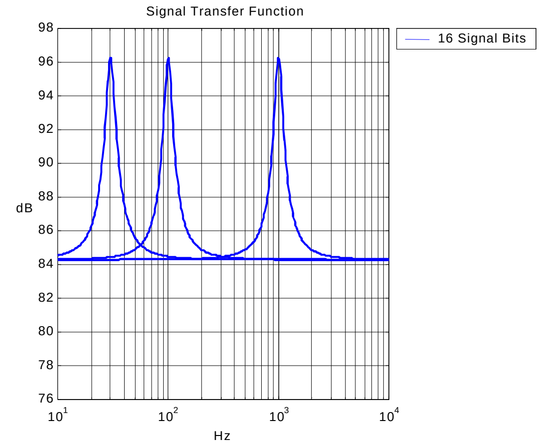


Figure 12 Equalization Filter Response for 16 bit 48 kHz Data

As we will see, in the case of parametric equalization filters, the magnitude of noise response increases when:

- Filter Q increases
- Sample rate increases
- Filter frequency decreases

The noise transfer function is used to show the amplitude of filter noise with respect to frequency. The noise transfer function is computed by collecting the terms of the noise difference equation shown in Equation 3 and taking the z-transform. The Noise Transfer Function (NTF) has the form:

$$NTF = \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2}};$$

Equation 4 Noise Transfer Function

For magnitude truncation, the characteritics of the noise source are:

- Uniform distribution over +/- 2^{-2b}, where b is the number of noise bits.
- The variance is: $s_q^2 = \frac{2^{-2b}}{\sqrt{3}}$

The variance at the filter output noise is:

$s_F^2 = s_q^2 \sum_{n=-\infty}^{\infty} h_q^2(n)$,where h_q are samples of the impulse response function of the NTF.

Figure 13 shows the Noise Transfer Function of the three parametric equalization filters using 8 and 16 noise bits.

The 0 dB line of the graph represents the lowest signal amplitude of interest, which is the lowest amplitude and the noise floor of the the ideal input signal. In this figure, we can see that 8 noise bits are sufficent to keep the 1000 Hz filter noise below the 0 dB threshold, but are not sufficient for the 100 or 30 Hz filters. Similarly, the 16 bits are sufficient to keep 1000 and 100 Hz filters below the 0 dB threshold. However, at 30 Hz, the noise cuts into the signal response by 11 dB.

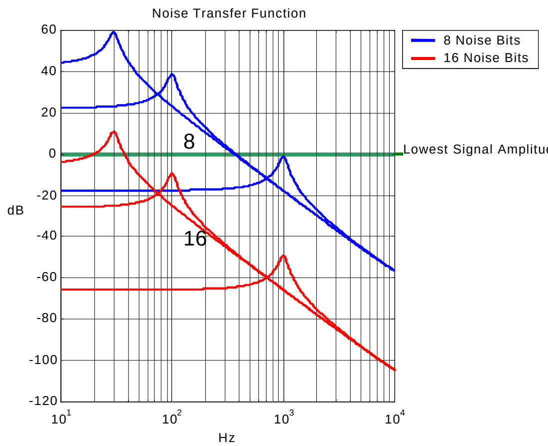
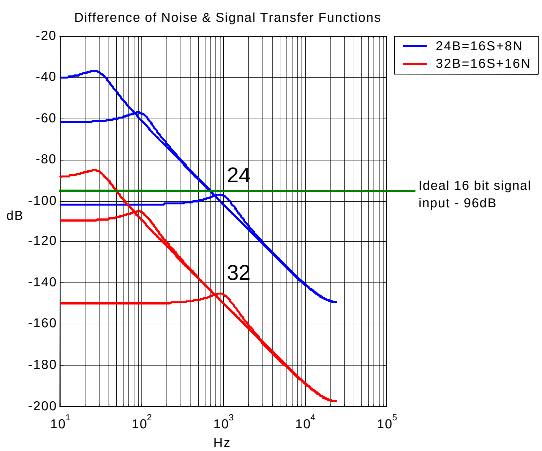


Figure 13 Equalization Filter Noise for 48 KHz Data

It is important to note that because the noise level is specified with respect to the number of noise bits – that Figure 13 is applicable for any signal precision.

Figure 14 shows the difference of the signal and noise transfer



functions for 16 bit data with 8 and 16 noise bits. This graph shows the signal-to-noise ratio of each filter at each frequency.

Figure 14 Difference Transfer Function for 16 bit 48 kHz Data

When the data precision is increased to 24 bits, the noise does not increase proportionately. It continues to occupy the same relative amplitude with respect to the minimum signal level as shown in Figure 15.

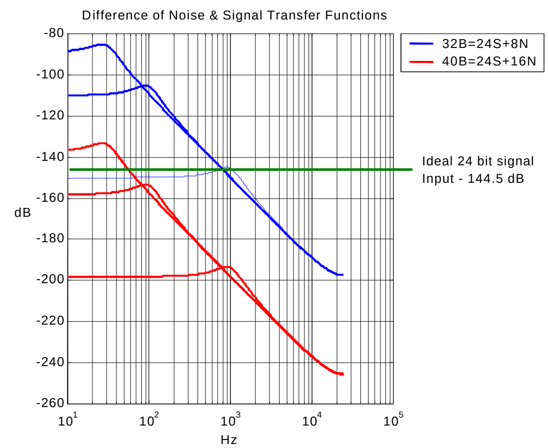


Figure 15 Difference Transfer Function for 24 bit 48 KHz data

Figure 16 and 17 show the increase in the noise amplitude when the sample rate is increased to 96KHz and 192 kHz.

These examples emphasize the importance of precision on digital performance.

For these particular filters, 16 or more noise bits are sufficient to preserve all of the input signal SNR for the 1000 and 100 Hz filters at sample rates of 48 kHz and 96 kHz. However, even with 16 noise bits there is some degradation of the input signal SNR at 30 Hz. To preserve all of the input SNR would take 20 noise bits. Similarly, for these filters at sample rates of 192 kHz, 20 noise bits are sufficient to preserve the input SNR for the 100 and 1000 Hz filters. At 192 kHz even with 20 bits, there is a some degradation of the input SNR at 30 Hz.

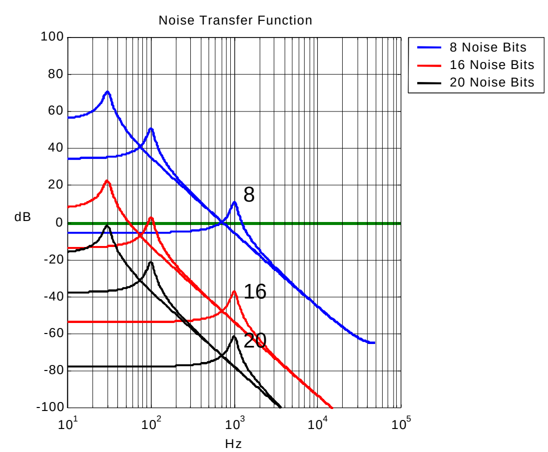


Figure 16 96 kHz Noise Transfer Function

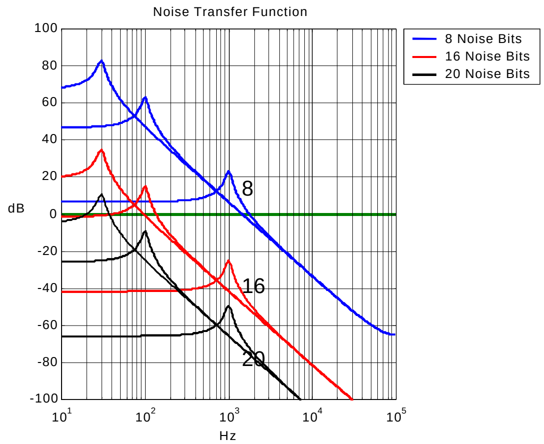


Figure 17 192 kHz Noise Transfer Function

It is important to note that while these examples are representative of equalization filters at these frequencies with Qs of 6.7, they do not represent the best or worst cases of signal or noise. Differences in filter parameters, center frequencies or filter types types will produce signal and noise characteristics that can be better or worse.

Zero Input Limit Cycles

The linear modeling that we have used for assessing the impact of quantization (or round-off) noise is sufficiently accurate for most analysis, except for one notable exception. This exception is the case of zero-input limit cycles, which are a non-linear phenomenon. Zero-input limit cycles produce periodic or "tone" components in the output in response to zero and small amplitude sinusoidal inputs. This behavior is best controlled by careful system design. The most effective means to prevent limit cycles relies upon using cascaded sections of second order Direct Form I filters and magnitude truncation quantization.

Overflow, Underflow, and Scaling

Overflow occurs during the calculation of the digital filter response when the result exceeds the largest number that can be represented. There are two principle instances were this occurs. Overflow can occur when the gain of one or more cascaded filters amplify the signal so that it exceeds the largest number that can be represented. An example of this case is when two cascaded filters are used to produce a twin peak response by subtraction, shown in Figure 18. If the positive gain filter precedes the negative gain filter, a full-scale signal input would be amplified by 20dB. This could produce an overflow condition.

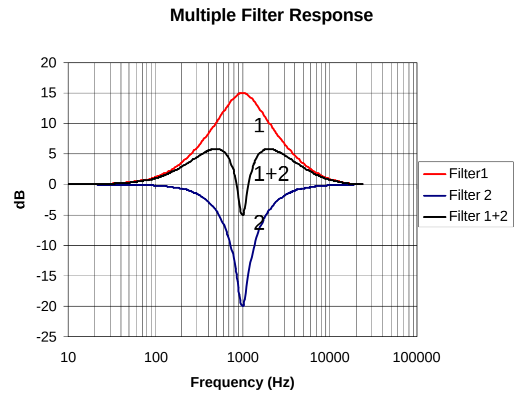


Figure 18 Summation of Filters

Overflow can also occur within the computation of a single filter stage for those filter types that have large coefficient values even though the filter has only a modest overall signal gain. In the second case, overflow occurs during the intermediate calculations as a result of the signal multiplication with one or more large coefficients.

There are several approaches that can be employed to prevent overflow:

Solution	Implementation Trade-off
In the case of cascaded filter gains – Reorder the order of the filters	Very useful approach. However, it is not sufficient to meet all cases
In the case of cascaded filter gains – Avoid the use of filters with positive gains	Using only negative gains can reduce the signal amplitude and compromise the SNR.
In the case of large coefficients – Avoid filters with large coefficient magnitudes or implement the response using two filters that have smaller coefficient magnitudes	This complicates the implementation, may not achieve the desired filter response, and consumes scarce processing resources
In both cases – Scale down the input signals	In the case of scaling with no increase in precision, the signal is shifted into the noise bit positions resulting in lowered SNR performance. I.e. we have fewer noise bits
In both cases – Add sufficient precision in the processing architecture to accommodate the calculations	No trade-off – simple implementation - good performance

The preferred solution is to have additional precision in the processing architecture to accommodate the range of expected range of amplitudes without degrading the SNR performance. This can be accomplished by either adding this additional precision as headroom bits to extend the internal maximum signal amplitude or by adding additional noise bits to reduce the noise floor. The difference between the two solutions is how frequently and when scaling is performed. Figure 19 shows how additional headroom bits permit filters with positive gains to be used to create a desired filter response.

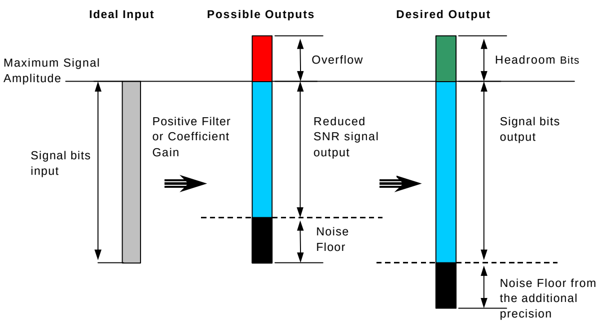


Figure 19 Importance of Added Precision for Overflow Conditions

The number of additional bits that are necessary to prevent overflow is dependent upon the permitted gains, filter types and filter parameters for a given sample rate. As previously discussed in the case of cascaded filters, intermediate gains of 18 to 24 dB are common. In the case of large coefficient values, although most filters have coefficient magnitudes that are between 0 and 2, there are a few frequently used audio filters that can have relatively large coefficients. The Treble Shelf is a commonly used audio filter that has relatively high coefficient magnitudes for relatively modest gains and frequencies. The transfer function of a 1 kHz Treble Shelf with a gain of 12 dB is shown in Figure 20.

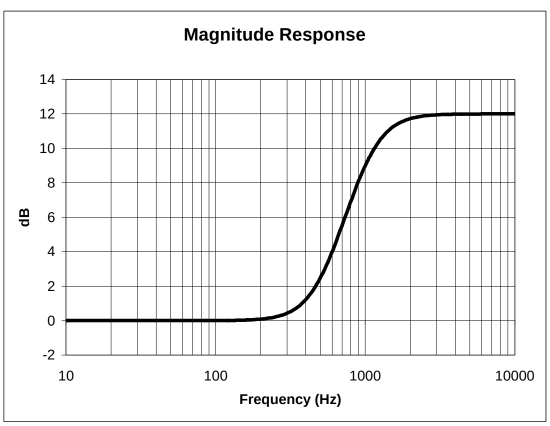


Figure 20 Treble Shelf Transfer Function

The coefficients for the Treble Shelf filter are:

b_0	3.795621	a_1	-1.80957
b_1	-7.226886	a_2	0.825762
b_2	3.44797		

From observation we can see that to accommodate the signal gain produced by the largest coefficient magnitude will require approximately 4 additional bits of precision. Although not shown here - subsequent dynamic analyses of the filter behavior indicate that one additional bit is required to prevent overflow for signals with high transient characteristics.

To accommodate both cascaded filters and large coefficient gains, 8 additional bits of precision appear to be sufficient to prevent overflow for most applications. These will be added to the most significant bit positions, as headroom bits. The advantage of adding headroom bits in comparison to noise bits is that the system is able to represent intermediate signal levels that are greater than the maximum input signal magnitude for typical cases. As a result, the headroom bits can eliminate the need to reduce the magnitude of input signal prior to filter processing, and then increase the magnitude of the result after filter processing, in many cases. The input to a filter or cascaded series of filters is reduced in exceptional cases where very large gains are used. The output signal magnitude is reduced when the total gain is greater than 1.

Underflow occurs when the result becomes so small that some signal information is irrecoverably lost. As in the case of overflow, this can occur for specific filters that have one or more small coefficients although the overall filter loss is modest. In these cases, the multiplication of the input and small coefficient values can produce an intermediate signal magnitude that is less than the smallest signal magnitude that can be represented with full input signal precision. There are two potential approaches that may be employed to prevent underflow:

Solution	Implementation Trade-off
The filter may be implemented in two filters that have larger coefficient values	This complicates the implementation and consumes additional processing resources
Have sufficient precision in the processing architecture to insure that the signal information is preserved	No trade-off – simple implementation - good performance

The preferred solution is to include sufficient precision in the processing architecture to insure that the minimum signal level can be preserved. The proposed solution to add 16 noise bits that is discussed in the section "Finite Precision Arithmetic in Filter Calculations" is sufficient to preserve SNR and prevent underflow for most applications. As we can see, the term "noise bits" is bit of a misnomer because these bits not only reduce the noise floor, but they also preserve signal information.

PRACTICAL APPLICATIONS OF DIGITAL FILTERS

To provide additional insight in the application and impact of precision in digital filter, examples of two practical digital applications are shown. The first example is an equalization of a small monitor loudspeaker. The second example is an electronic crossover for a 3 way loudspeaker. The filter architecture chosen for both of these examples is a IIR filter structure which is composed of cascaded sections of second order Direct Form I filters that use magnitude truncation. The filters will use 28 bit coefficients. The processing architecture is a 48 bit fixed point architecture to support 24 data bits, 8 overhead and 16 noise bits. This architecture is illustrated in Figure 21. The multiplier accepts a 28 bit coefficient by 48 bit data multiplication to produces a 76 bit result. The **Q** block quantizes the signal from 76 bits to 48 bits. The adder and accumulators will support 76 bit add and accumulation. The signal path between each of the IIR filters is 48 bits. At the conclusion of processing, the result is truncated to the desired output length of 24 or 32 bits. Unless other wise specified, the data sample rate is 48 kHz for the examples.

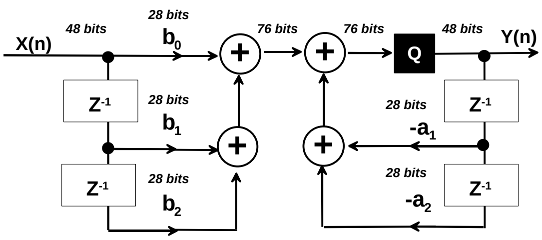


Figure 21 Processing Architecture

Loudspeaker Equalization Example

In this example, a two-way bass reflex (or ported) monitor loudspeaker is equalized to provide a frequency response that is flat with a gradual –5 dB de-emphasis of the high frequencies. The monitor has a 5 "inch woofer and a 1" soft dome tweeter. The port tuning is at 85 Hz. The uncorrected frequency response of the loudspeaker is shown in Figure 22.

The objectives of this equalization is to reduce the variations in the overall frequency response, extend the low frequency response, and improve the power handling of low frequency information. The improvement in the power handling of low frequency information is needed to eliminate distortion and noise that is produced by the loudspeaker by 50 Hz and lower tones when played at levels exceeding 15 watts.

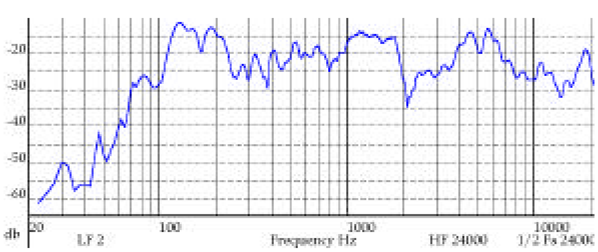


Figure 22 Frequency Response of Monitor Loudspeaker

The equalization developed for this example uses 8 cascaded second order IIR filters. Figure 23 shows the corrected loudspeaker response (shown in Red). Tables 2 a and b contain the filter parameters.

Filter Type	Freq
Butterworth High Pass	40
Butterworth High Pass	50

Table 2 a High Pass Filters

Filter Type	Freq	Gain	BW
Parametric EQ	85	15.0	20
Parametric EQ	135	- 13.0	20
Parametric EQ	188	-5.0	20
Parametric EQ	1300	- 7.5	522
Parametric EQ	2100	14.0	102
Parametric EQ	5500	- 10.0	750

Table 2 b Parametric Filters

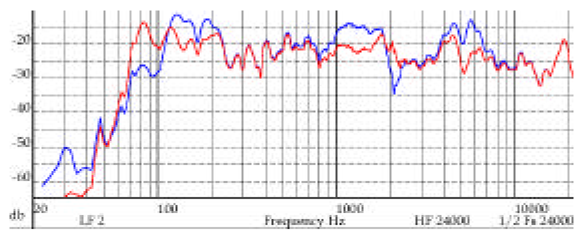


Figure 23 Equalized Response

The two high pass Butterworth filters provide a fourth order high pass filter that improves the loudspeaker power handling at low frequencies. This filter reduces low frequency energy that is sent to the loudspeaker for frequencies below the loudspeaker resonant frequency. For frequencies that are above the resonant frequency of the loudspeaker, the cabinet provides an acoustic load to the woofer, which dampens the woofer's motion. However, below the resonant frequency the woofer becomes acoustically unloaded. At these frequencies only a relatively modest amount of energy is necessary to cause the woofer to move out to the suspension limits, thereby producing noise and distortion.

The equalization filters at 85, 135, and 188 Hz flatten and extend the low frequency response of the loudspeaker over the interval of 75 Hz to 210 Hz.

The 1300, 2100 and 5500 Hz equalization filters compensate for the response irregularities of the woofer and tweeter on either side of the crossover frequency of 3200 Hz.

The following set of figures show the individual and collective characteristics of the filters that are used to perform the equalization.
Figure 24 shows the transfer functions of the 8 individual filters that were used to produce the equalized response.

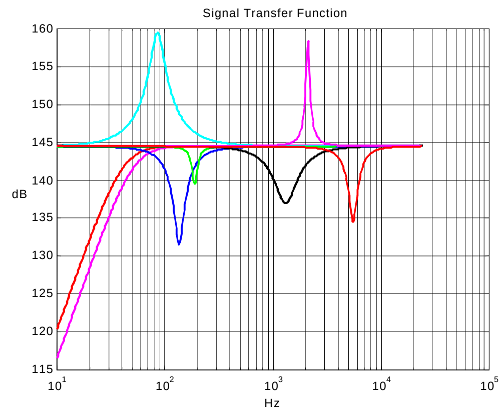


Figure 24 Transfer functions of the 8 individual filters

Figure 25 shows the noise transfer functions of the 8 individual filters that were used to produce the equalized response.
Figure 26 shows the difference transfer functions of the 8 individual filters that were used to produce the equalized response.

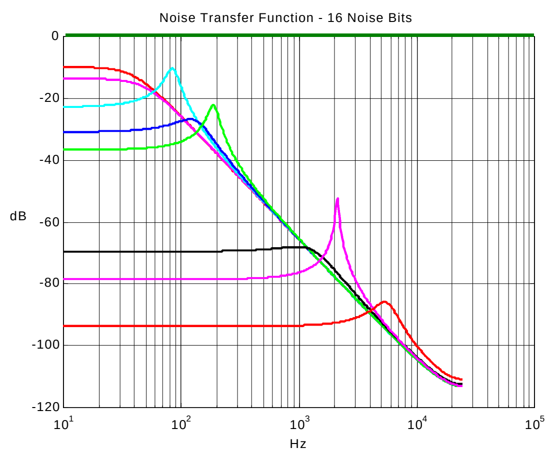


Figure 25 Noise Transfer Functions of the 8 individual filters – 16 Noise Bits

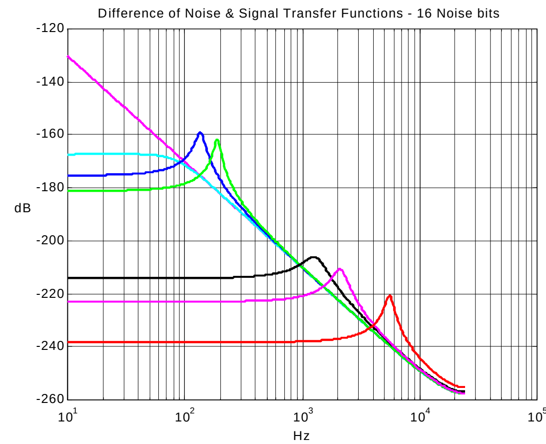


Figure 26 Difference Transfer Functions of the 8 individual filters – 16 Noise Bits

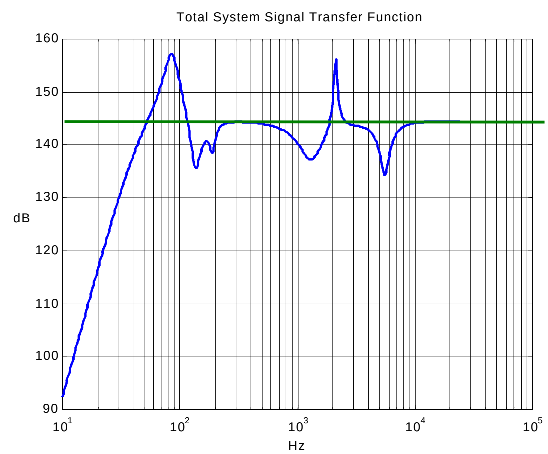


Figure 27 Total System Signal Transfer Function

As shown Figure 28, the noise produced by the filter is less than the noise floor of an ideal 24 bit input (144.49 dB) for all frequencies. Figure 29 shows the signal to noise difference is better than an ideal input for all frequencies above 21 Hz. The SNR drops below 144 dB below 21 Hz as a result of the signal attenuation from two 2nd order high pass filters at 40 and 50 Hz. This excellent performance is a result of using 28 bit coefficient and 48 bit data word (24 data bits, 8 headroom bits, and 16 noise bits).

To illustrate the impact of 16 versus 8 noise bits, the total system noise and total system difference transfer functions are shown in Figures 30 and 31 for an 8 noise bit system. These figures illustrate that, by using only 8 noise bits, the filter will only achieve 144 dB performance for frequencies above 500 Hz.

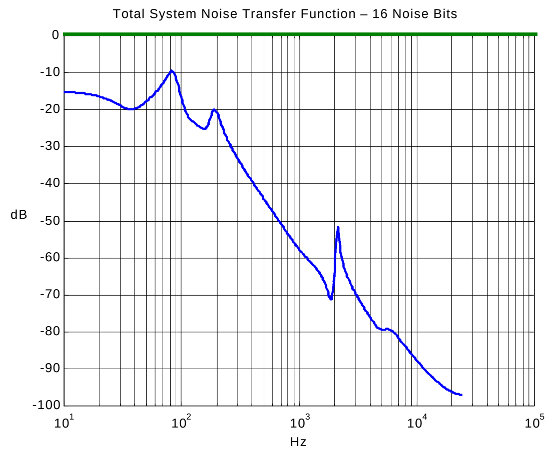


Figure 28 Total System Noise Transfer Function – 16 Noise Bits

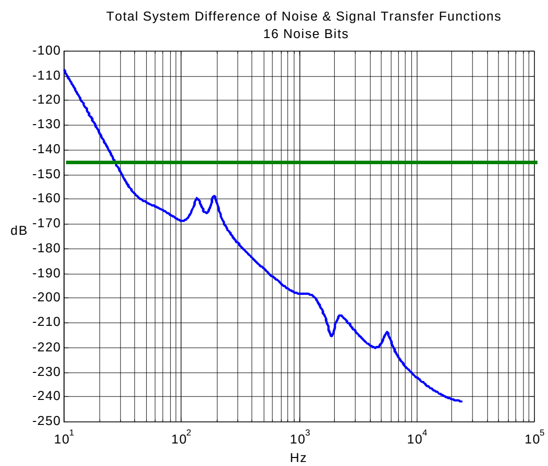


Figure 29 Total System Difference Transfer Function – 16 Noise Bits

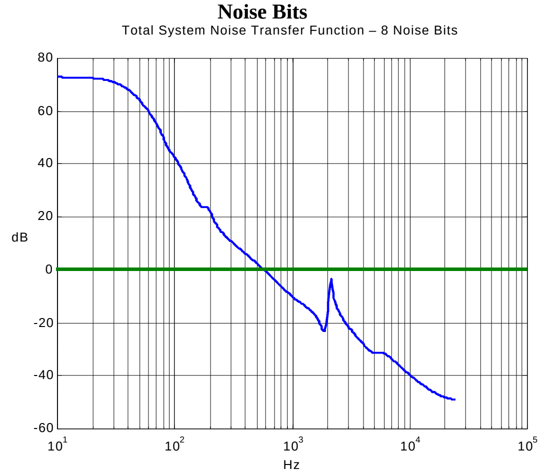


Figure 30 Total System Noise Transfer Function – Using 8 Noise bits

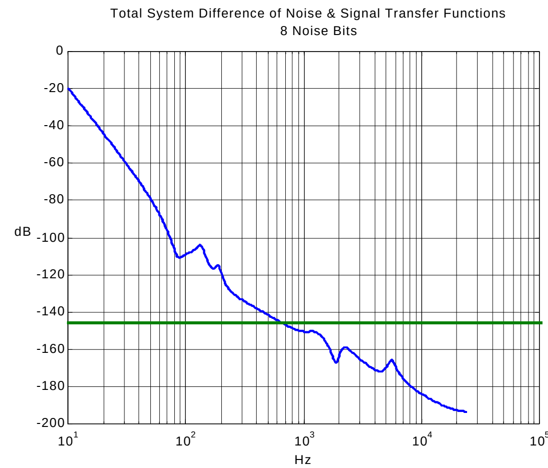


Figure 31 Total System Difference Transfer Function – Using 8 Noise bits

Loudspeaker Crossover Example

To gain additional insight into the effect of precision on filter performance, a digital crossover application is investigated. In this example a digital crossover is developed for a three-way loudspeaker. The loudspeaker is composed of a 12 inch woofer, a 5 inch midrange and a one inch soft dome tweeter. The woofer is in a vented cabinet. The midrange is in an acoustic suspension cabinet. The design objective is to develop a maximally flat response using crossovers that provide second order acoustic responses. The following figure depict the crossover and equalization filters that were developed for the three loudspeaker transducers.

Figure 32 shows the uncorrected woofer response (blue) the target response (black), and the maximum and minimum thresholds for the target response (magenta).

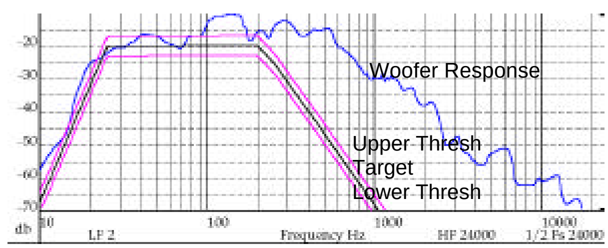


Figure 32 Woofer Response and Target Response

Figure 33 shows the preceding figure with the addition of the shaped woofer response (red).

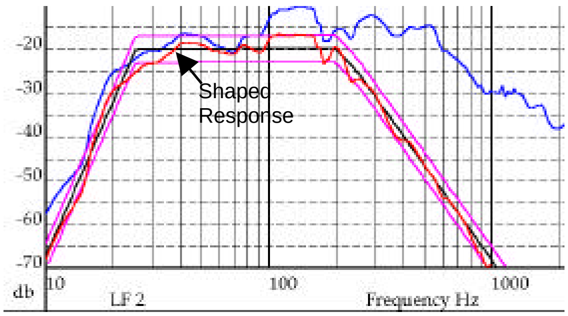


Figure 33 Woofer Target Crossover and Equalization

Table 3 contains the filter descriptions used to develop the woofer response. The filter complement for the woofer contains a 15 Hz high pass Linkwitz Riley second order high-pass filter to decrease the electrical energy the woofer that is below the acoustic resonance of the tuned cabinet. The woofer high frequency response is shaped by the 100 Hz second order Linkwitz Riley low-pass plus the 200 and 532 Hz Equalization filters. Figure 34 shows the individual woofer signal transfer functions.

Type	Gain	Freq	BW
Linkwitz Riley HP		15	
EQ	3	70	20
Linkwitz Riley LP		100	
EQ	10	200	50
EQ	-4.9	532	92

Table 3 Woofer Crossover and Equalization filters

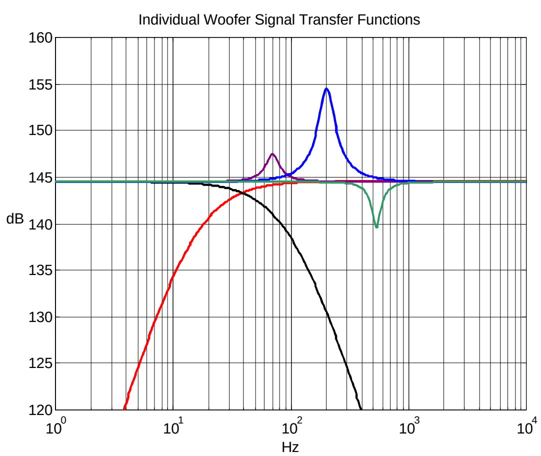


Figure 34 Woofer Signal Transfer Functions

Figure 35 shows the uncorrected midrange response (blue) and the shaped midrange response (red).

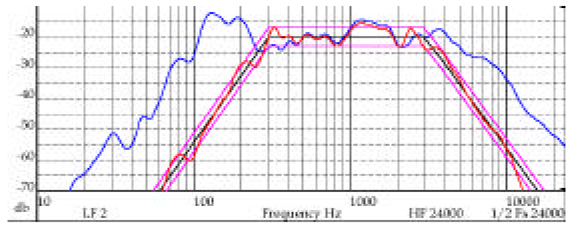


Figure 35 Midrange Target Crossover and Equalization

Table 4 contains the filter descriptions used to develop the midrange response. The midrange low frequency response is shaped by the 250 Hz second order Linkwitz Riley high-pass plus the 125, 240, and 321 Hz Equalization filters. The midrange high frequency response is shaped by the 3600 Hz second order Linkwitz Riley low-pass and the 7397 Hz Equalization filters.

Filter Type	Gain	Freq	BW
Linkwitz Riley HP		250	
Equalization	-22	125	30
Equalization	8	240	300
Equalization	14.5	321	76
Linkwitz Riley LP		3600	
Equalization	-5.3	7397	2242

Table 4 Midrange Crossover and Equalization filters

Figure 36 shows the individual midrange signal transfer functions.

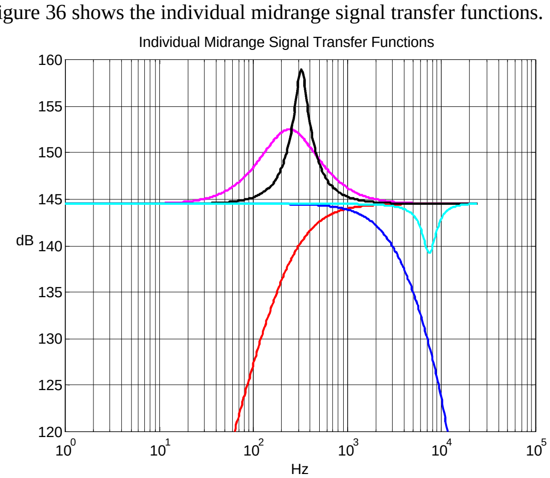


Figure 36 Midrange Signal Transfer Functions

Figure 37 shows the uncorrected tweeter response (blue) and the shaped tweeter response (red).

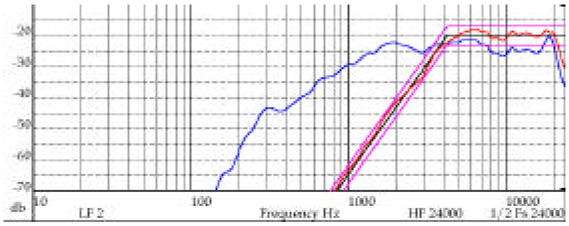


Figure 37 Tweeter Target Crossover and Equalization

Table 5 contains the filter descriptions used to develop the tweeter response. The tweeter low frequency response is shaped by the 3600 Hz second order Linkwitz Riley high-pass plus the 2000 Hz bass shelf filters. The high frequency response of the tweeter is shaped by the 3000 Hz treble shelf and the 18939 Hz equalization filters.

Filter Type	Gain	Freq	Bandwidth
Bass Shelf	-8	2000	
Linkwitz Riley HP		3600	
Treble Shelf	5	3000	1000
Equalization	-4	18936	1655

Table 5 Tweeter Crossover and Equalization filters

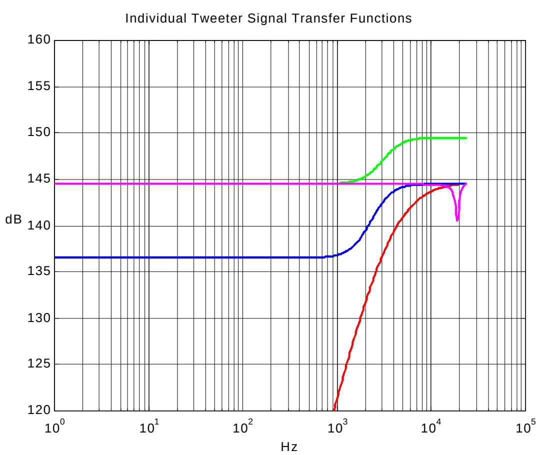


Figure 38 Tweeter Signal Transfer Functions

Figures 39 a, b and c show the woofer, midrange and tweeter signal transfer functions.

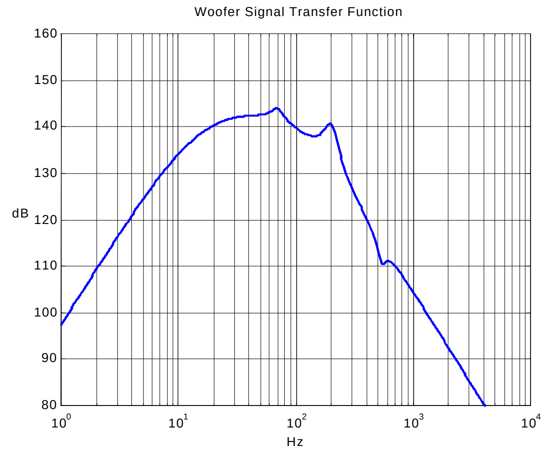


Figure 39 a Woofer Signal Transfer Function

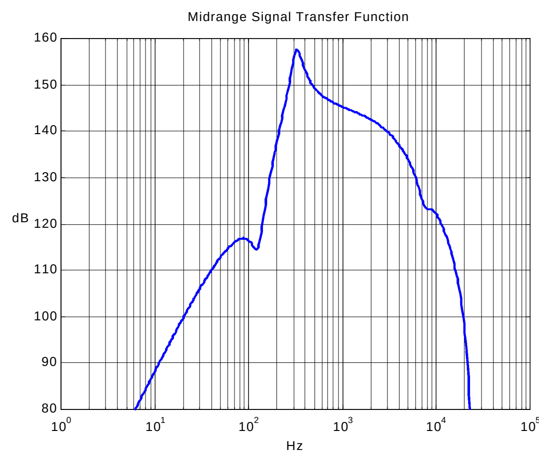


Figure 39 b Midrange Signal Transfer Function

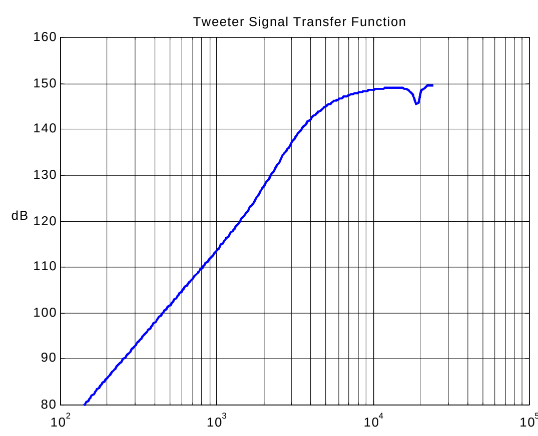


Figure 39 c Tweeter Signal Transfer Function

Figures 40 a, b and c show the woofer, midrange and tweeter noise transfer functions.

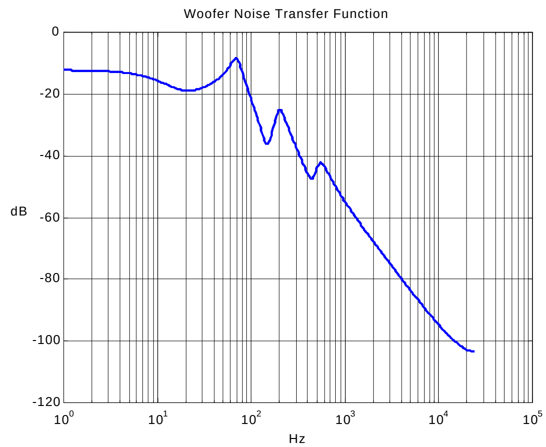


Figure 40 a Woofer Noise Transfer Function

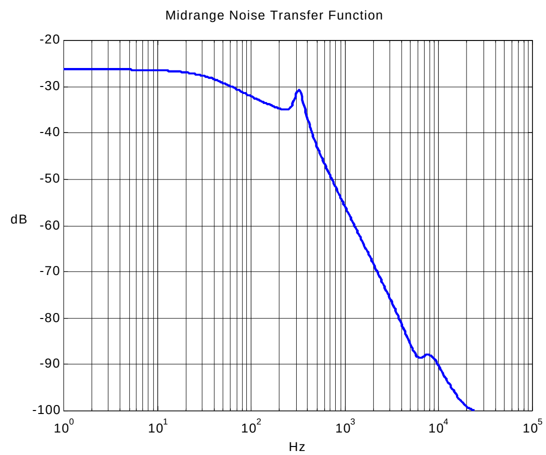


Figure 40 b Midrange Noise Transfer Function

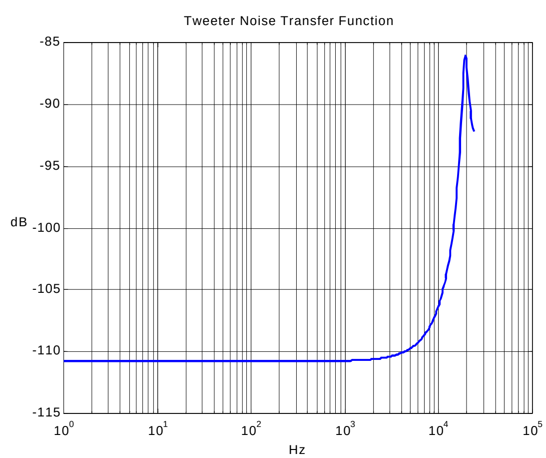


Figure 40 c Tweeter Noise Transfer Function

Figures 41 a, b and c show the woofer, midrange and tweeter difference transfer functions.

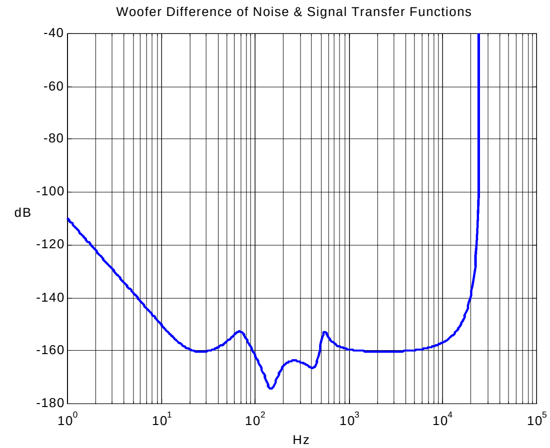


Figure 41 a Woofer Difference Transfer Function

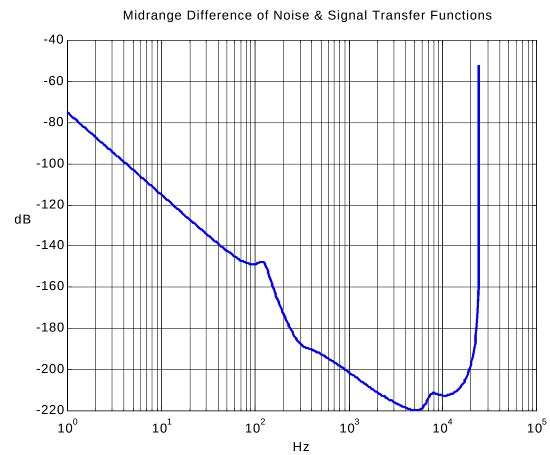


Figure 41 b Midrange Difference Transfer Function

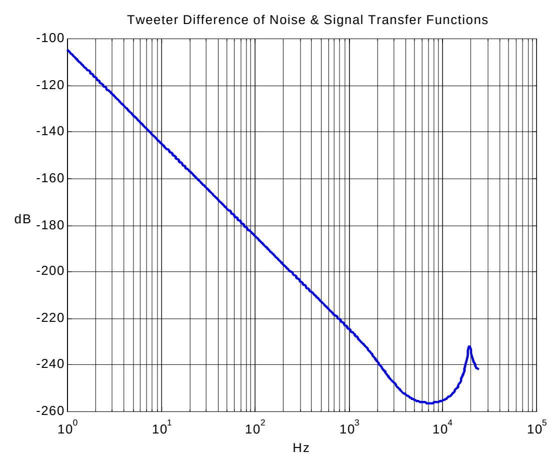


Figure 41 c Tweeter Difference Transfer Function

As can been seen in the Noise Transfer Function plots, Figures 40 a, b and c, the noise produced by the filters is less than the noise floor of an ideal 24 bit input (144.49 dB) for all frequencies. This is substantiated in the difference transfer functions, Figures 41 a, b and c, which show that the system performance drops below an ideal 24 bit input only at very low signal output levels. This excellent performance is a result of using 28 bit coefficients and 48 bit data word (24 data bits 8 head room bits, and 16 noise bits).

SUMMARY AND CONCLUSIONS

Digital filters are becoming ubiquitous in audio applications. As a result, good digital filter performance is important to audio system design. Digital filters differ from conventional analog filters by their use of finite precision to represent signals and coefficients and finite precision arithmetic to compute the filter response. The precision that is used determines the digital filter's response accuracy and the filter signal to noise ratio.

The coefficient precision determines the accuracy of the digital filter response in comparison to an ideal filter. As was shown, a second order Direct Form 1 filter using 24 bit coefficients can achieve a 1 dB or better response accuracy in implementing a modest parametric equalization filter with a 6 dB gain and a Q of 6 over the range of 50 Hz to 20,000 Hz at a 48kHz sample rate. However, when the sample rate is increased to 96 or 192 kHz, additional precision, 28 bits, is required to obtain similar performance. When less than this precision is used, the resulting filter response deviates substantially from the desired response.

Similarly, the precision that is used in computing the filter response plus the computation method determines the signal to noise performance of the filter. Noise arises in a digital filter as a result of the quantization that occurs in computing the filter response. The precision that is maintained throughout the computation determines the noise amplitude. As was shown, 16 noise bits are sufficient to preserve the SNR performance of the input signal for a parametric equalization filter with a gain of 12 dB and a Q of 6.7 from 100 to 20,000 Hz at sample rates of 48 and 96 kHz. If fewer noise bits are used, the SNR performance can degrade substantially. An example of this degradation is for the filter at 100 Hz, where 8 noise bits produce a loss of 40 dB in SNR at a 48 kHz sample rate and a loss of 50 dB SNR at a 96 kHz sample rate.

Finite precision representation also imposes a limitation on the maximum signal magnitude that can be represented. To permit positive gains to be used in forming intermediate values within and between cascaded stages, additional precision is required to avoid numeric overflow. As was discussed, 8 bits of additional precision provides sufficient headroom for a majority of cases.

In conclusion, the quality of 24 bit data at 48 and 96 kHz sample rates can be preserved during digital filtering applications by using:

- 48 bits of data precision (24 bits of data, 8 headroom bits and 16 noise bits)
- 28 bit coefficients
- An IIR filter structure which is composed of cascaded sections of second order Direct Form I filters that use magnitude truncation.

Credits:

The author is very grateful to Dr. Rusty Allred for his contributions to this paper.

References:

A.Oppenheim and R.Schafer, *Discrete-Time Signal Processing*, Prentice – Hall, 2nd Edition, 1998, ISBN 0-13-754920-2

P. Kraniauskas, *Transforms in Signals and Systems*, Addison – Wesley, 1993, ISBN 0-201-10694-8

E.C. Ifeachor and B.W. Jervis, *Digital Signal Processing A Practical Approach*, Addison – Wesley, 1993, ISBN 0-201-54413-X