

**A FUNDAMENTAL INTRODUCTION  
TO THE COMPACT DISC PLAYER**

Grant M. Erickson  
Department of Electrical Engineering  
University of Minnesota

November 29, 1994  
EE 3011  
Professor: Dr. Kevin M. Buckley

## TABLE OF CONTENTS

THE NEED FOR DIGITAL AUDIO.....	1
Strengths of the Digital Domain .....	1
Developments Facilitating the Compact Disc Player .....	2
PRINCIPLES OF DIGITAL AUDIO .....	3
Sampling.....	3
Aliasing .....	4
Quantization.....	5
Dither .....	6
Jitter.....	7
IMPLEMENTATION.....	8
System Overview .....	8
Digital-to-analog Converters.....	10
Multi-bit Converters.....	10
Low-bit Converters.....	11
Digital Filtering, Oversampling, and Noise Shaping .....	14
REFERENCES.....	19

## LIST OF FIGURES

Table 1.	Red Book specifications for the compact disc system.....	9
Figure 1.	Block diagram of a compact disc player.....	9
Figure 2.	Block diagram of a PWM/MASH digital-to-analog converter.....	12
Figure 3.	Block diagram of a PDM digital-to-analog converter.....	13
Figure 4.	Use of a transversal filter to achieve oversampling .....	15
Figure 5.	Effect of zeros interleaving and oversampling on a signal. ....	15
Figure 6a.	Various noise-density distributions as a function of frequency. ....	16
Figure 6b.	The PDM and PWM distributions in the audio band. ....	16
Figure 7.	Third-order noise shaper used in the PWM converter.....	17

## **A Fundamental Introduction to the Compact Disc Player**

The compact disc player has become one of the most ubiquitous pieces of consumer electronics equipment in use today. Tens of millions of players have been sold to date. However, as pervasive as the compact disc player's presence is, the beauty and complexity of its design and operation are underappreciated by most users. This brief text attempts to inform the reader of the basic fundamentals of the compact disc player. It is assumed that the reader has a basic knowledge of the fundamentals of signal processing, although it is certainly not a prerequisite for learning a great deal from the reading.

### **THE NEED FOR DIGITAL AUDIO**

#### ***Strengths of the Digital Domain***

Since Thomas Edison made the first audio recording on a foil covered cylinder in 1877, the field of audio recording has grown and matured. Edison's process and many others that followed were all based on a common process; the reproduction of an audio signal from a mechanical or electrical contact with the recording media—this is the realm of analog audio. After nearly 100 years, analog audio has reached a mature state and nearly all of its shortcomings have been addressed to the point that further improvements become financially prohibitive for the average consumer.

The very nature of the analog signal leads to its own shortcomings. In the analog domain, *any* waveform is allowable; therefore the playback mechanism has no means to differentiate noise and distortion from the original signal. Further, in an analog system every copy made introduces more noise than its parent. This fact is due to both the playback and recording mechanism which must physically contact the media, further damaging it after every pass. Every analog system also carries the side effect that the total system noise is the summation of all distortion and noise from each component in the signal path. Finally, analog equipment is of limited performance, exhibiting: an uneven frequency response (which requires extensive equalization), a limited 60 dB dynamic range, and a 30 dB channel separation—which affects stereo imaging and staging.

The need for a new audio format is apparent, and digital audio fills the performance shortcomings of analog audio. The beauty of the digital audio signal is that noise and distortion can be separated from the audio signal. A digital audio signal's quality is not a function of the reading mechanism nor the media in a properly engineered system. Performance parameters such as frequency response, linearity, and noise are only functions of the digital-to-analog converter (DAC). Performance parameters indicative of a digital audio system include full audio band frequency response of 5 ~ 22,000 Hz, 90+ dB dynamic range, and a flat response across the entire audio band.

The final strength of digital audio is the circuitry upon which it is built. First, due to a large degree of circuit integration digital circuits do not degrade with time as analog circuits do. Further, for all practical purposes, a digital signal will suffer no degradation until distortion and noise has become so great that the signal is out of its voltage threshold. However, this threshold has been made intentionally large expressly for this reason. The high level of circuit integration also means that for the same given task, the digital circuitry will cost far less than its analog counterpart.

The only real theoretical limitation to the accuracy of a digital signal is the quantity of numbers in the signal representation and the accuracy of those numbers. These are both known and controllable design parameters.

### ***Developments Facilitating the Compact Disc Player***

As staggering as the release of the compact disc player was in 1982, the technology and theories which allowed it to be born were long in development. In 1841, the great mathematician Augustin-Louis Cauchy first proposes the sampling theorem. Nearly 80 years later J.R. Carson publishes a mathematical analysis of time sampling in communications. In a 1928 lecture at the American Institute of Electrical Engineers Harry Nyquist provides proof of the sampling theorem in "Certain Topics in Telegraph Transmission Theory". In 1937, A. Reeves proposes pulse code wave modulation (PCM). In 1948, John Bardeen, William Shockley, and Walter Brattain invent the bipolar junction transistor at Bell Labs—compact digital circuitry is a reality. Two years later, in 1950 Richard W. Hamming publishes significant work on error correction and detection codes. In 1958 C.H. Townes and A.L. Shawlow invent the laser. In 1960 R.C. Bose

publishes binary group error correction codes. That same year I.S. Reed and G. Solomon publish error correction codes to be used in the CD player 22 years later. Also early computer music experiments take place at Bell Labs. Fifteen years before consumers see the first player, NHK Technical Research Institute publicly demonstrates a PCM digital audio recorder with a 30 kHz sampling rate and 12-bit resolution. Two years later, Sony Corporation demonstrates a PCM digital audio recorder with a 47.25 kHz sampling rate and 13-bit resolution. A hemisphere away, Dutch physicist Klaas Compaan uses a glass disc to store black and white holographic images using frequency modulation at Philips Laboratories. Four years later, in 1973 Philips engineers begin to contemplate an audio application for their “video” disc system. A prototype disc with a 44 kHz sampling rate is run through a 14-bit digital-to-analog converter and exhibits a signal-to-noise (S/N) ratio of 80 dB in monaural. Now a research frontier, Mitsubishi, Sony, and Hitachi all demonstrate digital audio discs at the Tokyo Audio Fair in 1977. One year later, Philips joins with its recording subsidiary Polygram Records to develop a worldwide digital audio standard. In March 1979, Philips demonstrates a prototype compact disc player in Europe. Sony joins the Philips/Polygram coalition after Matsushita declines. In June of 1980, the coalition formally proposes their CD standard. A year later in 1981, Sharp successfully mass produces the semiconductor laser. This step was crucial to delivering a consumer product. In Fall of 1982 nearly 150 years of work comes to fruition and Sony and Philips introduce their respective players to consumer in Europe. The following spring, the player is introduced in the United States. Twelve years later, the improvement of digital audio continues at a rapid pace and the analog format that was so prevalent in 1982 has all but disappeared.

## **PRINCIPLES OF DIGITAL AUDIO**

### ***Sampling***

Given an analog audio signal, a process is needed to bring it into the digital domain. This process is *sampling*, and it is dictated by the Nyquist sampling theorem which states how quickly samples must be taken to ensure an accurate representation of the analog signal.

The sampling theorem is quite simple. It states that the sampling frequency must be greater than or equal to the highest frequency in the original analog

signal. The relationship is given by Equation 1; note that the theorem can also be expressed in terms of the sampling period.

$$f_s \geq 2f \text{ or } T_s \leq \frac{T}{2} \quad (1)$$

The sampling theorem is simple enough, but to use it in a digital audio system, two constraints must be observed. The first is that the original signal must be bandlimited to half the sampling frequency by being passed through an ideal low-pass filter; the second is that the output signal must again be passed through an ideal low-pass filter to reproduce the analog signal. These constraints are crucial to sampling, and if not observed will lead to an unwanted effect known as *aliasing*.

### ***Aliasing***

Aliasing is a system's erroneous response that manifests itself when the constraints of the sampling theorem are not observed. Aliasing will surface in the audio signal as audible distortion. For the limiting case of a frequency at exactly half the sampling frequency, there will be only two samples generated—this is the minimum required to represent any waveform. For signals greater than  $f_s/2$ , the process of sampling can be thought of as modulating the input signal. The modulation creates image frequencies centered around integer multiples of  $f_s$ . These newly generated frequencies are then imaged or aliased back into the audible band. The frequency to which these will be aliased to can be computed by Eq. 2, where  $f_a$  is the alias frequency,  $f$  is the actual frequency,  $f_s$  is the sampling frequency, and  $k$  is an odd integer that satisfies the inequality.

$$f_a = \left| f - \frac{(k+1)f_s}{2} \right| \quad \frac{kf_s}{2} \leq f \leq \frac{(k+2)f_s}{2} \quad (2)$$

We can then easily compute for a sampling rate of 44.1 kHz, a signal of 23 kHz will be aliased to 21.1 kHz. More precisely, the frequency will be folded back across half the sampling frequency by the amount it exceed half the sampling frequency—in this case by 950 Hz.

Hence the use of a brickwall filter—one with a sharp cutoff characteristic—on the input signal is necessary. The need for placing a filter after the DAC in the player may not be intuitively obvious. Imagine the limiting case of a sine wave at half the sampling frequency. There will be two samples generated for this wave,

however the DAC will represent this as a square wave of the same frequency. From the Fourier series expansion, we know that a square wave consists of infinite harmonics. The DAC has now created frequencies that did not previously exist. Because the input signal was bandlimited, we know that it is reasonable to pass the output signal through another low-pass filter with the same characteristic as that used in the sampling process. This low-pass filter strips the higher-order harmonics from the square wave and we are left with the sine wave we started with. Due to its actions, this low-pass filter is often referred to as an anti-aliasing filter in the frequency domain and as a reconstruction filter in the time domain. A linear phase low-pass filter is characterized by having a symmetrical impulse response. In particular, the impulse response of a low-pass filter is the  $\sin(x)/x$  function. When the reconstruction filter is excited by an amplitude varying impulse train from the DAC, the output is a linear combination of the individual amplitude modulated impulse responses.

### ***Quantization***

Once sampling has taken place, we are far from done converting the analog signal to a digital one. In order to represent each sample as a binary series of bits, the infinitely varying voltage amplitude of the analog signal must be assigned a discrete value. This process of assignment is known as *quantization*. It is important to note that quantization and sampling are complementary processes. If we sample the time axis, then we must quantize the amplitude axis and vice versa. It is unfortunately common practice to refer to sampling and quantization as just quantization; this is, however, incorrect. The combined process is referred to as digitization.

In a 16-bit audio format, we can represent a sinusoidally varying voltage audio signal by  $2^{16}$  or 65,536 discrete levels. It is apparent then that quantization is a limiting performance factor in the overall digital audio system, by the number of bits allowed to the quantizing system. The system designer is faced with determining how many bits create a sufficient model of the original signal.

Because of this limiting design factor, quantizing is ideally imperfect in its signal representation, whereas sampling is theoretically perfect. There is then an error inherent in the quantization process regardless of the ideality of the rest of the system.

To visualize what this error is, imagine a digital thermometer on your oven. When the temperature reads 425° F, that value may or may not be accurate. The temperature in the oven may indeed be 425°, but it might also be as much as 425.4° or as little as 424.5°. A similar occurrence occurs with the quantizer in digital audio equipment. While quantizing, it determines the level in which the voltage for a given sample belongs. This quantized level may differ by as much as  $\pm \frac{Q}{2}$ , where  $Q$  is the width of the quantized level.

It is the difference between the actual voltage to be represented and the quantized voltage level that induces quantization error. The magnitude of the error may never exceed the voltage represented by “half” of the least-significant bit (LSB) in the data word. A measurement of the error in a digitization system can be made, and it is expressed as the signal-to-error (S/E) ratio. This ratio is given by Eq. 3, where  $n$  is the number of bits in the data word.

$$S/E(dB) = 6.02n + 1.76 \quad (3)$$

Hence, the theoretical S/E ratio for a 16-bit system is 98 dB. Keep in mind that this value is strictly theoretical and will be lowered and raised by many other performance parameters. For the most part, quantization error manifests itself as noise at high signal levels. However, quantization error becomes quite significant when a low-level signal approaches the level of the LSB, then the quantizing error actually *becomes* the signal, and therefore is an audible component of the output. In many types of music, these types of signals are common and distortion caused by quantization error is both unacceptable and irremovable. Fortunately, in practical systems this adverse effect can be effectively eliminated through the use of *dither*.

### ***Dither***

Dither is the process of adding low-level analog noise to a signal, to randomize or “confuse” the quantizer’s small-signal behavior. Dither specifically aims to address two problems in quantization. The first of which is that a reverberating, decaying signal can fall below the lower limit of the system resolution. That is to say that an attempt to encode a signal below the LSB results in nothing getting encoded. Clearly, information is lost. The second, as discussed in the previous section, is that system distortion increases as a percent of a decreasing input

signal. It is important to note that not only does dither remove some quantization error from the signal, it effectively *removes* it.

The concept might seem initially counterintuitive, but it is really quite simple. Dither relies on some special behavior of the human ear. The ear can detect a signal masked by particularly broadband noise. In some cases, the ear can easily detect a midrange signal buried as much as 10 to 12 dB *below* the level of broadband noise<sup>1</sup>. Those who still find the effects of dither questionable, might want to try the following interesting test<sup>2</sup>.

Let the text on this page represent the amplitude of the signal to be quantized. Also, let the space between your slightly spread fingers represent valid quantization intervals. Now place your hand across the text. Amplitude information has been irrecoverably lost due to quantization. Now provide dither to the signal by quickly moving your hand up and down along the plane of the page. The amplitude information that was lost has been retrieved at the expense of adding a slight amount of noise to the system—your blurred fingers.

So even though some noise has been added, we have eliminated the distortion due to quantization error with the result being a cleaner, more accurate signal.

### ***Jitter***

Although rarely observed in a well designed player, *jitter* is a worthy topic of discussion because of both its misconceptions and the large amount of press it has received. Jitter is basically defined as time instability. It occurs in both analog-to-digital and digital-to-analog conversion. The latter instance is the only concern here. Jitter occurs in the compact disc player when samples are being read off the disc. These reads are controlled by the pulses of a crystal oscillator. If the system clock pulse inaccurately (an unlikely event), if there is a glitch in the digital hardware, or if there is noise on a signal control line, the actual reading time will vary from sample to sample thus inducing noise and distortion in the extreme case.

A great deal of money has been made by shrewd marketeers preying on the fears of the consumer worried about jitter. Such products marketed include disc stabilizer rings to reduce rotational variations, highly damped rubber feet for the players, and other snake oil remedies. However, the careful engineer has beaten

the marketer to the punch by having the samples read off the disc into a RAM buffer. As the buffer becomes full, a local crystal oscillator can then “clock-out” the samples in a reliable manner, independent of the transport and reading mechanisms. This process is referred to as timebase correction and as stated before, any quality piece of equipment will implement it.

## IMPLEMENTATION

### *System Overview*

The compact disc player as a sound reproduction device fulfills the loop begun in the recording studio, returning the audio signal back to its original analog form. If all the theoretical guidelines have been followed in the equipment and processes between the musician and your audio system, the sound you hear is exactly the sound that was heard in the recording studio.

The specifications for the compact disc and compact disc players were jointly developed by Sony, Philips, and Polygram as mentioned previously. This specification is contained in their standards document referred to as the *Red Book*. A summary of this standard is seen in Table 1.

---

<b>DISC</b>	
Playing time:	74 minutes, 33 seconds maximum
Rotation:	Counter-clockwise when viewed from readout surface
Rotational speed:	1.2–1.4 m/sec. (constant linear velocity)
Track pitch:	1.6 $\mu\text{m}$
Diameter:	120 mm
Thickness:	1.2 mm
Center hole diameter:	15 mm
Recording area:	46 mm – 117 mm
Signal area:	50 mm – 116 mm
Material:	Any acceptable medium with a refraction index of 1.55
Minimum pit length:	0.833 $\mu\text{m}$ (1.2 m/sec) to 0.972 $\mu\text{m}$ (1.4 m/sec)
Maximum pit length:	3.05 $\mu\text{m}$ (1.2 m/sec) to 3.56 $\mu\text{m}$ (1.4 m/sec)
Pit depth:	~0.11 $\mu\text{m}$
Pit width:	~0.5 $\mu\text{m}$

### OPTICAL SYSTEM

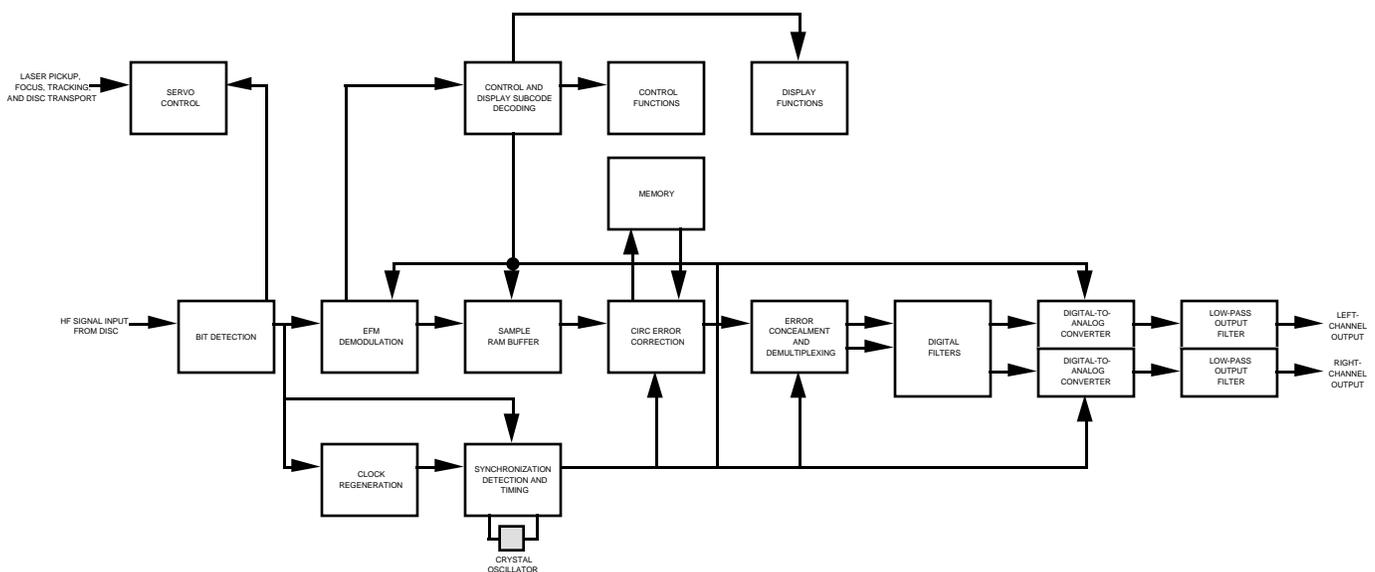
Standard wavelength:	$\lambda = 780 \text{ nm (7,800 \text{ \AA})}$
Focal depth:	$\pm 2 \text{ }\mu\text{m}$

### SIGNAL FORMAT

Number of channels:	2 channels (4 channel recording possible)
Quantization:	16-bit linear
Quantizing timing:	Concurrent for all channels
Sampling frequency:	44.1 kHz
Channel bit rate:	4.3218 Mb/sec
Data bit rate:	2.0338 Mb/sec
Data-to-channel bit ratio:	8:17
Error correction code:	Cross Interleave Reed-Solomon Code (with 25% redundancy)
Modulation system:	Eight-to-fourteen Modulation (EFM)

**Table 1.** *Red Book* specifications for the compact disc system<sup>3</sup>.

The compact disc player contains two main subsystems: the audio data processing system and the servo/control system. The servo, control, and display system orchestrate the mechanical operation of the player and include such items as the spindle motor, auto-tracking, lens focus, and the user interface. The audio data processing section covers all other player processes. A block diagram of the compact disc player is shown in Figure 1.



**Figure 1.** Block diagram of a compact disc player.

Since the introduction of the compact disc player in 1982, the market has seen three generations of players. First generation players were characterized by multi-bit DAC's used with brickwall reconstruction filters. Second generation players used the same multi-bit DAC's but took advantage of digital oversampling filters placed upstream of the DAC along with a *gentle* analog reconstruction filter. Finally, current players make use of low-bit DAC's along with oversampling filters and the gentle analog output filter. In the following sections, each of these DAC types and filtering methods will be investigated.

### ***Digital-to-analog Converters***

The very first demonstration players made by Sony, Philips, and others used 14-bit converters, which at the time were a vast improvement over analog equipment, but nonetheless were poor quality by today's standards. By the time the first consumer players were released in 1982, 16-bit converters were the standard. By 1989, many manufacturers touted the use of 18 and 20-bit converters.

**MULTI-BIT CONVERTERS**—At the digital hardware level, multi-bit converters may be designed in several ways. The most common of these include the ladder network converter, integrating converter, and dynamic element matching converter. The discussion of these implementations is beyond the scope of this text, so the ambitious reader is referred to the reference material.

The number of bits in a DAC is a poor method of determining its performance and accuracy. A better measure of performance is the accuracy of the actual bits themselves. Under ideal circumstances, a 16-bit converter would exactly convert all 16-bits of the sample data word in a linear fashion. However, this is seldom possible. In practice a 16-bit DAC is less than sufficient for accurate conversion.

The error in a 16-bit (or any multi-bit) converter relies on the accuracy of the most significant bit (MSB) of the data word. Inaccuracy in this bit can result in an error of half the signal's amplitude—a significant error by any measure. This in mind, manufacturers reasoned that converters with high bit rates could overcome this shortcoming along with others through sheer numbers. In addition to ensuring the accuracy of the MSB by having more than 16-bits, they can also improve quantization performance by adding  $2^{x-16}$  more quantization

levels than a 16-bit converter. Now, any nonlinearity in the conversion process would be a far smaller fraction of the overall signal and the more quantization levels result in a greater S/E ratio by virtue of Eq. 1. The extra bits used by these converters may be either thrown away, be left unused, or be put to other intelligent uses that will be discussed later. Unfortunately, it is a misconception that the use of an 18- or 20-bit DAC gives true 18 or 20-bit audio performance.

Despite the fantastic performance benefits of these  $n^{\text{th}}$  generation multi-bit converters, they are still plagued by many errors. Linearity was already mentioned, but they are also plagued by gain error, slew-rate distortion, and zero-crossing distortion. All of these error and distortion types introduce severe harmonic distortion and group delay; thereby perturbing signal stability, imaging, and staging.

Two methods of output reconstruction have been used with the multi-bit DAC's. The first of these employed the use of the "brickwall" filter. These filters had a very sharp cutoff characteristic and held the signal gain close to unity almost to cutoff. This was necessitated because the data was at a frequency such that aliasing and noise artifacts existed immediately above the audible band. The inherent problem with such a filter design was that they had tremendous phase nonlinearities at high frequencies, and high-frequency group delay—change in phase shift with respect to frequency. The second method of output reconstruction deals with an oversampling digital filter prior to the DAC and a gentle analog filter. By gentle, it is meant that a cutoff slope of 12 dB/octave and a -3 dB point of 30-40 kHz can be used. Its design then is noncritical and low-order—which guarantees excellent phase linearity. In fact, for most practical reconstruction filters, phase distortion can be held at  $\pm 0.5^\circ$  over the entire audio band. The discussion of this is pertinent to both multi- and low-bit DAC's, so the topic will be covered after the next section.

**LOW-BIT CONVERTERS**—To combat the problems of multi-bit converters, two competing technologies were developed, the first by Matsushita and the second by Philips. Rather than converting whole data words in parallel at the sampling frequency, both methods involve converting far shorter word lengths at far higher rates. This serial data conversion is an inherently digital process and has been made possible in part by the powerful digital signal processors available today.

Matsushita's method is based on pulse-width modulation (PWM). In this design, the width of the signal pulse represents the unique data word, thus it is critical that the PWM steps have exact width and minimum jitter to maximize accuracy and linearity of the output. The commercial name for the process used is MASH (Multi-stAge noise SHaping). A MASH converter is made of a 4-times oversampling digital filter, followed by first- and second-order noise shapers in parallel. The output from the noise shapers is then fed into a PWM converter, whose output is then low-pass filtered. A block diagram of the MASH system is shown in Fig. 2.

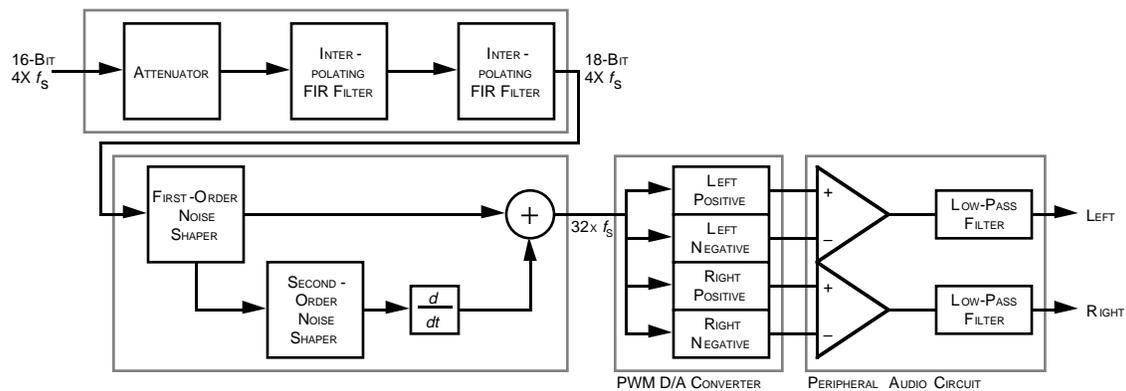


Figure 2. Block diagram of a PWM/MASH digital-to-analog converter.

A digital finite impulse response (FIR) filter produces 18-bit data from a 16-bit input sample after 4-times oversampling. The noise shapers then convert this 18-bit data into an 11-step quantized format for the PWM after 8-times oversampling. The PWM system is operated at 768 times the original sampling frequency (33.868 MHz). If it were to actually do a 1-bit conversion of 16-bit signals, 65,536 pulses would be needed to represent each amplitude. However, this would require the converter to operate at speeds in excess of 2.98 GHz—faster than the currently available bipolar transistor technology. This restraint imposes the requirement that the 18-bit data be reduced to the 11-step output. In practice the MASH converter can only be considered a “3.5-bit” converter.

The second low-bit conversion technique by Philips is known as pulse-density modulation (PDM) or Bitstream conversion. In this technique, the density ratio of the sign of the pulses is related to the original 16-bit data word. The PDM converter is a true 1-bit technology. This signal representation may not seem immediately obvious. A simple model helps illustrate what is happening<sup>4</sup>. If a

light is on, then the room is brightly lit; if the light switch is off, the room is dark. But if the switch is cycled rapidly on and off, an intermediate intensity can be created. The sample data from the decoder chip is first passed to a low-pass non-recursive 4-times oversampling FIR interpolation filter. This type of filter yields higher quality because it is phase-linear. First-order noise shaping is performed by the accumulator of the multiplier in the filter. The second filtering stage consists of a 32-times oversampling linear interpolator and a 2-times oversampling sample and hold circuit. At this stage, a 352 kHz digital dither signal at -20 dB is added to the sample signal. This reduces nonlinearities induced by quantization noise. At this point, the total oversampling is 256-times and the data word has increased to 17-bits. The data is then fed at a frequency of 11.2896 MHz into the second order noise shaper. The noise shaper reduces the 17-bit data to a 1-bit stream by using  $\Sigma$ - $\Delta$  modulation. In this process quantization noise is redistributed away from the audio frequency by as much as 2 orders of magnitude. The bitstream is then converted to an analog form by a switched capacitor network. A block diagram of the PDM converter is shown in Fig. 3.

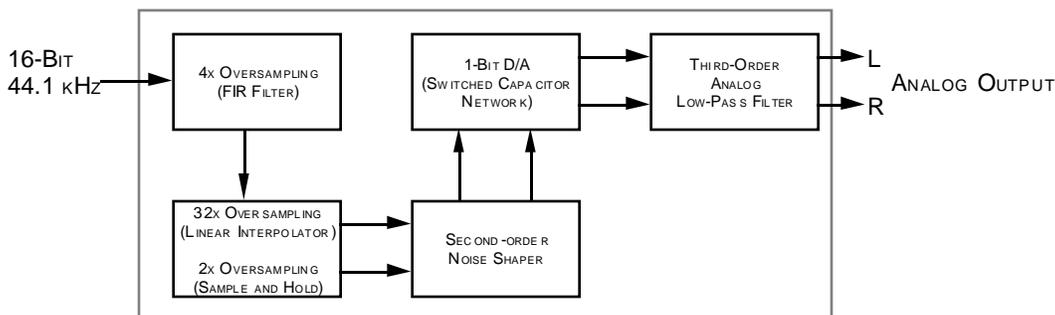


Figure 3. Block diagram of a PDM digital-to-analog converter.

Because there are only two voltage references in the PDM converter, there is no level matching requirement for improved accuracy. Therefore the linearity errors associated with it are eliminated.

Comparisons of THD and linearity error for various 16-, 18-, 20-, and 1-bit converters yield interesting results. PWM and PDM converters show  $< \pm 1$  dB linearity for input signals from -100 to -80 dB and are virtually linear thereafter. Some of the most expensive players on the market with 18- and 20-bit converters using 4-, 8-, 16-, and even 32-times oversampling yield up to  $\pm 4$  dB linearity error for signals as high as -75 dB. In the THD tests performed with a -60 dB 1 kHz sine wave test signal, the expensive multi-bit players showed harmonics up to the 13th

at levels greater than  $-110$  dB<sup>5</sup>. Only the PDM converter was able to hold all non-fundamental harmonics under  $-110$  dB.

### ***Digital Filtering, Oversampling, and Noise Shaping***

Oversampling is not mandated by any theorem discussed previously, but its use yields tremendous performance gains regardless of the type of converter used. Oversampling quite simply means using a sampling frequency greater than that dictated by the Nyquist theorem. By exceeding the Nyquist frequency, many of the precision demands made by the theorem can be relaxed (like the brickwall filter). In addition to the benefits seen at the output filter, the signal-to-noise ratio is boosted greatly and quantization noise is reduced in the audio band. The latter is decreased by an incredible amount when oversampling is used in conjunction with noise-shapers, which will be discussed shortly.

The oversampling process is well suited to a digital signal processor (DSP), which essentially takes in audio samples, performs an operation on them, and then outputs audio samples. Because the samples are modified, the DSP is in effect a digital filter. The DSP is beneficial because the operations it performs are precise and repeatable, not otherwise possible with analog techniques, and result in lower noise and distortion than with analog techniques. The oversampling process can be viewed simply as interleaving zeros between each sample with additional samples. In practice, these new samples are produced by using a shift register (which acts as a delay line), coefficient multipliers, and an adder. The shift register has taps after each delay element. The output of each tap is taken and then multiplied by a coefficient stored in ROM associated with the impulse response of the low-pass filter. These delayed multiples are then summed to generate a new sample. An example of this can be seen in Fig. 4.

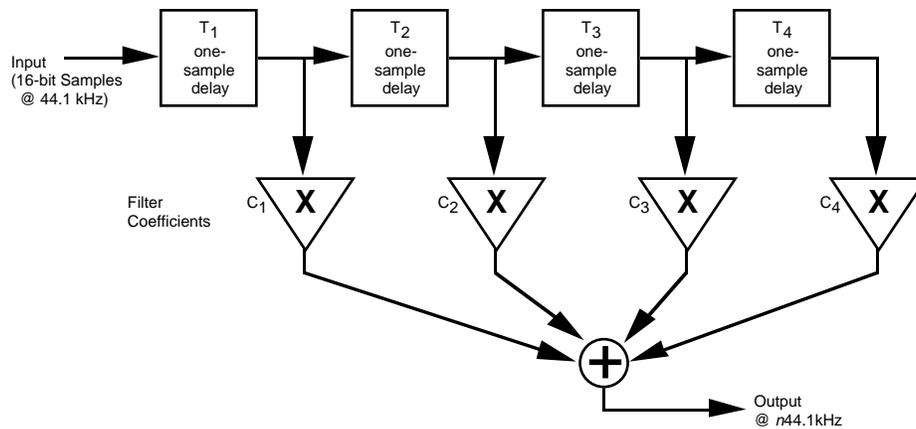


Figure 4. Use of a transversal filter to achieve oversampling .

The total result of this process is that new interpolated samples are created at each interleaved zero-value. This is shown graphically in Fig. 5.

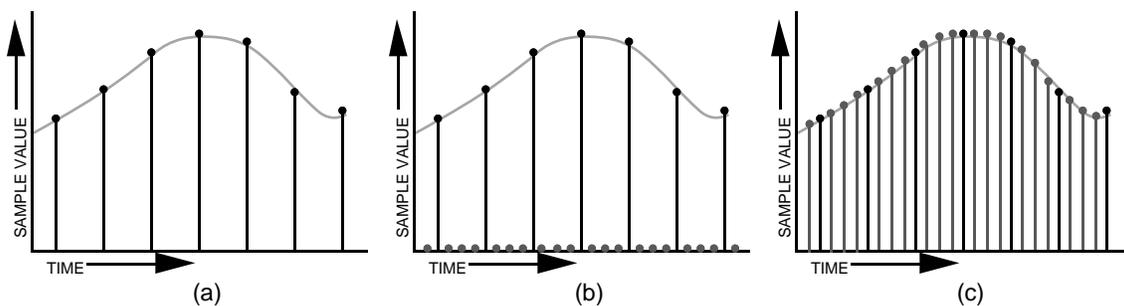


Figure 5. Effect of zeros interleaving and oversampling on a signal. Original signal and samples (a) with: interleaved zeros (b) and interpolated new samples (c).

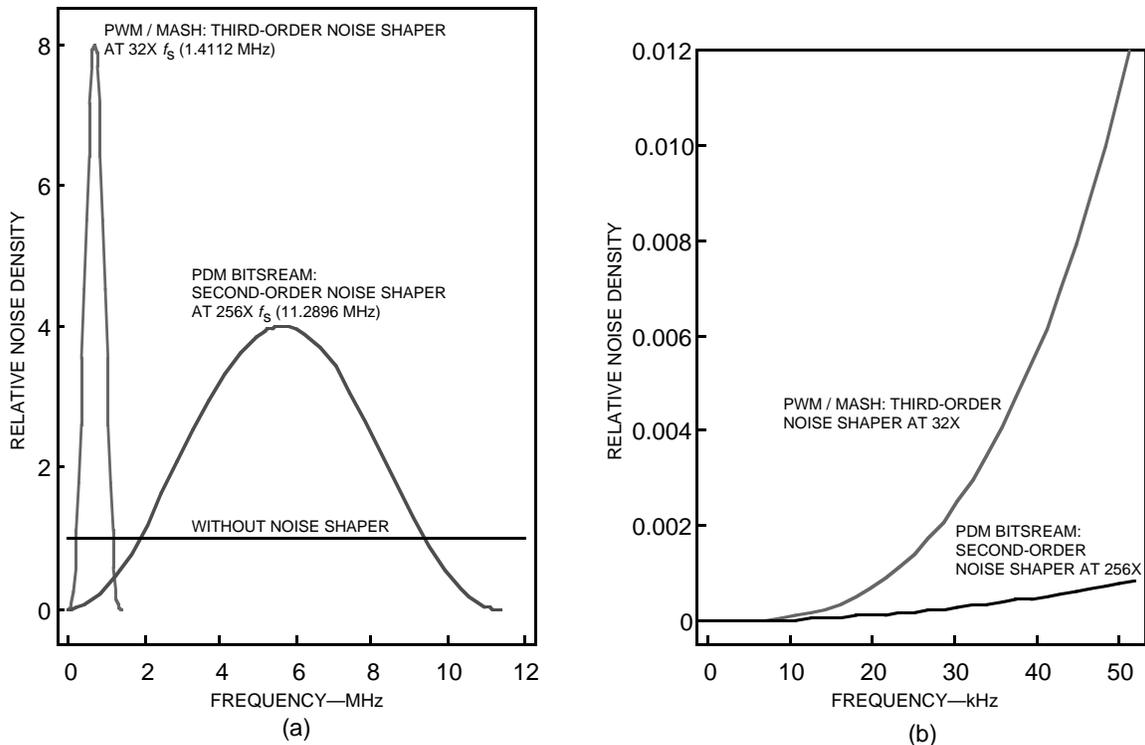
As a result of this, the sampling frequency has increased by whatever amount of oversampling occurred, and the data word length has grown. Because the sampling frequency has risen, the noise in the audio band has been shifted out by a greater amount than it was before. Noise shaping is then implemented to reduce the data word size and further exaggerate the amount noise is moved out of the audio band.

As stated previously, the primary job of the noise shaper is to alter the frequency spectrum of the error signals so as to move most of the quantization error out of the audible frequency range. Noise shaping reduces quantization noise by using a negative feedback technique. In effect, the shaper attempts to reduce quantization error by using its known qualities to actually subtract from the signal. The power behind a low-bit conversion technique relies on the power of its noise-shaping

algorithm. In general, the more complex the noise-shaper, the lower the audio band noise. Thus the performance of the noise shaper is determined by the order of the shaper and its operating frequency. The latter parameter is a function of how much oversampling is performed prior to shaping. The first relationship we can extract from these parameters is the higher the order of the shaper, the higher the slope of the noise redistribution and hence the lower the audible noise. The drawback is that sideband noise is increased so much that the analog filters could be overburdened. The second relationship is that the higher the operating frequency, the higher in the frequency domain the noise is shifted. These two relationships are defined by the noise-density distribution equation which is shown in Eq. 4, where  $f_s$  is the original sampling frequency and  $n$  is the shaper order..

$$\rho_{noise}(f) = \left[ 2 \sin\left(\frac{f}{f_s}\right) \right]^n \quad (4)$$

The relationship is also illustrated in Fig. 6a. The only limitation in operation speed is the available speed of digital logic. Therefore, the conscientious designer aims for the proper balance between shaping order and oversampling.



**Figure 6.** Various noise-density distributions as a function of frequency (a). The PDM and PWM distributions in the audio band (b).

As a footnote, the operating frequency has the greatest effect of the two parameters on noise density distribution. This is clearly visible in a much more detailed look at the noise distributions in Fig. 6b. Clearly, the PDM has significantly lower audio band noise and necessitates only a simple analog reconstruction filter. A block diagram of the third-order noise shaper used in the MASH converter is shown in Fig. 7.

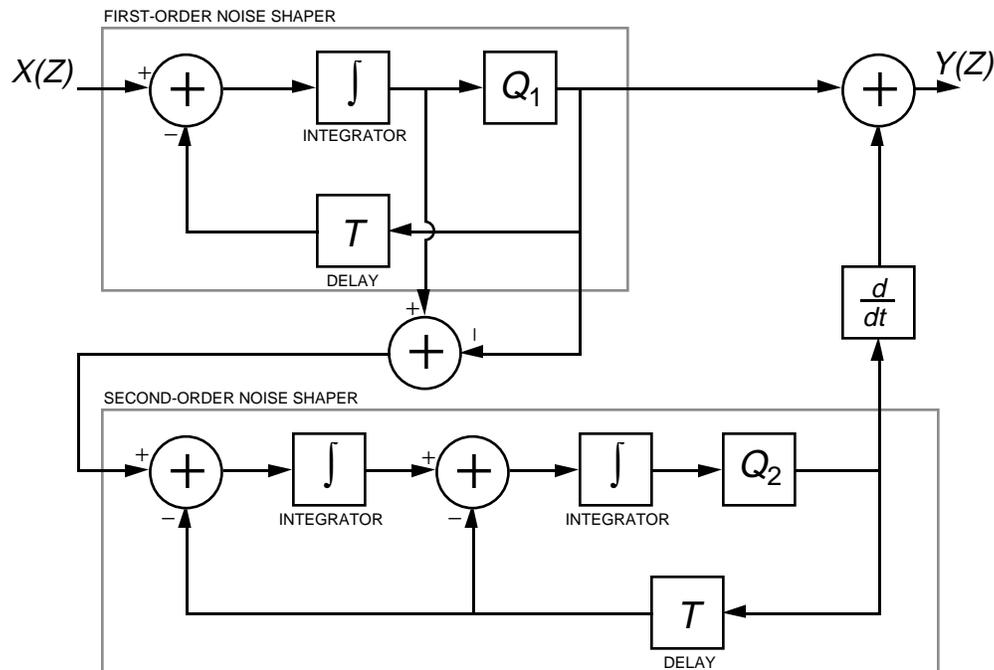


Figure 7. Third-order noise shaper used in the PWM converter.

In the shaper given in Fig. 6, the input signal is fed into quantizer  $Q_1$  after the residual error is subtracted from the delay block in the first order shaper. The residual signal is also fed into the second order noise shaper, where the output of the second quantizer,  $Q_2$ , is differentiated and then summed with the output of the first noise shaper to create the final output signal<sup>6</sup>.

The compact disc has only existed for about 13 years, and more than likely has as many years of useful life left. There are many advances that are still possible in the format and many of them are just in their infancy. However, many challengers have already entered the playing field; some by the original creators of the compact disc. Sony has created both the DAT standard as well as the Mini-Disc, and Philips has created the DCC (digital compact cassette). Regardless of

the compact disc's lifetime, it is certain that digital audio will remain, and analog will be reserved to the role of input at the microphone in the studio and output at the speaker in the listening environment.

This is by no means a complete or exhaustive analysis of the basic fundamentals of the compact disc player. Many issues such as error-correction, data encoding and decoding, and pickup design were neglected. However, the concepts covered here should provide the reader with a strong background, and incite some interest in learning more. For the reader who *is* interested in learning more, the *The Art of Digital Audio* by Watkinson is an extensive collection of knowledge on digital audio. It is at times very technical in nature, but the material introduced builds upon itself nicely. Pohlmann's book, *The Compact Disc Handbook*, focuses solely on the compact disc player and the compact disc itself along with all its diverse formats—of which audio is only one. His book is very thorough in its coverage and should leave no questions from the reader unanswered. Pohlmann's book has a fair amount of overlap with Watkinson's and would make a better starting point for those short on time.

#### ENDNOTES

<sup>1</sup>John Eargle. ("Bitter Jitter and Sweeter Dither," *Audio*. Vol. 76, No. 1). p. 24

<sup>2</sup>Ken C. Pohlmann. (*The Compact Disc Handbook—The Computer music and digital audio series*. 2nd Ed. Madison, WI: A-R Editions, Inc. 1992). p. 34-35.

<sup>3</sup>Pohlmann, p. 46.

<sup>4</sup>Pohlmann, p. 148.

<sup>5</sup>Prasanna Shah. ("Music of the Bitstream," *Audio*. Vol. 75, No. 1). p. 64.

<sup>6</sup>Shah, p. 61.

**REFERENCES**

Eargle, John. "Bitter Jitter and Sweeter Dither," *Audio*. Vol. 76, No. 1 (Jan. 1992). 24+

Oppenheim, Alan V. *Signals and Systems*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1983.

Pohlmann, Ken C. *The Compact Disc Handbook—The Computer music and digital audio series*. 2nd Ed. Madison, WI: A-R Editions, Inc. 1992.

Shah, Prasanna. "Music of the Bitstream," *Audio*. Vol. 75, No. 1 (Jan. 1991). 56 - 60+.

Strawn, John. *Digital Audio Engineering—The Computer music and digital audio series*. Los Altos, CA: William Kaufmann, Inc., 1985.

Strawn, John. *Digital Audio Signal Processing—The Computer music and digital audio series*. Los Altos, CA: William Kaufmann, Inc., 1985.

Watkinson, John R. *The Art of Digital Audio*. 2nd Ed.. Boston, MA: Focal Press, 1994.

Watkinson, John R. *Coding for Digital Recording*. Boston, MA: Focal Press, 1990.

Whyte, Bert. "Every Little Bit Helps," *Audio*. Vol. 76, No. 8 (Aug. 1992). 19-21.