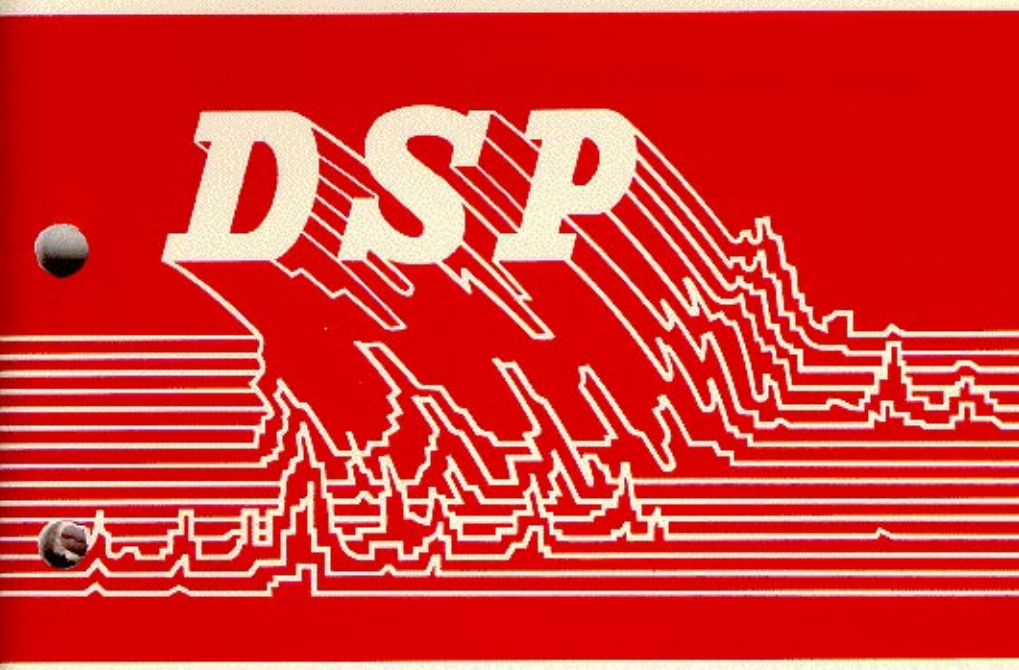
APR8/D
Rev. 1

Principles of Sigma-Delta Modulation for Analog-to-Digital Converters



MOTOROLA

Motorola Digital Signal Processors

Principles of Sigma-Delta Modulation for Analog-to-Digital Converters

by
Sangil Park, Ph. D.
Strategic Applications
Digital Signal Processor Operation

Table of Contents

Illustrations

Illustrations

Tables

SECTION 1

Introduction

"Since the Σ–Δ A/D converters are based on digital filtering techniques, almost 90% of the die is implemented in digital circuitry which enhances the prospect of compatibility."

The performance of digital signal processing and communication systems is generally limited by the precision of the digital input signal which is achieved at the interface between analog and digital information. Sigma-Delta (Σ–Δ) modulation based analog-to-digital (A/D) conversion technology is a cost effective alternative for high resolution (greater than 12 bits) converters which can be ultimately integrated on digital signal processor ICs.

Although the sigma-delta modulator was first introduced in 1962 [1], it did not gain importance until recent developments in digital VLSI technologies which provide the practical means to implement the large digital signal processing circuitry. The increasing use of digital techniques in communication and audio application has also contributed to the recent interest in cost effective high precision A/D converters. A requirement of analog-to-digital (A/D) interfaces is compatibility with VLSI technology, in order to provide for monolithic integration of both the analog and digital sections on a single die. Since the Σ–Δ A/D converters are based on digital filtering techniques, almost 90% of the die is implemented in digital circuitry which enhances the prospect of compatibility.

Additional advantages of such an approach include higher reliability, increased functionality, and reduced chip cost. Those characteristics are commonly required in the digital signal processing environment of today. Consequently, the development of digital signal processing technology in general has been an important force in the development of high precision A/D converters which can be integrated on the same die as the digital signal processor itself. The objective of this application report is to explain the Σ–Δ technology which is implemented in the DSP56ADC16, and show the superior performance of the converter compared to the performance of more conventional implementations. Particularly, this application note discusses a third-order, noise-shaping oversampling structure.

Conventional high-resolution A/D converters, such as successive approximation and flash type converters, operating at the Nyquist rate (sampling frequency approximately equal to twice the maximum frequency in the input signal), often do not make use of exceptionally high speeds achieved with a scaled VLSI technology. These Nyquist samplers require a complicated analog lowpass filter (often called an anti-aliasing filter) to limit the maximum frequency input to the A/D, and sample-and-hold circuitry. On the other hand, Σ–Δ A/D converters use a low resolution A/D converter (1-bit quantizer), noise shaping, and a very high oversampling rate (64 times for the DSP56ADC16). The high resolution can be achieved by the decimation (sample-rate reduction) process. Moreover, since

precise component matching or laser trimming is not needed for the high-resolution Σ–Δ A/D converters, they are very attractive for the implementation of complex monolithic systems that must incorporate both digital and analog functions. These features are somewhat opposite from the requirements of conventional converter architectures, which generally require a number of high precision devices. This application note describes the concepts of noise shaping, oversampling, and decimation in some detail. ■

SECTION 2

Conventional Analog-to-Digital Converters

"Most A/D converters can be classified into two groups according to the sampling rate criteria: Nyquist rate converters... and oversampling converters..."

Signals, in general, can be divided into two categories; an analog signal, x(t), which can be defined in a continuous-time domain and a digital signal, x(n), which can be represented as a sequence of numbers in a discrete-time domain as shown in Figure 2-1. The time index n of a discrete-time signal x(n) is an integer number defined by sampling interval T. Thus, a discrete-time signal, x*(t), can be represented by a sampled continuous-time signal x(t) as:

$$x^*(t) = \sum_{n=-\infty}^{\infty} x(t)\delta(t - nT) \qquad \text{Eqn. 2-1}$$

where:

$$\delta(t) = 1, \quad t = 0,$$

$$0, \quad \text{elsewhere}$$

A practical A/D converter transforms x(t) into a discrete-time digital signal, x*(t), where each sample is expressed with finite precision. Each sample is approximated by a digital code, i.e., x(t) is transformed

into a sequence of finite precision or quantized samples x(n). This quantization process introduces errors which are discussed in **SECTION 3 Quantization Error in A/D Converters**.

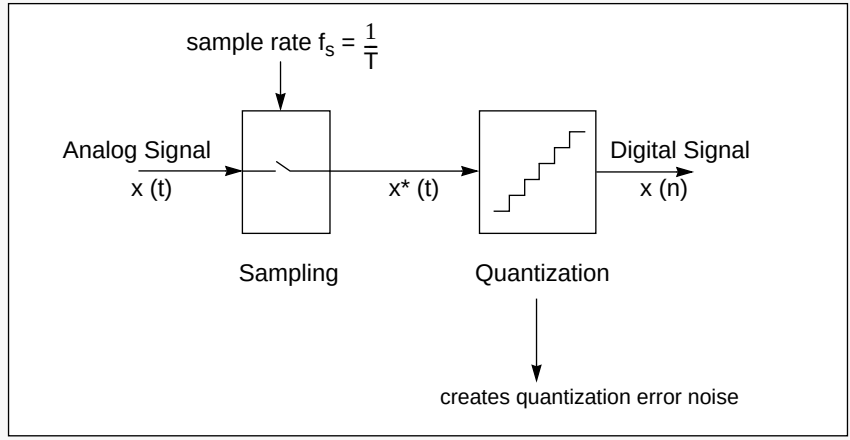


Figure 2-1 *Generalized Analog-to-Digital Conversion Process*

Most A/D converters can be classified into two groups according to the sampling rate criteria. Nyquist rate converters, such as a successive approximation register (SAR), double integration, and oversampling converters, sample analog signals which have maximum frequencies slightly less than the Nyquist frequency, $f_N = f_s/ 2$, where fs is the sampling frequency [2]. Meanwhile, oversampling converters perform the sampling process at a much higher rate, $f_N << F_s$ (e.g., 64 times for the DSP56ADC16), where F_s denotes the input sampling rate.

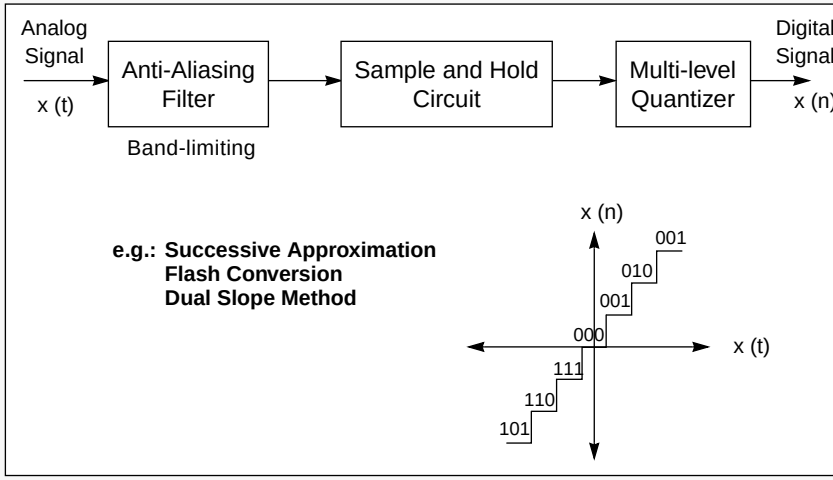


Figure 2-2 *Conventional Analog-to-Digital Conversion Process*

Figure 2-2 illustrates the conventional A/D conversion process that transforms an analog input signal x(t) into a sequence of digital codes x(n) at a sampling rate of $f_s = 1/T$, where T denotes the sampling interval. Since $\delta(t - nT)$ in Eqn. 2-1 is a periodic function with period T, it can be represented by a Fourier series given by [5]:

$$\sum_{n=-\infty}^{\infty} x(t)\delta(t-nT) = \frac{1}{T}\sum_{n=-\infty}^{\infty} x(t)e^{(j2n\pi t)/T} \qquad \text{Eqn. 2-2}$$

Combining Eqn. 2-1 and Eqn. 2-2, we get:

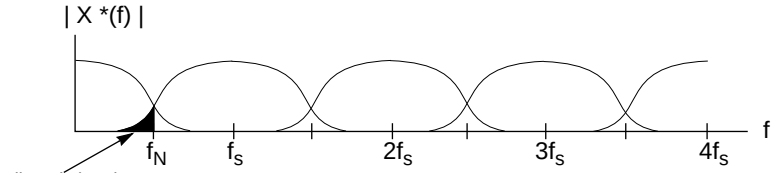
$$x^*(t) = \frac{1}{T}\sum_{n=-\infty}^{\infty} x(t)e^{(j2n\pi t)/T} = \frac{1}{T}\sum_{n=-\infty}^{\infty} x(t)e^{j2\pi f_s nt} \qquad \text{Eqn. 2-3}$$

Eqn. 2-2 states that the act of sampling (i.e., the sampling function):

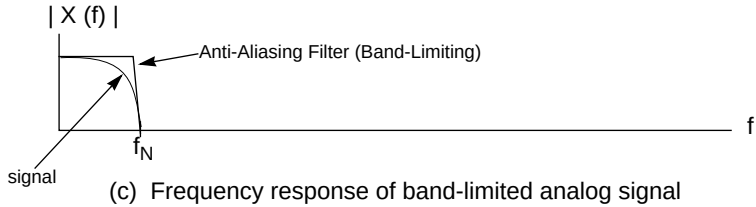
$$\sum_{n=-\infty}^{\infty} x(t)\delta(t-nT)$$

is equivalent to modulating the input signal by carrier signals having frequencies at 0, f_s, $2f_s$,. . .. In other words, the sampled signal can be expressed in the frequency domain as the summation of the original signal component and signals frequency modulated by integer multiples of the sampling frequency as shown in Figure 2-3. Thus, input signals above the Nyquist frequency, fn, cannot be properly converted and they also create new signals in the base-band, which were not present in the original signal. This non-linear phenomenon is a signal distortion frequently referred to as aliasing [2]. The distortion can only be prevented by properly lowpass filtering the input signal up to the Nyquist frequency. This lowpass filter, sometimes called the anti-aliasing filter, must have flat response over the frequency band of interest (baseband) and attenuate the frequencies above the Nyquist frequency enough to put them under the noise floor. Also, the non-linear phase distortion caused by the anti-aliasing filter may create harmonic distortion and audible degradation. Since the analog anti-aliasing filter is the limiting factor in controlling the bandwidth and phase distortion of the input signal, a high performance anti-aliasing filter is required to obtain high resolution and minimum distortion.

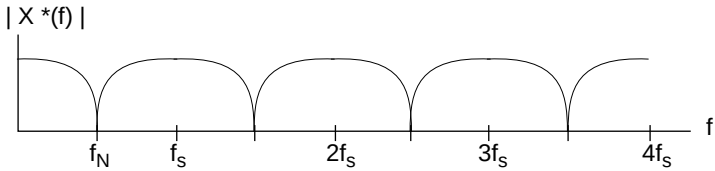
(a) Frequency response of unlimited signal



(b) Frequency response of sampled unlimited signal



(c) Frequency response of band-limited analog signal



(d) Frequency response of sampled digital signal

Figure 2-3 *Spectra of Analog and Sampled Signals*

In addition to an anti-aliasing filter, a sample-and-hold circuit is required. Although the analog signal is continuously changing, the output of the sample-and-hold circuitry must be constant between samples so the signal can be quantized properly. This allows the converter enough time to compare the sampled analog signal to a set of reference levels that are usually generated internally [3]. If the output of the sample-and-hold circuit varies during T, it can limit the performance of the A/D converter subsystem.

Each of these reference levels is assigned a digital code. Based on the results of the comparison, a digital encoder generates the code of the level the input signal is closest to. The resolution of such a converter is determined by the number and spacing of the reference levels that are predefined. For high-resolution Nyquist samplers, establishing the reference voltages is a serious challenge.

For example, a 16-bit A/D converter, which is the standard for high accuracy A/D converters, requires 216 - 1 = 65535 different reference levels. If the converter has a 2V input dynamic range, the spacing of these levels is only 30 mV apart. This is beyond the limit of component matching tolerances of VLSI technologies [4]. New techniques, such as laser trimming or self-calibration can be employed to boost the resolution of a Nyquist rate converter beyond normal component tolerances. However, these approaches result in additional fabrication complexity, increased circuit area, and higher cost.

■

SECTION 3

Quantization Error in A/D Conversion

"For an input signal which is large compared to an LSB step, the error term e(n) is a random quantity in the interval (-q/2, q/2) with equal probability."

The process of converting an analog signal (which has infinite resolution by definition) into a finite range number system (quantization) introduces an error signal that depends on how the signal is being approximated. This quantization error is on the order of one least-significant-bit (LSB) in amplitude, and it is quite small compared to full-amplitude signals. However, as the input signal gets smaller, the quantization error becomes a larger portion of the total signal.

When the input signal is sampled to obtain the sequence x(n), each value is encoded using finite word-lengths of B-bits including the sign bit. Assuming the sequence is scaled such that $|x(n)| \leq 1$ for fractional number representation, the pertinent dynamic range is 2. Since the encoder employs B-bits, the number of levels available for quantizing x(n) is 2^B. The interval between successive levels, q, is therefore given by:

$$q = \frac{1}{2^{B-1}} \qquad \text{Eqn. 3-1}$$

which is called the quantization step size. The sampled input value $x^*(t)$ is then rounded to the nearest level, as illustrated in Figure 3-1.

From Eqn. 3-2, it follows that the A/D converter output is the sum of the actual sampled signal $x^*(t)$ and an error (quantization noise) component e(n); that is:

$$x(n) = x^*(t) + e(n) \qquad \text{Eqn. 3-2}$$

For an input signal which is large compared to an LSB step, the error term e(n) is a random quantity in the interval (-q/2, q/2) with equal probability. Then the noise power (variance), σ_e^2, can be found as [5]:

$$\sigma_e^2 = \mathbf{E}[e^2] = \frac{1}{q}\int_{(-q)/2}^{q/2} e^2 de = \frac{q^2}{12} = \frac{2^{-2B}}{3} \qquad \text{Eqn. 3-3}$$

where: E denotes statistical expectation

We shall refer to σ_e^2 in Eqn. 3-3 as being the steady-state input quantization noise power. Figure 3-2 shows the spectrum of the quantization noise. Since the noise power is spread over the entire frequency range equally, the level of the noise power spectral density can be expressed as:

$$N(f) = \frac{q^2}{12f_s} = \frac{2^{-2B}}{3f_s} \qquad \text{Eqn. 3-4}$$

The concepts discussed in this section apply in general to A/D converters. ■

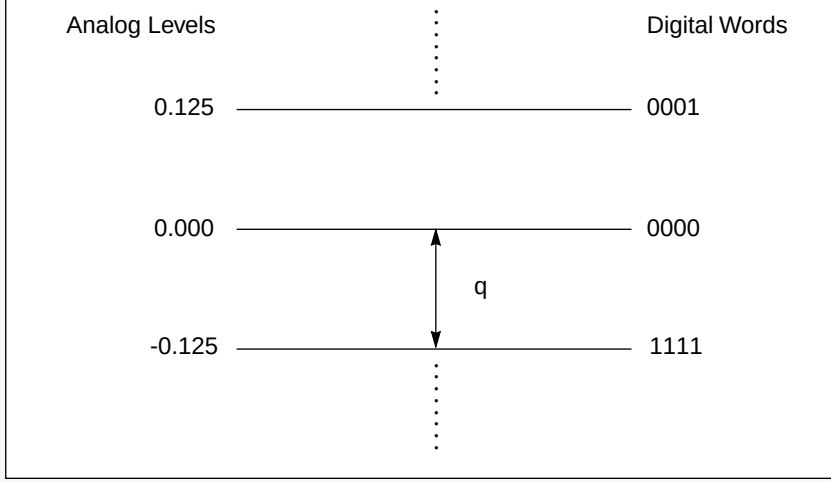


Figure 3-1 *Quantization Error*

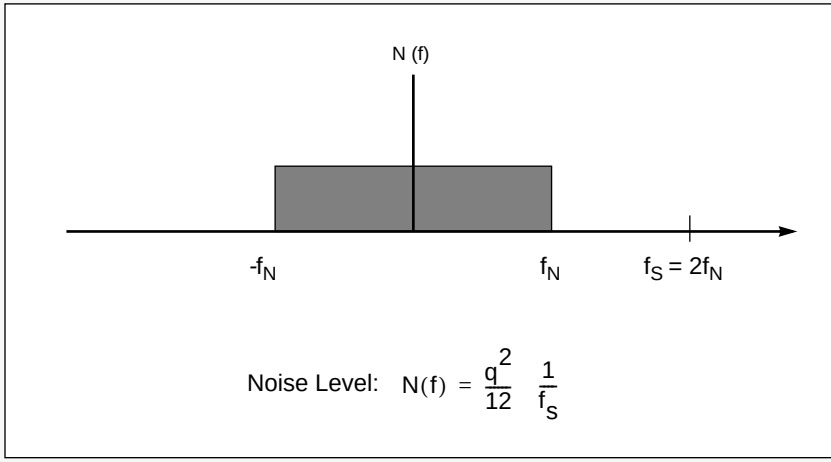


Figure 3-2 *Noise Spectrum of Nyquist Samplers*

SECTION 4

Oversampling and Decimation Basics

"The benefit of oversampling is more than an economical anti-aliasing filter. The decimation process can be used to provide increased resolution."

The quantization process in a Nyquist-rate A/D converter is generally different from that in an oversampling converter. While a Nyquist-rate A/D converter performs the quantization in a single sampling interval to the full precision of the converter, an oversampling converter generally uses a sequence of coarsely quantized data at the input oversampling rate of $F_S = Nf_S$ followed by a digital-domain decimation process to compute a more precise estimate for the analog input at the lower output sampling rate, f_S, which is the same as used by the Nyquist samplers.

Regardless of the quantization process oversampling has immediate benefits for the anti-aliasing filter. To illustrate this point, consider a typical digital audio application using a Nyquist sampler and then using a two times oversampling approach. Note that in the following discussion the full precision of a Nyquist sampler is assumed. Coarse quantizers will be considered separately.

The data samples from Nyquist-rate converters are taken at a rate of at least twice the highest signal frequency of interest. For example, a 48 kHz sampling rate allows signals up to 24 kHz to pass without

aliasing, but because of practical circuit limitation, the highest frequency that passes is actually about 22 kHz. Also, the anti-aliasing filter in Nyquist A/D converters requires a flat response with no phase distortion over the frequency band of interest (e.g., 20 kHz in digital audio applications). To prevent signal distortion due to aliasing, all signals above 24 kHz for a 48 kHz sampling rate must be attenuated by at least 96 dB for 16 bits of dynamic resolution.

These requirements are tough to meet with an analog low-pass filter. Figure 4-1(a) shows the required analog anti-aliasing filter response, while Figure 4-1(b) shows the digital domain frequency spectrum of the signal being sampled at 48 kHz.

Now consider the same audio signal sampled at $2f_S$, 96 kHz. The anti-aliasing filter only needs to eliminate signals above 74 kHz, while the filter has flat response up to 22 kHz. This is a much easier filter to build because the transition band can be 52 kHz (22k to 74 kHz) to reach the -96 dB point. However, since the final sampling rate is 48 kHz, a sample rate reduction filter, commonly called a decimation filter, is required but it is implemented in the digital domain, as opposed to anti-aliasing filters which are implemented with analog circuitry. Figure 4-1(d) and Figure 4-1(e) illustrate the analog anti-aliasing filter requirement and the digital-domain frequency response, respectively. The spectrum of a required digital decimation filter is shown in Figure 4-1(f). Details of the decimation process are discussed in **SECTION 7 Digital Decimation Filtering.**

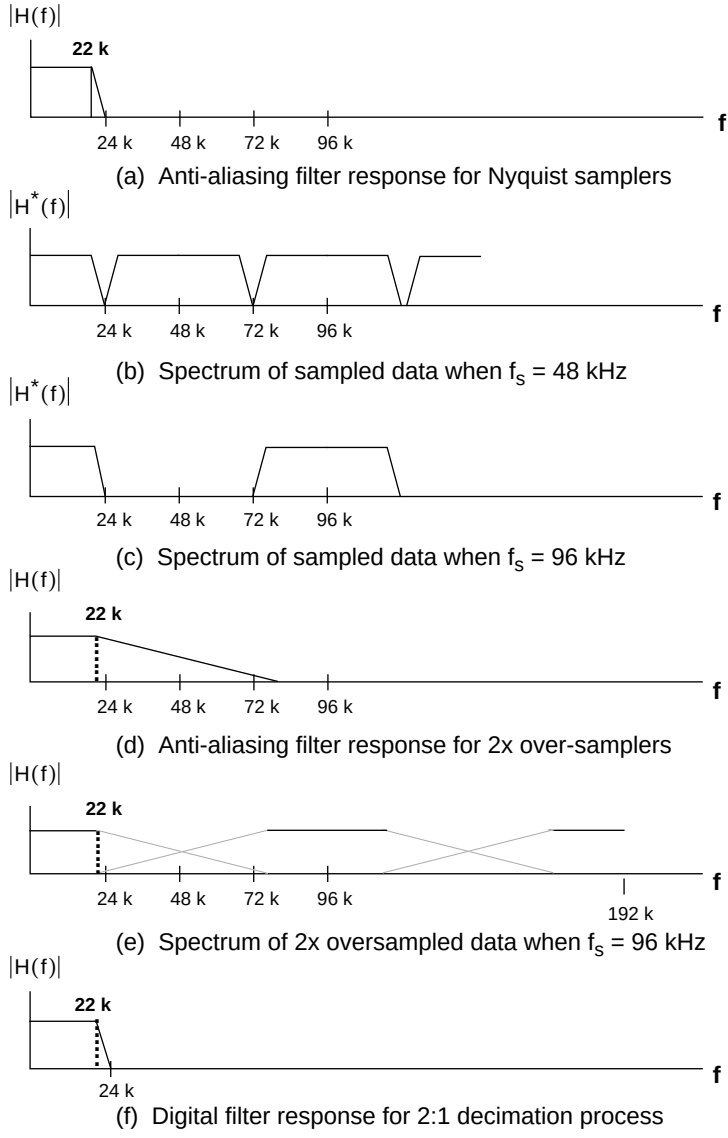


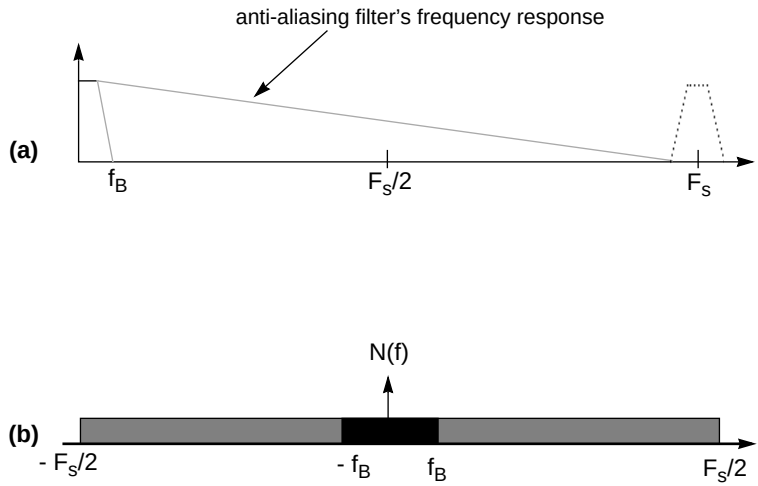
Figure 4-1 *Comparison Between Nyquist Samplers and 2X Oversamples*

This two-times oversampling structure can be extended to N times oversampling converters. Figure 4-2(a) shows the frequency response of a general anti-aliasing filter for N times oversamplers, while the spectra of overall quantization noise level and baseband noise level after the digital decimation filter is illustrated in Figure 4-2(b). Since a full precision quantizer was assumed, the total noise power for oversampling converters and one times Nyquist samplers are the same. However, the percentage of this noise that is in the bandwidth of interest, the baseband noise power NB is:

$$N_B = \int_{-f_B}^{f_B} (f)df = \frac{2f_B}{F_s}\frac{q^2}{12} \qquad \text{Eqn. 4-1}$$

which is much smaller (especially when F_s is much larger than f_B) than the noise power of Nyquist samplers described in Eqn. 3-4.

Figure 4-3 compares the requirements of the anti-aliasing filter of one times and N times oversampled Nyquist rate A/D converters. Sampling at the Nyquist rate mandates the use of an anti-aliasing filter with very sharp transition in order to provide adequate aliasing protection without compromising the signal bandwidth f_B .



Overall Noise Level: $N(f) = \frac{q^2}{12} \quad \frac{1}{F_s}$

In-Band Noise Level: $N_B = \int_{-f_B}^{f_B} N(f)df = \frac{q_{f_B}^2}{12} \frac{2}{F_s}$

Note: One R-C lowpass filter is sufficient for the anti-aliasing filter

Figure 4-2 *Anti-Aliasing Filter Response and Noise Spectrum of Oversampling A/D Converters*

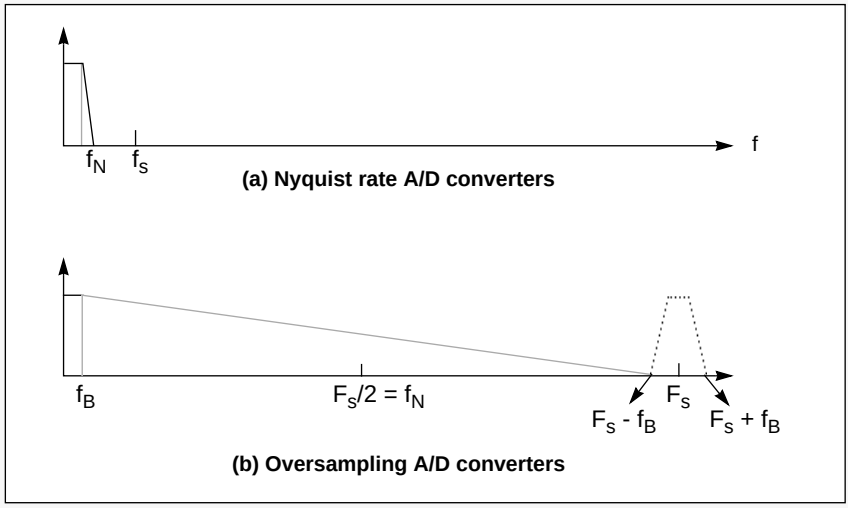


Figure 4-3 *Frequency Response of Analog Anti-Aliasing Filters*

The transition band of the anti-aliasing filter of an oversampled A/D converter, on the other hand, is much wider than its passband, because anti-aliasing protection is required only for frequency bands between $NF_S - f_B$ and $NF_S + f_B$, when N = 1, 2, ..., as shown in Figure 4-2(b). Since the complexity of the filter is a strong function of the ratio of the width of the transition band to the width of the passband, oversampled converters require considerably simpler anti-aliasing filters than Nyquist rate converters with similar performance. For example, with N = 64, a simple RC lowpass filter at the converter analog input is often sufficient as illustrated in Figure 4-2(a).

The benefit of oversampling is more than an economical anti-aliasing filter. The decimation process can be used to provide increased resolution. To see how this is possible conceptually, refer to Figure 4-4, which shows an example of 16:1 decimation process with 1-bit input samples. Although the input data resolution is only 1-bit (0 or 1), the averaging method (decimation) yields more resolution (4 bits [24 = 16]) through reducing the sampling rate by 16:1. Of course, the price to be paid is high speed sampling at the input — speed is exchanged for resolution.

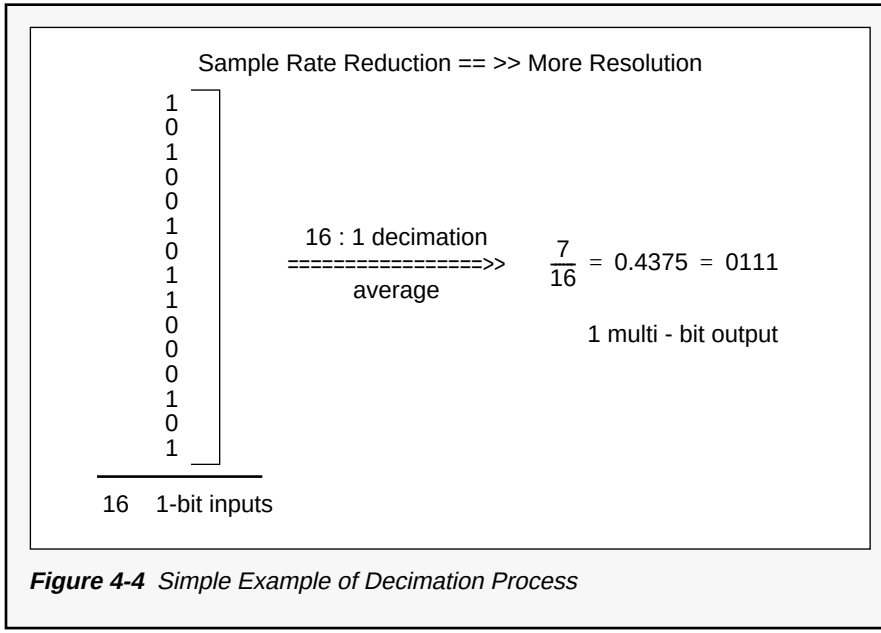


Figure 4-4 *Simple Example of Decimation Process*

SECTION 5

Delta Modulation

"Delta modulators, furthermore, exhibit slope overload for rapidly rising input signals, and their performance is thus dependent on the frequency of the input signal."

The work on sigma-delta modulation was developed as an extension to the well established delta modulation [6]. Let's consider the delta modulation/ demodulation structure for the A/D conversion process. Figure 5-1 shows the block diagram of the delta modulator and demodulator. Delta modulation is based on quantizing the change in the signal from sample to sample rather than the absolute value of the signal at each sample.

Since the output of the integrator in the feedback loop of Figure 5-1(a) tries to predict the input x(t), the integrator works as a predictor. The prediction error term, $x(t) - \bar{x}(t)$, in the current prediction is quantized and used to make the next prediction. The quantized prediction error (delta modulation output) is integrated in the receiver just as it is in the feed back loop [7]. That is, the receiver predicts the input signals as shown in Figure 5-1.

The predicted signal is smoothed with a lowpass filter. Delta modulators, furthermore, exhibit slope overload for rapidly rising input signals, and their performance is thus dependent on the frequency of the input signal. In theory, the spectrum of quantization noise of the prediction error is flat and the noise level is set by the 1-bit comparator. Note that the signal-to-noise ratio can be enhanced by decimation processes as shown

in Figure 4-2. The Motorola MC3417 Continuously Variable Slope Delta Modulator operation is based on delta modulation.

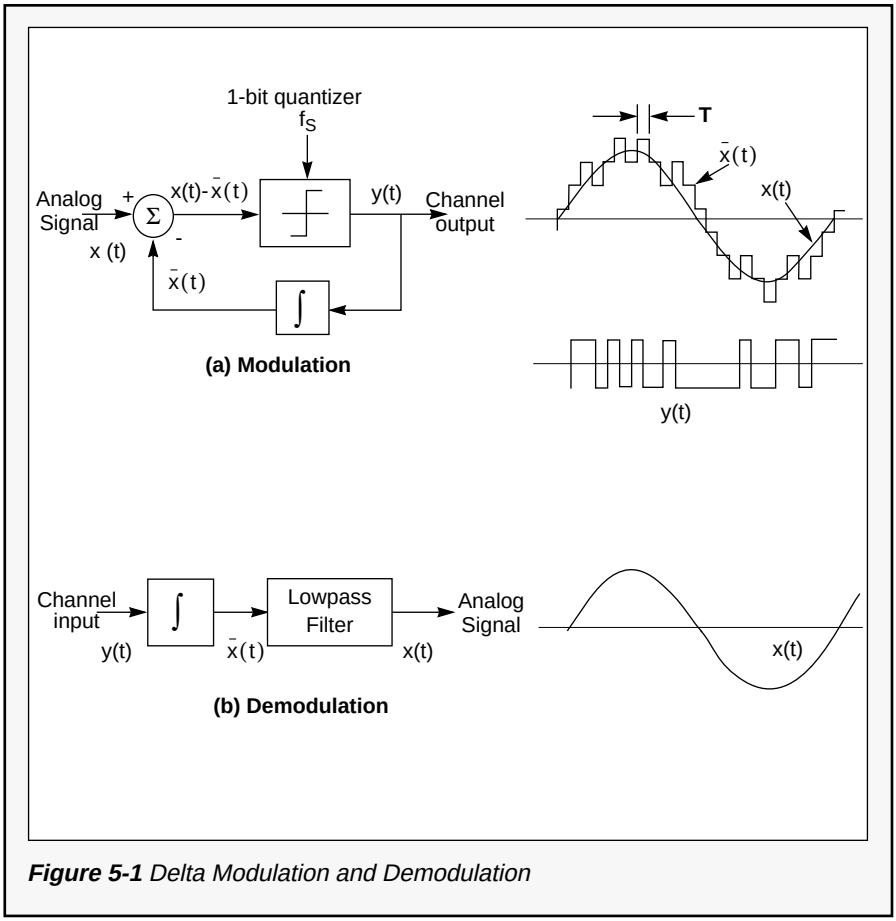


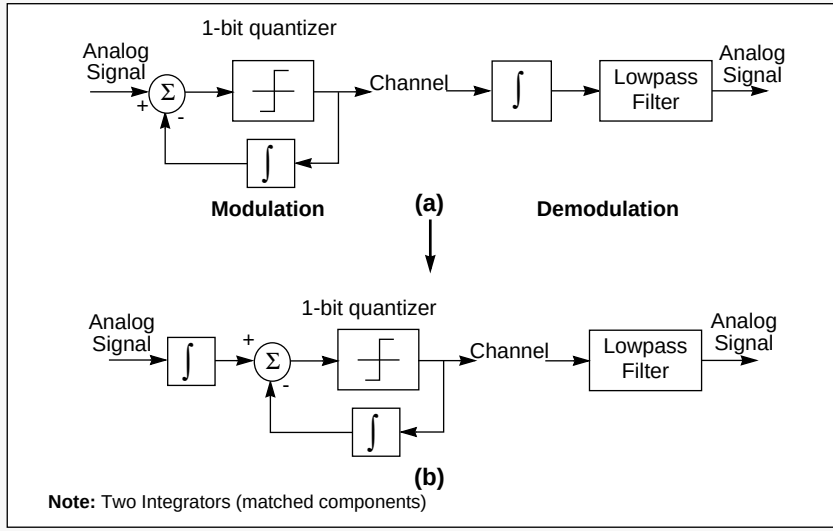
Figure 5-1 *Delta Modulation and Demodulation*

SECTION 6

Sigma-Delta Modulation and Noise Shaping

"The noise shaping property is well suited to signal processing..."

Delta modulation requires two integrators for modulation and demodulation processes as shown in Figure 6-1(a). Since integration is a linear operation, the second integrator can be moved before the modulator without altering the overall input/output characteristics. Furthermore, the two integrators in Figure 6-1 can be combined into a single integrator by the linear operation property.



Note: Two Integrators (matched components)

Figure 6-1 *Derivation of Sigma-Delta Modulation from Delta Modulation*

The arrangement shown in Figure 6-2 is called a *Sigma-Delta* (Σ–Δ) *Modulator* [1]. This structure, besides being simpler, can be considered as being a "*smoothed version*" of a 1-bit delta modulator.

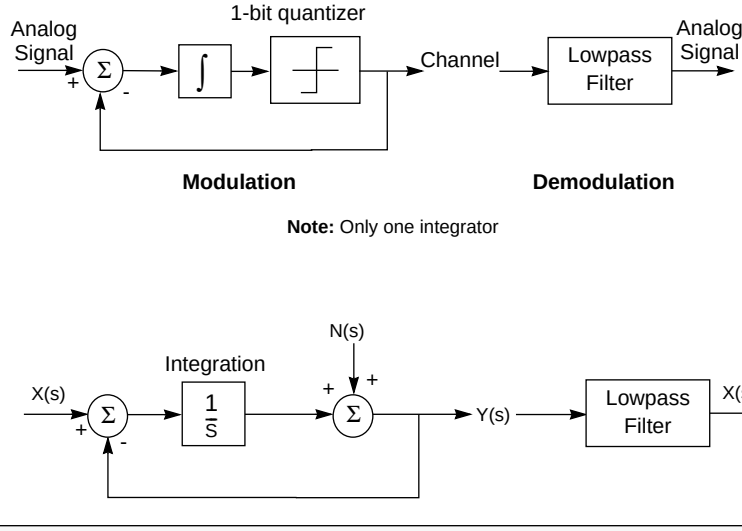


Figure 6-2 *Block Diagram of Signa-Delta Modulation*

The name *Sigma-Delta* modulator comes from putting the integrator (sigma) in front of the delta modulator. Sometimes, the Σ–Δ modulator is referred to as an interpolative coder [14]. The quantization noise characteristic (noise performance) of such a coder is frequency dependent in contrast to delta modulation. As will be discussed further, this noise-shaping property is well suited to signal processing applications such as digital audio and communication. Like delta

modulators, the Σ–Δ modulators use a simple coarse quantizer (comparator). However, unlike delta modulators, these systems encode the integral of the signal itself and thus their performance is insensitive to the rate of change of the signal.

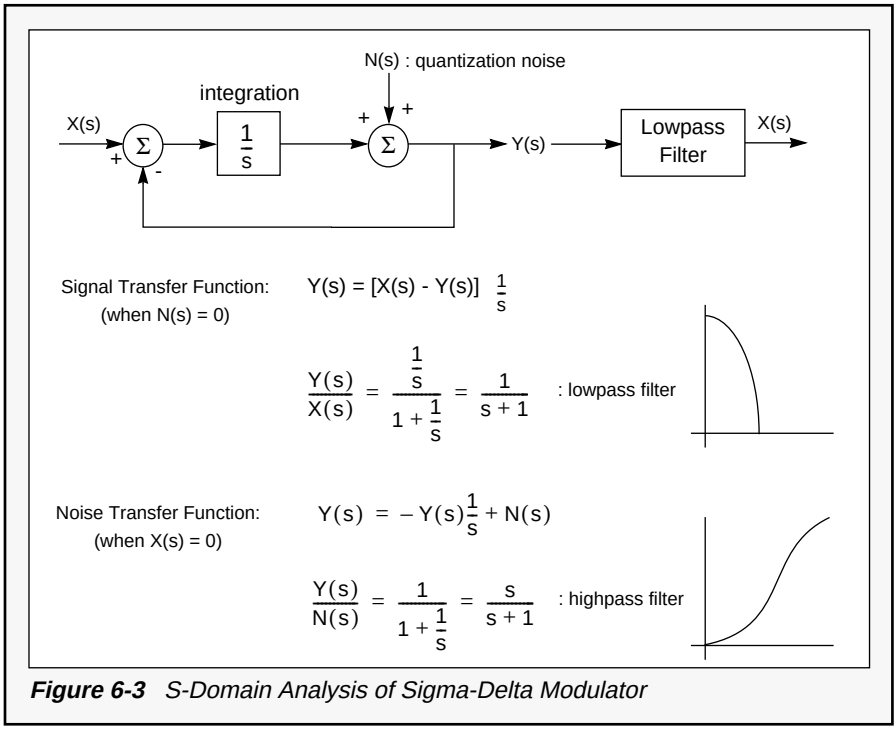


Figure 6-3 *S-Domain Analysis of Sigma-Delta Modulator*

The noise-shaping principle is illustrated by a simplified *"s-domain"* model of a first-order Σ–Δ modulator shown in Figure 6-3. The summing node to the right of the integrator represents a comparator. It's here that sampling occurs and quantization noise is added into the model. The signal-to-noise (S/N)

transfer function shown in Figure 6-3 illustrates the modulator's main action. As the loop integrates the error between the sampled signal and the input signal, it lowpass-filters the signal and highpass filters the noise. In other words, the signal is left unchanged as long as its frequency content doesn't exceed the filter's cutoff frequency, but the Σ–Δ loop pushes the noise into a higher frequency band. Grossly oversampling the input causes the quantization noise to spread over a wide bandwidth and the noise density in the bandwidth of interest (baseband) to significantly decrease.

Figure 6-4 shows the block diagram of a first-order oversampled Σ–Δ A/D converter. The 1-bit digital output from the modulator is supplied to a digital decimation filter which yields a more accurate representation of the input signal at the output sampling rate of f_S. The shaded portion of Figure 6-4 is a first-order Σ–Δ modulator. It consists of an analog difference node, an integrator, a 1-bit quantizer (A/D converter), and a 1-bit D/A converter in a feed back structure. The modulator output has only 1-bit (two-levels) of information, i.e., 1 or -1. The modulator output y(n) is converted to $\bar{x}(t)$ by a 1-bit D/A converter (see Figure 6-4).

The input to the integrator in the modulator is the difference between the input signal x(t) and the quantized output value y(n) converted back to the predicted analog signal, $\bar{x}(t)$. Provided that the D/A converter is perfect, and neglecting signal delays,

this difference between the input signal x(t) and the fed back signal $\bar{x}(t)$ at the integrator input is equal to the quantization error. This error is summed up in the integrator and then quantized by the 1-bit A/D converter. Although the quantization error at every sampling instance is large due to the coarse nature of the two level quantizer, the action of the Σ–Δ modulator loop is to generate a ± 1 output which can be averaged over several input sample periods to produce a very precise result. The averaging is performed by the decimation filter which follows the modulator as shown in Figure 6-4.

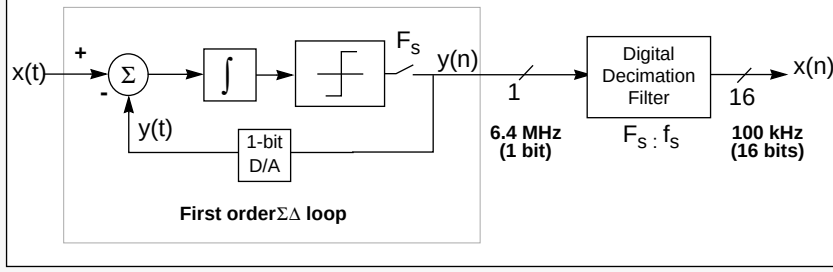


Figure 6-4 *Block Diagram of First-Order Sigma-Delta A/D converter*

The waveforms of x(t) and y(n) for a first-order Σ–Δ modulator are illustrated in Figure 6-5 when the input signal is a sinusoid. The modulator performs both the sampling and the quantization operation in this example, as is typical for circuit implementations of Σ–Δ modulators. In each clock cycle, the value of the output of the modulator is either plus or minus full scale, according to the results of the 1-bit A/D conversion. When the sinusoidal input to the

modulator is close to a plus full scale, the output is positive during most clock cycles. A similar statement holds for the case when the sinusoid is close to minus full scale. In both cases, the local average of the modulator output tracks the analog input. When the input is near zero, the value of the modulator output varies rapidly between a plus and a minus full scale with approximately zero mean.

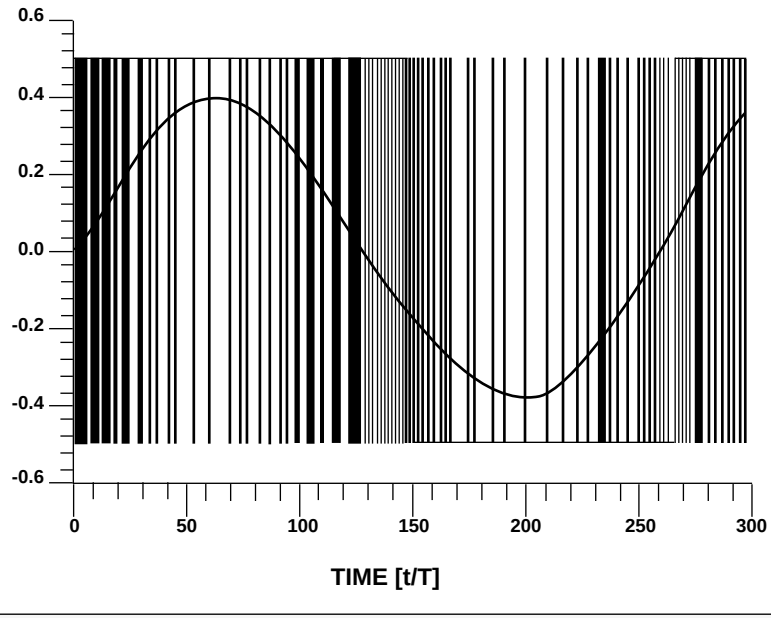


Figure 6-5 *Input and Output of a First-Order Sigma-Delta Modulator*

6.1 Analysis of Sigma-Delta Modulation in the Z-Transform Domain

Consider the first-order loop shown in Figure 6-6. The z-domain transfer function of an integrator is denoted by I(z) and the 1-bit quantizer is modeled as an additive noise source. Standard discrete-time signal analysis yields:

$$Y(z) = Q(z) + I(z)[X(z) - z^{-1}Y(z)]$$

Eqn. 6-2

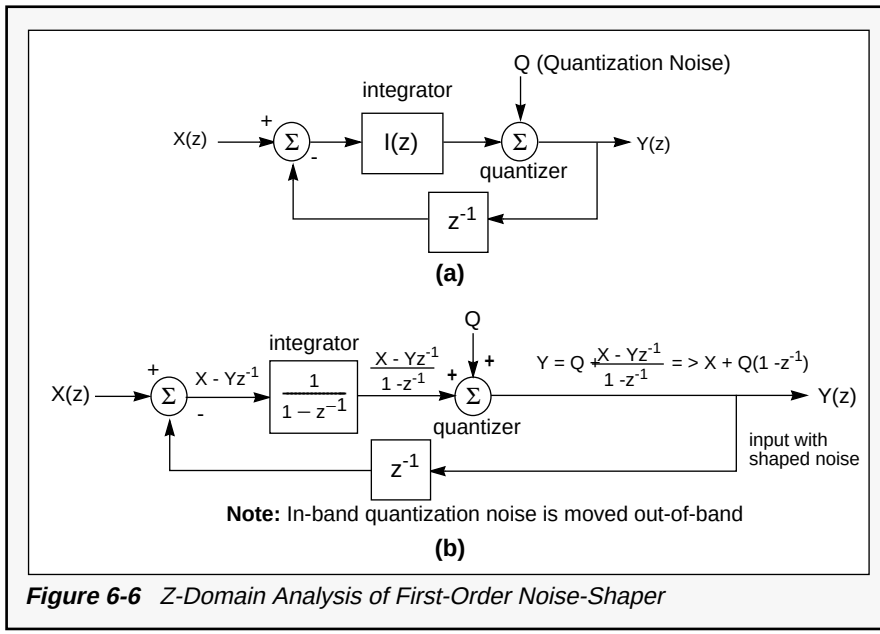


Figure 6-6 *Z-Domain Analysis of First-Order Noise-Shaper*

and can be solved for Y(z) as:

$$Y(z) = X(z)\frac{I(z)}{1 + I(z)z^{-1}} + Q(z)\frac{1}{1 + I(z)z^{-1}} \qquad \text{Eqn. 6-3}$$

Since an ideal integrator is defined as:

$$I(z) = \frac{1}{1 - z^{-1}} \qquad \text{Eqn. 6-4}$$

the first-order Σ–Δ loop output can be simplified to:

$$Y(z) = X(z) + (1 - z^{-1})Q(z) \qquad \text{Eqn. 6-5}$$

Since the quantization noise is assumed to be random, the differentiator ($1\text{-}z^{-1}$) shown in Eqn. 6-5 doubles the power of quantized noise. However, the error has been pushed towards high frequencies due to the differentiator, ($1\text{-}z^{-1}$), factor. Therefore, provided that the analog input signal to the modulator, x(t), is oversampled, the high-frequency quantization noise can be removed by digital lowpass filters without affecting the input signal characteristics residing in baseband. This lowpass filtering is part of the decimation process.

That is, after the digital decimation filtering processes, the output signal has only the frequency components from 0 Hz to f_B. Thus, the performance of first-order Σ–Δ modulators can be compared to the conventional 1-bit Nyquist samplers and the delta-modulation type oversamplers.

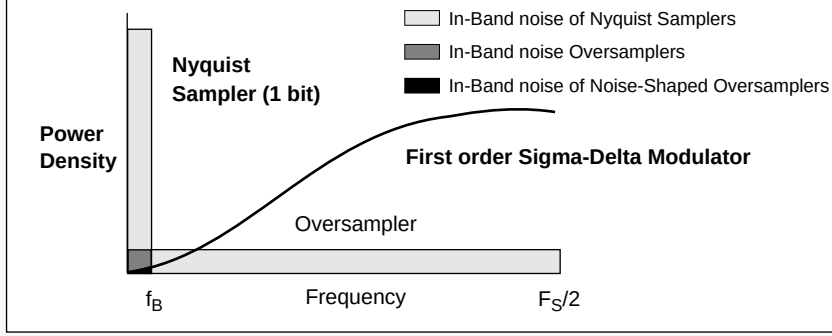


Figure 6-7 *Spectrum of a First-Order Sigma-Delta Noise Shaper*

Figure 6-7 shows the spectrum of a first-order Σ–Δ noise-shaper described in Figure 6-6. As shown in Figure 6-7, the baseband (up to f_B) noise of the Σ–Δ converters appears to be much smaller than Nyquist samplers or delta modulators. However, for the first order modulator discussed, the baseband noise can not reach below the -96 dB signal-to-noise ratio needed for 16-bit A/D converters.

The higher order cascaded (feed-forward) Σ–Δ modulators have been introduced and implemented [8-10]. The block diagrams of second and third order Σ–Δ modulators are shown in Figure 6-8. Since these cascaded structures use a noise feed-forward scheme, the system is always stable and the analysis is simpler compared to the second-order feedback Σ–Δ modulators [11-13] and higher order interpolative coders with feedback loops [14,15]. When multiple first-order Σ–Δ loops are cascaded to obtain higher order modulators, the signal that is passed to the successive loop is the error term from

the current loop. This error is the difference between the integrator output and the quantization output.

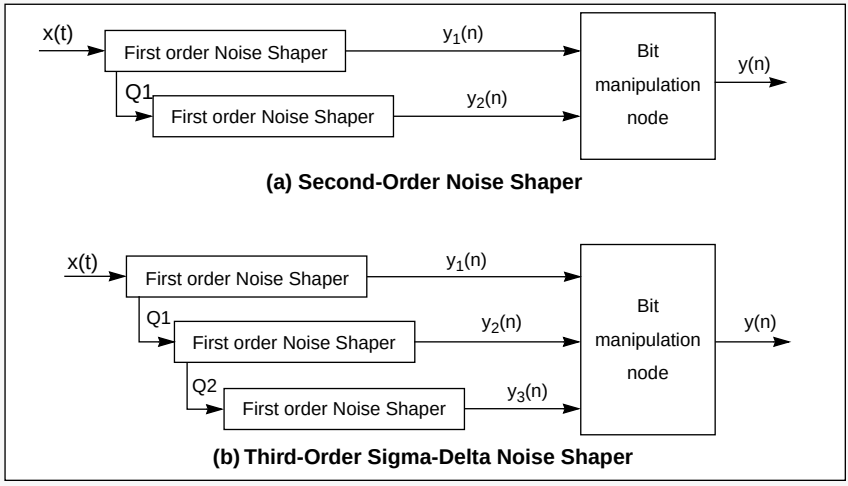


Figure 6-8 *Second-Order and Third-Order Sigma-Delta Noise Shapers*

If the input signals to the second and third stage Σ–Δ loops are Q_1 and Q_2, respectively, the quantization output for the second order Σ–Δ modulator is given by:

$$Y_2(z) = Q_1(z) + (1 - z^{-1})Q_2(z)$$

Eqn. 6-6

which yields the second-order Σ–Δ modulator output shown in Figure 6-8(a) as:

$$Y(z) = X(z) + (1 - z^{-1})^2 Q_2(z)$$

Eqn. 6-7

where: Y(z) is the z-transform of y(n) which is the sampled and quantized signal of x(t) at t = n

Similarly, for the third order Σ–Δ modulator:

$$Y_3(z) = Q_2(z) + (1 - z^{-1})Q_3(z) \qquad \text{Eqn. 6-8}$$

and, after combining the individual output terms we obtain:

$$Y(z) = X(z) + (1 - z^{-1})^3 Q_3(z) \qquad \text{Eqn. 6-9}$$

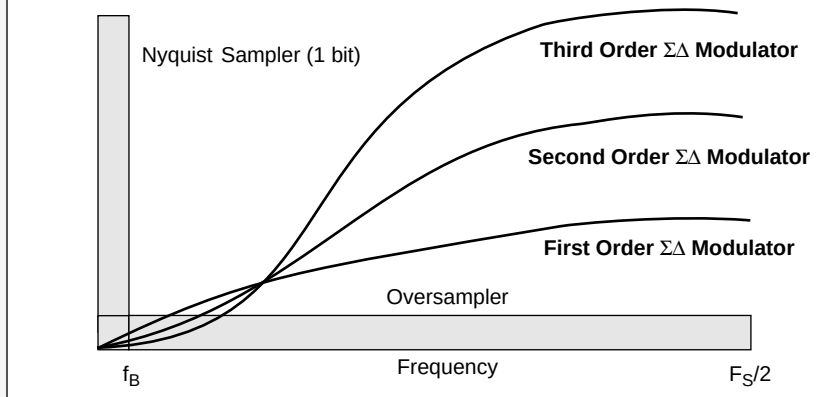
where: *Q_3 is the quantization noise from the third Σ–Δ loop*

Essentially, the noise shaping function in a Σ–Δ modulator is the inverse of the transfer function of the filter $[1\text{-}z^{-1}]^{-1}$ in the forward path of the modulator. A filter with higher gain at low frequencies is expected to provide better baseband attenuation for the noise signal. Therefore, modulators with more than one Σ–Δ loop such as the third-order system shown in Figure 6-8(b), perform a higher order difference operation of the error produced by the quantizer and thus stronger attenuation at low frequencies for the quantization noise signal. The noise shaping functions of second-order and third-order modulators are compared to that of a first-order system in Figure 6-9. The baseband quantization error power for the third-order system is clearly smaller than for the first-order modulator.

The above analysis can be extended to yield quantitative results for the resolution of Σ–Δ modulators, provided that the spectral distribution of the quantization error q(n) is known. It has been shown that the error generated by a scalar quantizer with quantization

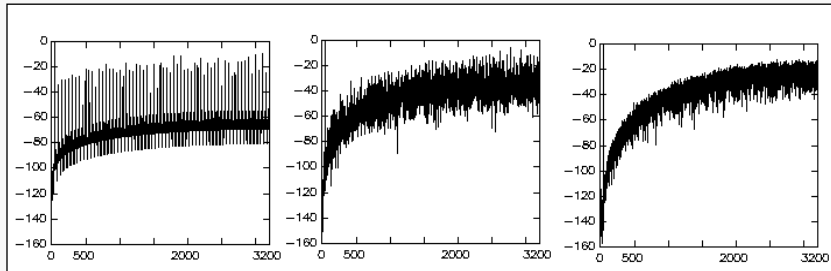
levels equally spaced by q is uncorrelated, assuming that the number of quantization levels is large and the quantized signal is active [16,17]. This result is not rigorously applicable to Σ–Δ modulators, however, because Σ–Δ quantizers only have two levels. Hence, the noise and signal are somewhat correlated. Nevertheless, analysis based on this uncorrelated assumption yields correct results in many cases. Often these analytical results provide a more intuitive interpretation for the operation of the modulator than those obtained from computer simulations. The latter, however, will always be presented to demonstrate the correctness of the analytical result in a particular case.

The performance of this triple cascaded structure can also be compared by inputting a single sinusoidal signal to the structure and plotting the spectra. Figure 6-10 shows the frequency response of the three modulator outputs. The frequency ranges of the x-axis is up to half of the input sampling frequency (3.2 MHz for 6.4 MHz input sampling rate). Note that the frequency of interest is up to 50 kHz which is a very small portion of the plots.



Note: Higher order Noise Shaper has less baseband noise

Figure 6-9 *Multi-Order Sigma-Delta Noise Shapers*



(a) First order sigma-delta (b) Second order sigma-delta (c) Third order sigma-delta

NOTE: Frequency band of interest (in-band): 0 - 5 kHz
Rest of frequency band will be removed by digital decimation filters

Figure 6-10 *Spectra of Three Sigma-Delta Noise Shapers*

SECTION 7

Digital Decimation Filtering

"The simplest and most economical filter to reduce the input sampling rate is a* 'Comb-Filter' *because it does not require a multiplier."

Filtering noise which could be aliased back into the baseband is the primary purpose of the digital filtering stage. Its secondary purpose is to take the 1-bit data stream that has a high sample rate and transform it into a 16-bit data stream at a lower sample rate. This process is known as decimation. Essentially, decimation is both an averaging filter function and a rate reduction function performed simultaneously.

The output of the modulator is a coarse quantization of the analog input. However, the modulator is oversampled at a rate that is as much as 64 times higher than the Nyquist rate for the DSP56ADC16. High resolution is achieved by averaging over 64 data points to interpolate between the coarse quantization levels of the modulator. The process of averaging is equivalent to lowpass filtering in the frequency domain. With the high frequency components of the quantization noise removed, the output sampling rate can be reduced to the Nyquist rate without aliasing noise into the baseband.

Three basic tasks are performed in the digital filter sections:

1. Remove shaped quantization noise: The Σ–Δ modulator is designed to suppress quantization noise in the baseband. Thus, most of the quantization noise is at frequencies above the baseband. The main objective of the digital filter is to remove this out-of-band quantization noise. This leaves a small amount of baseband quantization noise and the band-limited input signal component. Reducing the baseband quantization noise is equivalent to increasing the effective resolution of the digital output.
2. Decimation (sample rate reduction): The output of the Σ–Δ modulator is at a very high sampling rate. This is a fundamental characteristic of Σ–Δ modulators because they use the high frequency portion of the spectrum to place the bulk of the quantization noise. After the high frequency quantization noise is filtered out, it is possible to reduce the sampling rate. It is desirable to bring the sampling rate down to the Nyquist rate which minimizes the amount of information for subsequent transmission, storage, or digital signal processing.
3. Anti-aliasing: In practice, the input signals are seldom completely band-limited. Since the modulator is sampling at a rate much higher than the output Nyquist rate, the analog anti-aliasing filter before the modulator can roll off gradually. When the digital processor reduces the sampling rate down to the Nyquist rate, it needs to provide the necessary additional aliasing rejection for the input signal as opposed to the internally generated quantization noise.

There are a number of factors that make it difficult to implement the digital decimation filter. The input sampling rate of the modulator is very high and the

digital decimation filter must perform computationally intensive signal processing algorithms in real time. Furthermore, higher order modulators produce highly shaped noise as indicated in the spectrum shown in Figure 6-7 and Figure 6-8. Thus, the decimation filter must perform very well to remove the excess quantization noise. Applications like high quality audio conversion impose the additional constraint that the digital signal processing must perform its task without distorting the magnitude and phase characteristics of the input signal in the baseband. The goal is to implement the digital filter in a minimum amount of logic and make it feasible for monolithic implementation.

The simplest and most economical filter to reduce the input sampling rate is a *"Comb-Filter"*, because such a filter does not require a multiplier. A multiplier is not required because the filter coefficients are all unity. This comb-filter operation is equivalent to a rectangular window finite impulse response (FIR) filter. However, the comb-filter is not very effective at removing the large volume of out-of-band quantization noise generated by the Σ–Δ modulators and is seldom used in practice without additional digital filters. Also, the frequency response of the comb-filter can cause substantial magnitude drooping at the upper region of baseband. For many applications which cannot tolerate this distortion, the comb-filter must be used in conjunction with one or more additional digital filter stages.

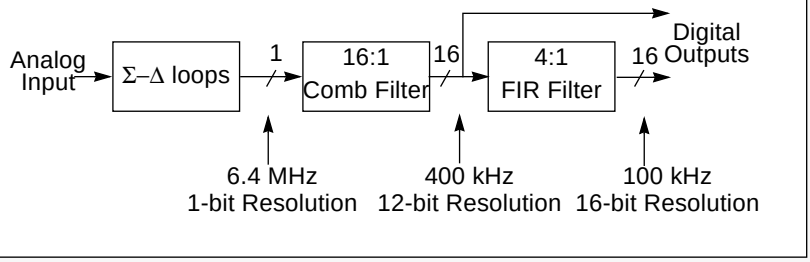


Figure 7-1 *Digital Decimation Process*

In the DSP56ADC16 a total of two filter sections were used as follows [10]:

1. The sampling rate from the Σ–Δ modulator is reduced by a factor of 16 with the comb-filter section as shown in Figure 7-1. A four-stage comb-filter was used to decimate the third order modulator on the DSP56ADC16 [17,18].
2. The second section is a FIR lowpass filter with symmetric coefficient values to maintain a linear-phase response. It provides 4:1 decimation and magnitude compensation for the magnitude change (droop) from the comb-filter output. The FIR filter coefficients are also computed to equalize the baseband frequency response to within ±0.001 dB of a flat response.

7.1 Comb-Filter Design as a Decimator

A comb-filter of length N is a FIR filter with all N coefficients equal to one. The transfer function of a comb-filter is:

$$H(z) = \sum_{n=0}^{N-1} z^{-n} = \frac{Y(z)}{X(z)} \qquad \text{Eqn. 7-1}$$

For N = 4 Eqn. 7-1 becomes:

$$y(n) = x(n) + x(n-1) + x(n-2) + x(n-3) \qquad \text{Eqn. 7-2}$$

Clearly the filter is a simple accumulator which performs a moving average. Using the formula for a geometric sum Eqn. 7-1 can be expressed in closed form as:

$$H(z) = \frac{1 - z^{-N}}{1 - z^{-1}} = \frac{Y(z)}{X(z)} \qquad \text{Eqn. 7-3}$$

or, in the discrete-time domain for N = 4:

$$y(n) = x(n) - x(n-4) + y(n-1) \qquad \text{Eqn. 7-4}$$

Using this recursive form the number of additions has been reduced to become independent of N. The closed form solution in Eqn. 7-3 can be factored into two separate processes-integration followed by differentiation as shown in Eqn. 7-5 :

$$y(z) = \left[\frac{1}{1 - z^{-1}}\right][1 - z^{-N}]x(z) \qquad \text{Eqn. 7-5}$$

Now, since the comb filter will be followed by an N:1 decimator, the differentiation function can be done at the lower rate. This discussion for a general N:1 comb-filter decimator is shown in Figure 7-2. Note that the accumulator can overflow however, as long as the final output does not overflow (i.e., the filter is scaled properly for unity gain) and two's complement or *"wrap around"* arithmetic is used, the accumulator overflow will not cause an error [18,19].

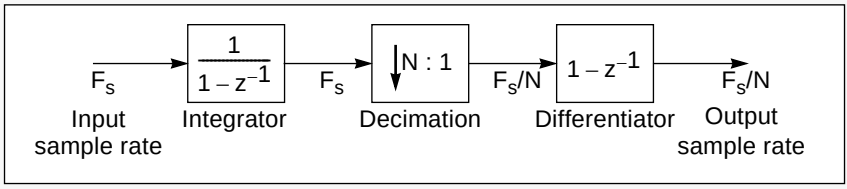


Figure 7-2 *Block Diagram of One-Stage Comb Filtering Process*

The transfer function and magnitude response of a comb-filter with the filter window length N = 16 followed by a 16:1 decimation process are shown in Figure 7-3. In summary the comb-filter decimator has the following advantages:

- *no multipliers are required*
- *no storage is required for filter coefficients*
- *intermediate storage is reduced by integrating at the high sampling rate and differentiating at the low sampling rate, compared to the equivalent implementation using cascaded uniform FIR filters*
- *the structure of comb-filters is very "regular" consisting of two basic building blocks*
- *little external control or complicated local timing is required*

- *the same filter design can easily be used for a wide range of rate change factors, N, with the addition of a scaling circuit and minimal changes to the filter timing*

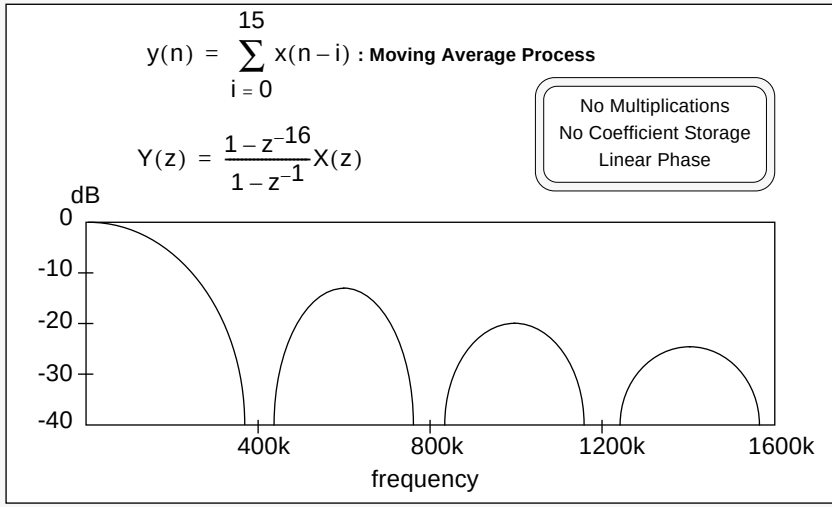


Figure 7-3 *Transfer Function of a Comb-Filter*

A single comb-filter stage usually does not have enough stop-band attenuation in the region of interest to prevent aliasing after decimation. However, cascaded comb-filters can be used to give enough stop-band attenuation. Figure 7-4 shows a structure of four cascaded comb filter sections and the resulting spectrum compared to lower order comb-filter stages. This realization needs eight registers for data and 4(N+1) additions per input for computation.

There are considerable advantages in both storage and arithmetic for this kind of comb decimator realization. This approach is also attractive because the comb decimation filter delivers not only a much faster sample rate with 12 bits of dynamic resolution, but also an untruncated or arithmetically undistorted 16-bit output. **SECTION 8** discusses an application which takes advantage of this comb-filter output.

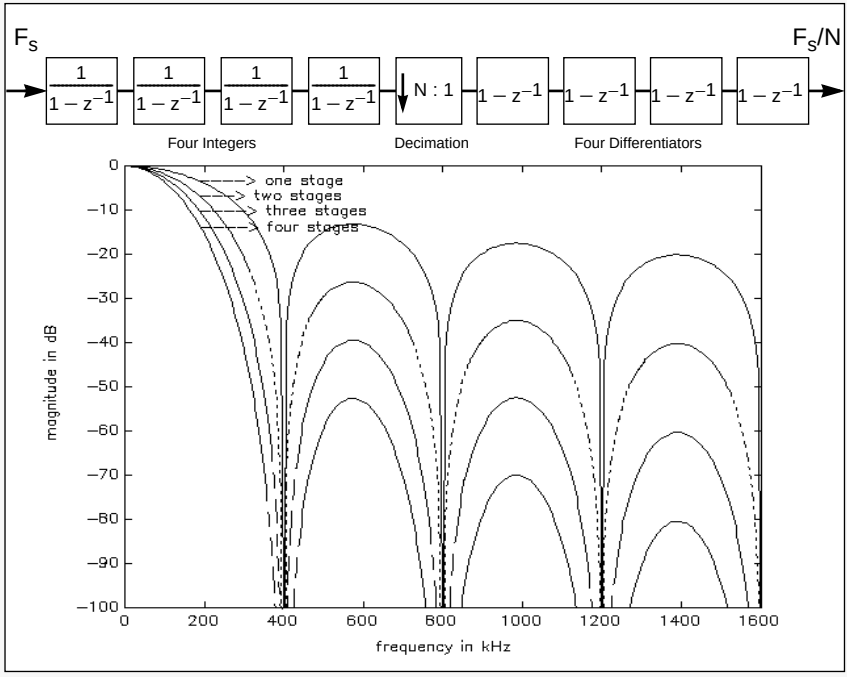
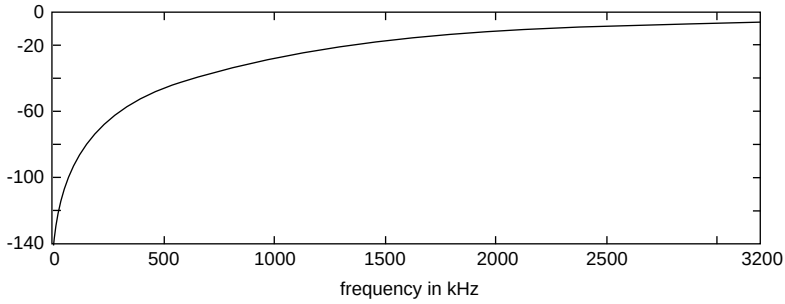
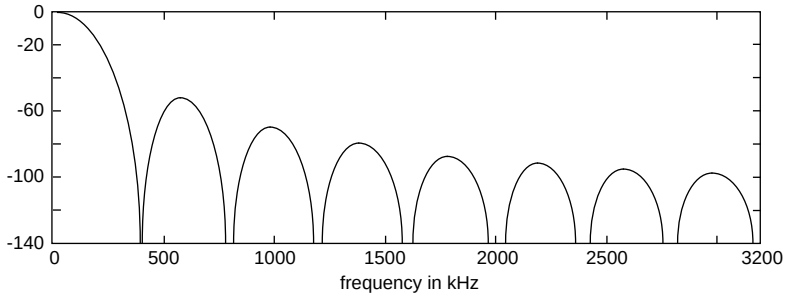


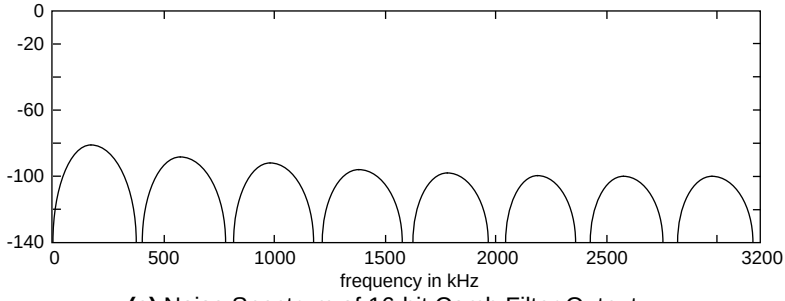
Figure 7-4 *Cascaded Structure of a Comb-Filter*



(a) Noise Spectrum of 1-Bit Input to comb Filter



(b) Spectrum of 4-Stage Comb Filter



(c) Noise Spectrum of 16-bit Comb Filter Output
(-72 dB Noise Power: 12-bit Dynamic Resolution)

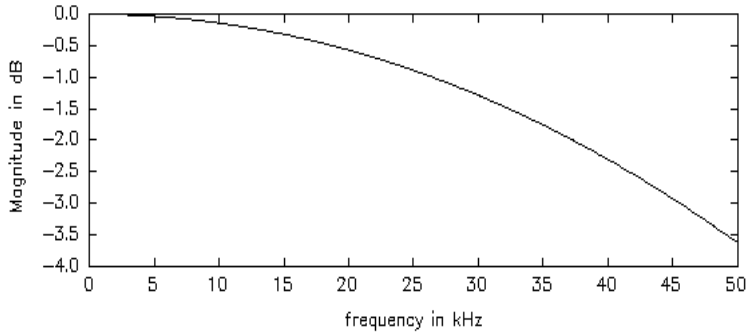
Figure 7-5 *Aliased Noise in Comb-Filter Output*

Figure 7-5(a) shows the third-order Σ–Δ noise shaper output spectrum; the four-stage comb-filter magnitude response is illustrated in Figure 7-5(b). Thus, the theoretical quantized noise filtered by the four-stage comb-filter can be computed as shown in Figure 7-5(c). After a 16:1 decimation process, the noise in the frequency band from 200 kHz to 1.6MHz will be aliased back to baseband of up to 200 kHz. The numerical summation of total noise in Figure 7-5(c) is around -72 dB S/N ratio which yields a linear 12-bit dynamic resolution.

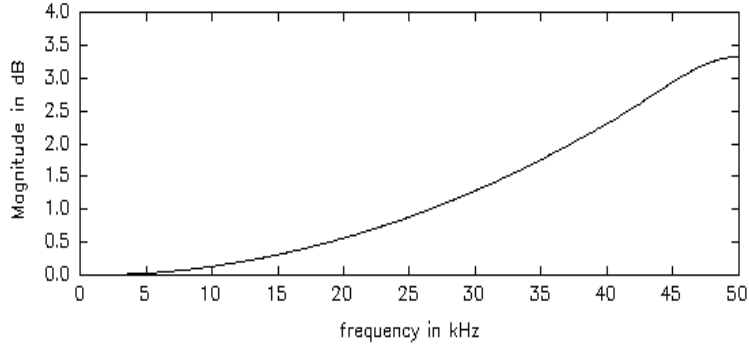
7.2 Second Section Decimation FIR Filter

The first comb-filter section decimates the Σ–Δ modulator output to the intermediate rate of 400 kHz (assuming a master clock at 12.8 MHz) and the second filter section provides the sharp filtering necessary to reduce the frequency aliasing effect when the sampling rate is decimated down to 100 kHz while the resolution becomes 16 bit. The second filter section has two different tasks to perform. The first task is 4:1 decimation lowpass filtering with 0.001 dB ripple up to 45.5 kHz and more than -96 dB cutoff attenuation above 50 kHz. The stop-band attenuation of -96 dB is sufficient because the quantization noise level has not yet reached its full magnitude as shown in Figure 6-10(c). The second task is the passband response compensation for the droop introduced by the *"comb-filter" s*ection. Figure 7-6(a) shows the magnitude droop

by the fourth order cascaded comb filter section, while Figure 7-6(b) shows the compensation filter response up to cutout frequency.



(a) *Comb-Filter Magnitude Response in Baseband*



(b) *Compensation FIR Filter Magnitude Response*

Figure 7-6 *Magnitude Response*

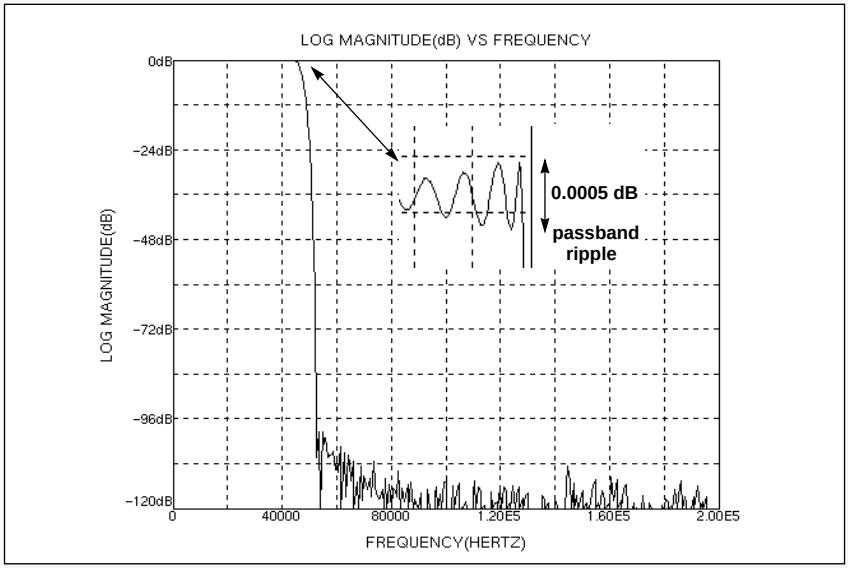


Figure 7-7 *FIR Filter Magnitude Response*

Although FIR filters take more time to perform the decimation filtering process than infinite impulse response (IIR) filters for the same passband and out-of-band frequency characteristics, FIR filters can be designed to have a linear-phase response which is required for high-end audio and instrumentation applications.

Design techniques for calculating the FIR filter coefficients with linear-phase response are well documented [5,20]. Figure 7-7 shows the frequency response of a low-pass FIR filter with 255 symmetric coefficients using a computer optimized procedure. This filter provides a stop-band attenuation of -96 dB with a maximum passband ripple of 0.001dB which meets the filter requirement as defined in the **DSP56ADC16 datasheet** [21]. However, the

actual second section FIR filter coefficients are computed by the combination of the compensation filter and the lowpass filter.

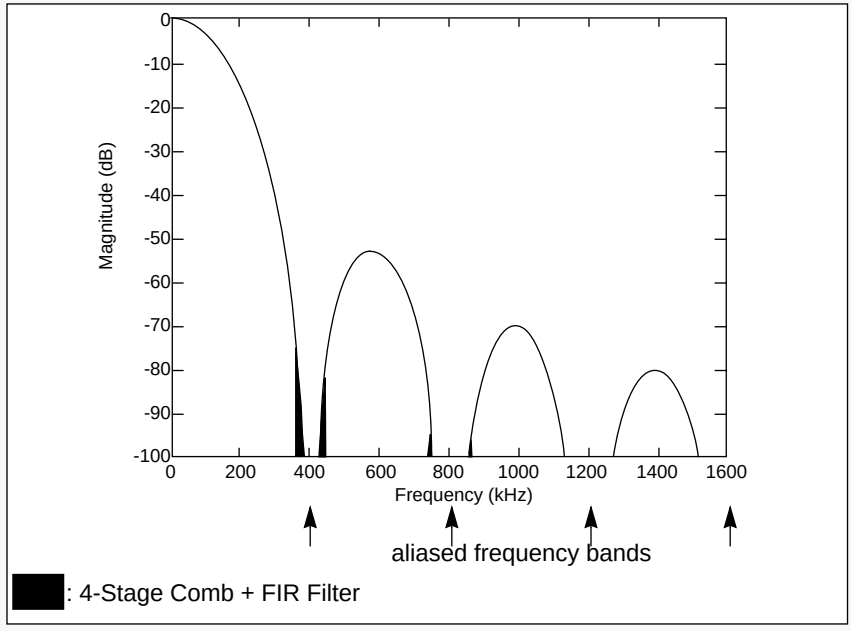


Figure 7-8 *Aliased Noise Bands of FIR Filter Output*

Figure 7-8 shows the frequency bands and the final quantized noise level which would have aliased back to baseband after 64:1 decimation by the comb-filter section and the second 4:1 decimation section. When the one-pole RC analog lowpass filter is implemented at the analog input terminals as discussed in **SECTION 4**, the power of the aliased frequency band shown in Figure 7-8 will be less than -96 dB. In fact, the theoretical and numerical computations show that total aliased power after the final decimation process is around -110 dB S/N ratio. ■

SECTION 8

Mode Resolution by Filtering the Comb-Filter Output with Half-Band Filters

"The half-band filter is based on a symmetrical FIR design and approximately half of the filter coefficients are exactly zero. Hence, the number of multiplications in implementing such filters is one-fourth of that needed for arbitrary FIR filter designs."

This section explains how to obtain 18-20 bit resolution from the comb-filter output. In particular, a series half-band filter structure is suggested to take advantage of a third-order noise-shaper combined with oversampling. Figure 8-1 shows the spectrum of a third-order noise-shaper whose input-output equation is defined in Eqn. 6-8. The transfer function $(1 - z^{-1})^3$ in Eqn. 6-8 for the quantization noise is essentially a highpass filter so that the noise is shifted to higher, out-of-band, frequencies where it is then digitally filtered out.

At the final arithmetic operation for the FIR filter output, a 38-bit accumulator value is convergently rounded to fit into the 16-bit output data format. Thus, the shaped noise shown in Figure 8-1 becomes almost flat at the 16-bit level due to the arithmetic rounding noise. Although the comb-filter output has only 12-bit resolution, the 16-bit output value is not yet arithmetically rounded or truncated. Thus, the input

signal and the baseband spectral shape of the third-order noise shaper are unchanged, as shown in Figure 8-2. Since the comb-filter is designed to obtain maximum attenuation only on the higher frequency components which will be aliased into the frequency band of interest after 16:1 decimation [18], the characteristic of the analog input signal is preserved, while the out-of-band shaped noise shown in Figure 8-1 has been attenuated.

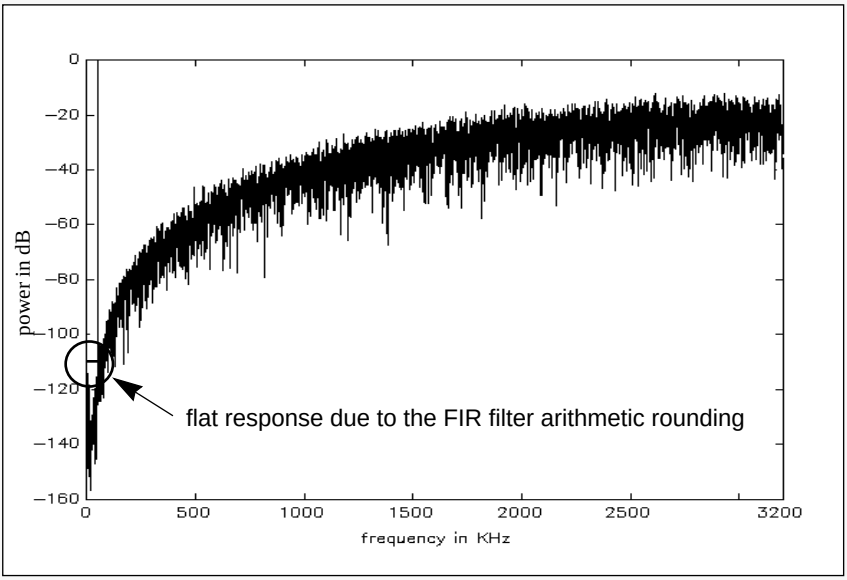


Figure 8-1 *Spectrum of a Third-Order Noise Shaper (16384 FFT bins)*

The output from most conventional converters including the FIR filter output of the DSP56ADC16, has a flat background noise due to the quantization noise as well as arithmetic rounding noise. Thus, further decimation processes can only gain 3 dB or

1/2 bit more resolution per octave. In other words, a further 16:1 or 256:1 decimation process of the 16-bit resolution output signal is required to obtain 18-bit resolution at 6.25 kHz sample rate or 20-bit resolution at 400 Hz sample rate, respectively, which is very impractical. By taking advantage of the fact that the noise is shaped in the comb-filter output, 9 dB or 1.5 bit more resolution per octave can be theoretically achieved. Thus, a further 16:1 or 64:1 decimation process can provide 18-bit resolution at 25 kHz sample rate or 20-bit resolution at 6.25 kHz sample rate, respectively. (This assumes quantization noise is dominant.)

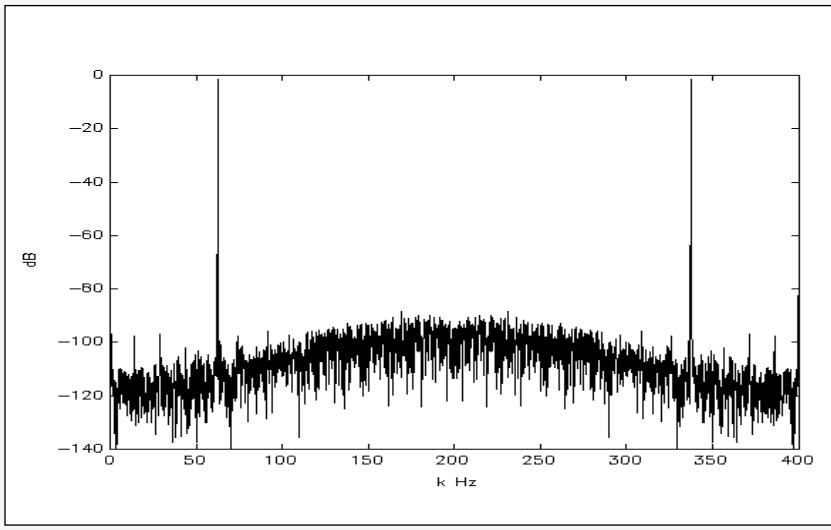


Figure 8-2 *Spectrum of Typical Comb-Filter Output (4096 FFT bins)*

Figure 8-3 illustrates the cascaded half-band filter design specification for a 64:1 decimation process. The number of instructions to run the filters on the DSP56001 and the memory requirements are tabulated in Table 8-1. A cascaded half-band filter structure is used for computational simplicity [22]. The half-band filter is based on a symmetrical FIR design and approximately half of the filter coefficients are exactly zero. Hence, the number of multiplications in implementing such filters is one-fourth of that needed for arbitrary FIR filter designs. Since a half-band filter can only implement a 2:1 decimation, a series of such filters may be cascaded to perform a higher decimation filter process.

Table 8-1 *Parameters for Designing Half-Band Filters*

Stage	Output sample rate	Passband	Stopband	# of taps	# of coef in RAM	# of MAC	# of instructions per second
1	200k	3k	197k	19	6	11	11x200k=2.2M
2	100k	3k	97k	19	6	11	11x100k=1.1M
3	50k	3k	47k	19	6	11	11x50k=550K
4	25k	3k	22k	23	7	13	13x25k=325K
5	12.5k	3k	9.5k	31	9	17	17x12.5k=212.5K
6	6.25k	2.9k	3.55k	199	51	99	99x6.25k=618.75 K
Total number of instructions for 6 half-band filters per second : 5.007 MIPS							

To obtain more than a 16-bit resolution signal output, a processor with more than 16-bit coefficient word-width is required. Fortunately, the DSP56001 general purpose DSP processor has a hardware multiply-accumulate unit that is able to multiply 24-bit data and 24-bit coefficient, and accumulate 56 bits in just one instruction, which is useful for half-band filter operations. The DSP56001 architecture also provides parallel data buses, circular buffers, and large on-chip memories along with a 75 ns instruction cycle, which fits nicely for the proposed filter structure [23]. Detailed discussion on this topic can be found in [24].

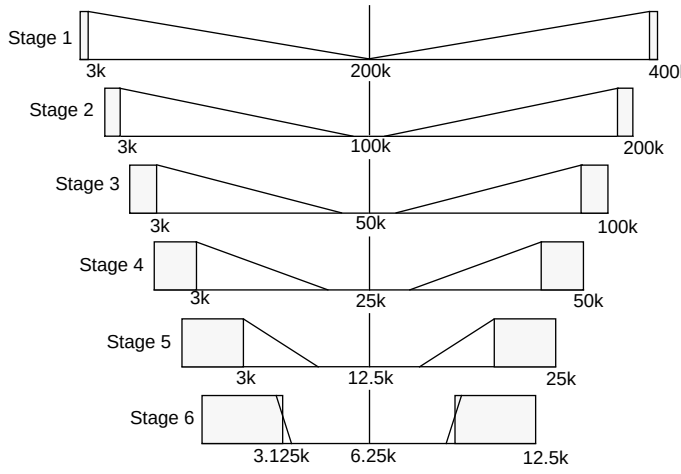


Figure 8-3 *Decimation Process using a Series of Half-Band Filters*

■

SECTION 9

Summary

". . . Σ–Δ technology offers system cost savings because the analog anti-aliasing filter requirements are considerably less complex . . ."

Sigma-Delta conversion technology is based on oversampling, noise shaping, and decimation filtering. There are many inherent advantages in Σ–Δ based analog-to-digital converters. The major advantage being that it is based predominantly on digital signal processing, hence the cost of implementation is low and will continue to decrease. Also, due to its digital nature Σ–Δ converters can be integrated onto other digital devices. Manufacturing technology notwithstanding, Σ–Δ technology offers system cost savings because the analog anti-aliasing filter requirements are considerably less complex and the sample-and-hold circuit is intrinsic to the technology due to the high input sampling rate and the low precision A/D conversion. Since the digital filtering stages reside behind the A/D conversion, noise injected during the conversion process, such as power-supply ripple, voltage-reference noise, or noise in the ADC itself, can be controlled. Also, Σ–Δ converters are inherently linear and don't suffer from appreciable differential non-linearity, and the background noise level which sets the system S/N ratio is independent of the input signal level. The last, but certainly not least, consideration is cost. Attaining a high level of performance at a fraction of the cost of hybrid and modular designs is probably the greatest advantage of all. ■

REFERENCES

[1] H. Inose, Y. Yasuda and J. Marakami, “A telemetering system by code modulation, delta-sigma modulation,” *IRE Trans. on Space, Electronics and Telemetry*, SET-8, pp. 204-209, Sept. 1962.

[2] H. Nyquist, “Certain topics in telegraph transmission theory,” *AIEE Trans.*, pp. 617-644, 1928.

[3] M. Armstrong, et al, “A COMS programmable self-calibrating 13b eight-channel analog interface processor,” *ISSCC Dig. Tech. Paper*, pp. 44-45, Feb. 1987.

[4] K. Lakshmikumar, R. Hadaway, and M. Copeland, “Characterization and modeling of mismatch in MOS transistors for precision analog design,” *IEEE J. Solid-State Circuits*, Vol. SC-21, pp. 1057-1066, Dec. 1986.

[5] N. Ahmed and T. Natarajan, *Discrete-Time Signals and Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[6] R. Steele, *Delta Modulation Systems*, Pentech Press, London, England, 1975.

[7] N. Scheinberg and D. Schilling, “Techniques for correcting transmission error in video adaptive delta-modulation channels,” *IEEE Trans. Commun.*, pp. 1064-1070, Sept. 1977.

[8] Y. Matsuya, et al, “A 16 bit oversampling A-to-D conversion technology using triple-integration noise shaping,” *IEEE J. of Solid-State Circuits*, Vol. SC-22, No. 6, pp. 921-929, Dec. 1987.

[9] M. Rebeschini et al, “A high-resolution CMOS Sigma-Delta A/D converter with 320 kHz output rate,” *Proc. ISCAS*, pp. 246-249, 1989.

[10] C. D. Thompson, “A VLSI sigma delta A/D converter for audio and signal processing applications,” *Proc. Int. Conf. on Acoustic, Speech and Signal Processing*, Vol. 4. pp. 2569-2572, Glasgow, Scotland, May 23-26, 1989.

[11] J. C. Candy, “A use of double integration in Sigma-Delta modulation,” *IEEE Trans. on Communications*, Vol. COM-33, No. 3, pp. 249-258, March 1985.

[12] R. Koch, et al, “A 12-bit Sigma-Delta analog-to-digital converter with a 15-MHz clock rate,” *IEEE J. of Solid-State Circuits*, Vol. SC-21, No. 6, pp. 1003-1010, Dec. 1986.

[13] B. Boser and B. Wooley, “Quantization error spectrum of Sigma-Delta modulators,” *Proc. International Symposium on Circuits and Systems*, pp. 2331-2334, June, 1988.

[14] W. L. Lee and C. G. Sodini, “A topology for higher order interpolative coders,” *Proc. International Symposium on Circuits and Systems*, pp. 459-462, May 1987.

[15] D. R. Welland, et al, “A stereo 16-bit delta-sigma A/D converter for digital audio,” *Proc. the 85th convention of Audio Engineering Society*, Vol. 2724 (H-12), Los Angeles, CA, Nov. 3-6, 1988.

[16] W. Bennett, “Spectra of quantized signals,” *Bell Syst. Tech. J.*, Vol. BSTJ-27, pp. 446-472, July 1948.

[17] B. Widrow, “A study of rough amplitude quantization by means of Nyquist sampling theory,” *IRE Trans. Circuit Theory*, Vol. CT-3, pp. 266-276, Dec. 1956.

[18] E. B. Hogenauer, “An economical class of digital filters for decimation and interpolation,”, *IEEE Trans, Acoust. Speech Signal Process.*, Vol. ASSP-29, No. 2, pp. 155-162, April 1981.

[19] S. Chu and C. S. Burrus, “Multirate filter designs using comb filters,” *IEEE Trans. on Circuits and Systems*, Vol. CAS-31, No. 11, pp. 913-924, Nov. 1984.

[20] A. V. Oppenheim and R. W. Schafer, *Discrete-time signal processing*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[21] Motorola Inc., *DSP56ADC16 Technical Data Sheet* (DSP56ADC16/D), 1989.

[22] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[23] Motorola Inc., *DSP56000/DSP56001 Digital Signal Processor User’s Manual*, Rev. 2, 1990.

[24] S. Park, “A real-time implementation of half-band filters to obtain 18-20 bit resolution from the DSP56ADC16 Sigma-Delta A/D converter,” *Proc. International Conference on Acoustics, Speech and Signal Processing*, Vol. 2. pp. 989-992, Albuquerque, NM, April 3-6, 1990. ■