

# High-Fidelity PWM Inverter for Digital Audio Amplification: Spectral Analysis, Real-Time DSP Implementation, and Results

César Pascual, Zukui Song, Philip T. Krein, *Fellow, IEEE*, Dilip V. Sarwate, *Fellow, IEEE*, Pallab Midya, *Senior Member, IEEE*, and William (Bill) J. Roeckner

**Abstract**—A complete digital audio amplifier has been developed, implemented and tested. The process is entirely computational, and the output load and filter are the only analog components in the system. The process makes use of digital signal processing and a switching power stage to provide both high fidelity and high efficiency, beginning with a digital audio data stream. The advantages of naturally-sampled pulse-width modulation (PWM) are discussed in depth, including spectral analysis and comparisons to uniformly-sampled PWM. It is shown that natural PWM does not introduce audible distortion at switching frequencies consistent with power electronics practice. Interpolation methods for sample data conversion to natural PWM are discussed, and error analysis is presented based on Lagrange's Expansion Theorem. Noise-shaping processes are used to support high fidelity with practical values of time resolution. A counter conversion process enforces switching dead time in the inverter gate signals. The experimental full-bridge inverter implementation demonstrates that miniaturization is possible. A complete test system delivered more than 50 W into an 8  $\Omega$  load with an efficiency of 80% and total harmonic distortion plus noise of 0.02%.

**Index Terms**—Class D amplifier, digital audio amplifier, noise-shaping, pulse-width modulation, uniform and natural sampling.

## NOMENCLATURE

$c(t)$	Differences between actual crossing times and noise-shaper output.
$d$	Duty ratio.
$f$	Generalized frequency.
$f_c$	Carrier waveform (switching) frequency.
$f_L$	Low-pass filter cutoff frequency.
$f_s$	Signal bandlimit frequency.
$j$	$\sqrt{-1}$ .
$k$	Integer cycle number; frequency index.
$m$	Smallest positive integer that gives $\lfloor (mf_s + f_L)/f_c \rfloor = 1$ .

$n$	Frequency index; integer.
$p(t)$	Signal-dependent portion of UPWM waveform.
$\hat{p}(t)$	Signal-dependent portion of NPWM waveform.
$pwm(t)$	Complete UPWM waveform.
$\widehat{pwm}(t)$	Complete NPWM waveform.
$sq(t)$	Square wave with values of $\pm 1$ and 50% duty ratio.
$t$	Time.
$\hat{t}$	Solution to generalized crossing-point equation.
$t_k$	Crossing-point for natural-sampled PWM.
$u(k)$	Quantizer input for noise-shaping.
$u(t)$	Heaviside step function.
$v[k]$	Noise-shaping output samples.
$x(t)$	Modulating signal, scaled to make $ x(t)  < 1$ .
$\hat{x}[k]$	Discrete-time samples from PCM-to-NPWM conversion (used as the noise-shaping input).
$\hat{x}(t)$	Distorted signal that, under uniform sampling, yields the same samples as $x(t)$ under natural sampling.
$y(t)$	Signal-dependent baseband portion of PWM waveform.
$y_k(t)$	Signal-dependent switching-harmonic portion of PWM waveform.
$z$	$z$ -domain operator.
$z(t)$	Demodulation output.
$H(z)$	Transfer function for noise-shaping.
$N(n)$	A function representing a bandlimit frequency ratio $N(n) = \lfloor (nf_s + f_L)/f_c \rfloor$ .
$P(f)$	Fourier transform of $p(t)$ .
$\hat{P}(f)$	Fourier transform of $\hat{p}(t)$ .
$PWM(f)$	Fourier transform of UPWM signal.
$\widehat{PWM}(f)$	Fourier transform of NPWM signal.
$S_n(f)$	Fourier transform of $x^n(t)$ .
$SQ(f)$	Fourier transform of square wave.
$T$	Switching period $1/f_c$ .
$X(f)$	Fourier transform of $x(t)$ .
$\hat{X}(f)$	Fourier transform of $\hat{x}(t)$ .
$Y(f)$	Fourier transform of $y(t)$ .
$\hat{Y}(f)$	Fourier transform of $\hat{y}(t)$ .
$Y_k(f)$	Fourier transform of $y_k(t)$ .
$Z(f)$	Fourier transform of $z(t)$ .
$\tau_k$	Pulse turn-off time.
$\omega$	Radian frequency of signal.

Manuscript received July 1, 2002; revised September 26, 2002. This work was supported through the joint University of Illinois/Motorola Center for Communications Research. Recommended by Associate Editor S. B. Leeb.

C. Pascual, Z. Song, P. T. Krein, and D. V. Sarwate are with the Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL 61801-2018 USA (e-mail: krein@ece.uiuc.edu).

P. Midya and W. J. Roeckner are with Motorola Labs, Schaumburg, IL, 60196 USA.

Digital Object Identifier 10.1109/TPEL.2002.807102

## I. INTRODUCTION

### A. Basis of Switching Amplifiers

**P**ULSE-WIDTH modulation (PWM) is well established in power electronics as a basis for inverters with sinusoidal output voltages. The application of PWM inverter circuits for switching audio amplification is also well known. It is useful to consider the fundamental basis for this approach. From a power delivery standpoint, PWM provides two crucial advantages. The first advantage is that it encodes a signal into a few discrete levels, with the information represented in pulse duty ratios. This coding characteristic permits energy to be delivered by switching among a small number of discrete power sources. The second advantage is the ability to recover the signal from its discrete-level form with a passive filter.

When the discrete power sources can be generated efficiently, PWM provides the basis for highly efficient signal delivery, especially to loads with low-pass characteristics. Thus a PWM signal prepared from an audio input can be used as a switching function for a bridge or half-bridge inverter, and a low-pass filter extracts the audio and delivers it to a loudspeaker. Overall, the process is nonlinear, and it is not obvious whether it can achieve the levels required of a high-fidelity audio system.

### B. Questions of Fidelity and Distortion

It is known that low-pass filter signal recovery is imperfect for PWM [1], [2]. In addition, it is natural to expect that the nonlinearities associated with real inverters and switching devices would tend to limit performance. With the extensive growth of digital audio, the digital characteristics of a signal provide a basis for defining high fidelity. For example, the 16-b signal from a CD-Audio (CDA) source has 1-b quantization error as the lower bound on noise and distortion. This is one part in  $2^{16}$ , or  $-96$  dB. A 24-b audio sampling range corresponds to a lower bound of  $-144$  dB. An amplifier that can reach these low levels is effectively perfect by comparison with audio signal quantization error. Commercial audio amplifiers do not achieve these levels in general. A high-end consumer audio amplifier might achieve low-power distortion levels of about 0.05% ( $-66$  dB). Expensive studio or THX-grade amplifiers can reach 0.02% ( $-74$  dB). Measured results from an available amplifier with digital input (Sony model STR-DE445) showed 0.2% distortion ( $-54$  dB) at a 60 W output level. It is important to recognize, however, that for music and similar audio signals, listeners are much more tolerant of harmonic distortion than of nonharmonic noise. An amplifier with 0.2% harmonic distortion, for example, produces good results if its signal-to-noise ratio is high.

In general, a concept of equivalent word length can be used to describe fidelity. Table I shows various bit levels, in comparison with distortion percentages and with distortion dB levels. In audio work, the term total harmonic distortion (THD) is sometimes used loosely: the distortion produced by a nonlinear process need not be harmonic. More generally, the measure at the output is the sum of distortion and noise, even though harmonic distortion has less effect on sound quality than noise. Table I reflects this definition. The equivalent word length is the power of 2 that is close to the inverse of

TABLE I  
EQUIVALENT WORD LENGTHS BASED ON DISTORTION LEVELS

Distortion+noise (% of signal)	Distortion+noise (dB relative to signal)	Equivalent word length (bits)
0.2	$-54$	9
0.1	$-60$	10
0.05	$-66$	11
0.02	$-74$	12
0.01	$-80$	13
0.0015	$-96$	16
0.0001	$-120$	20
0.000024	$-132$	22
0.000006	$-144$	24

the distortion plus noise fraction. The table only reflects one measure of performance. Dynamic range, intermodulation distortion, channel separation, signal-to-noise ratio and other factors in general favor a high word length in the data stream even if the amplifier linearity (especially in terms of harmonic distortion) is much lower.

Can a PWM process deliver an audio output with total harmonic distortion (THD) and nonharmonic distortion effects that can be considered high fidelity (approximately  $-60$  dB or better)? What are the important distortion effects in a real inverter, and how can they be addressed?

In this paper, it is shown how PWM can indeed be used to provide high fidelity. The distortion effects are analyzed in depth. Noise-shaping processes that reduce quantization errors in the process are described. The process is then implemented in a complete bridge inverter that processes information directly in digital form. The signal processing chain from digital input to the inverter gate drives is entirely digital, and the loudspeaker and its filters represent the only analog components in the system. Experimental results confirm that a PWM inverter can achieve high fidelity in practice. Limitations of dead-time and switching device imperfections can be mitigated to avoid direct effects on system distortion. The experimental circuit delivered 50 W into the load with total distortion plus noise of about 0.022% ( $-73$  dB) and overall power efficiency of more than 80%.

### C. PWM Sampling

In a typical class-D amplifier in an audio application, an analog signal to be amplified is compared against a high-frequency triangle or sawtooth. The resulting switching function drives a bridge or half-bridge inverter. When digital audio signals are available, the conventional approach is to convert first to analog, and then proceed with the comparison process.

Analog PWM is termed natural sampling, since the output time instants represent sample points at which the signal naturally crosses the comparison waveform. It is also possible to apply uniform sampling, in which pulse widths are determined from signal values sampled at a fixed rate. Analysis of natural sampling PWM (NPWM) with a few discrete tones as modulating input [3] has shown that the output spectrum includes the baseband audio signal, the switching frequency and its harmonics, and side bands formed by the analog signal mixing with

all harmonics of the switching frequency. Since the switching side bands are unlimited in width, switching distortion extends down into the low frequency baseband range. Thus it is widely accepted that natural sampling generates distortion [4], [5]. Fortunately, this distortion can be managed with the choice of a high switching frequency.

Uniform sampling PWM (UPWM) also gives rise to the baseband signal, the switching frequency and its harmonics, and switching sidebands. In addition, a UPWM waveform contains harmonics of the signal itself. These are considerably larger than the switching sidebands for practical values of switching frequency. Hence UPWM has more distortion in the audio band than NPWM in practice.

The primary objective of class-D amplifiers is high efficiency. Conventional audio amplifiers rarely exceed 20% efficiency in use. An amplifier based on a PWM inverter, in contrast, can reach 90% efficiency or more. In the past, it has often been assumed that this efficiency improvement benefit comes at the price of distortion. The potential efficiency improvement for battery-powered applications or for miniature amplifiers has driven the study of advanced PWM amplification techniques. It is now known that such a tradeoff is unnecessary. High quality can be achieved with PWM, and distortion can be reduced to acceptable levels.

#### D. Toward Digital PWM

The connection between natural sampling and distortion is not necessarily relevant for PWM amplification. In power electronics PWM practice, switching frequencies must be set far above the modulating waveform to simplify the filtering process and minimize the size of energy storage components. For a single tone modulating signal, the magnitude of various distortion components is proportional to ordinary Bessel function values [6], [7]. When the switching frequency is several times the modulating frequency, the baseband distortion is very low. For example, a switching frequency of 88 kHz and modulating frequency of 12 kHz produces distortion in the baseband (12 kHz and below) of a low  $-111$  dB. A switching frequency more than ten times the modulation frequency produces baseband distortion below  $-144$  dB. Considering the audio band to extend to about 20 kHz, a switching frequency of 200 kHz does not produce detectable distortion—based on a single-tone analysis.

An analog PWM process is susceptible to noise. Noise that appears in the modulating signal or the comparison waveform is troublesome, especially since noise causes dc bias effects in an inverter [8]. Another potential source of distortion is the comparison waveform: If the triangle waveform is not perfectly linear, the natural sampling times will contain error, and the end result will be distorted. With the growing use of digital audio in compact disks (CDs), DVDs, movie soundtracks, broadcasting, and computer applications, it is natural to pursue an entirely digital PWM process. Several groups are exploring digital implementations of PWM for audio amplifier applications [9]–[14].

High-quality digital PWM processes have been considered for more than a decade [15], although much of the work has not included the switching power stage. Only a few results at

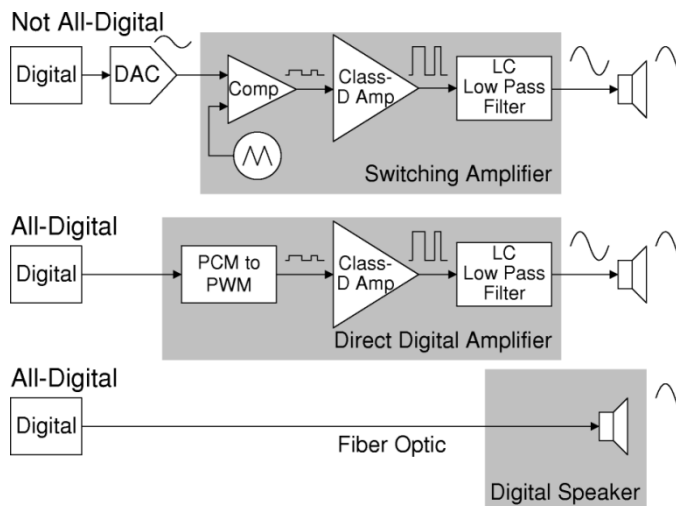


Fig. 1. Progression from analog PWM amplifier to digital speaker.

the power stage output have been reported [11], [14], with significant power levels only in [11]. In any case, it is well established that NPWM can provide the basis for high-fidelity audio processing. Power considerations, such as switch nonlinearities, dead-time requirements, and gate drive delays, are complicating issues that imply additional sources of distortion or other problems. Here the design is carried through to a complete experimental system, with performance that goes beyond the prior results.

Fig. 1 shows a progression from an analog PWM process through a fully digital process to the ultimate goal of a “digital speaker.” In a so-called digital speaker, an efficient local class D amplifier, combined with a digital processing stream, produces local output in response to both the digital audio input and to network commands such as volume and channel settings. We envision a flexible network-based audio system as the techniques mature.

#### E. Summary of Contributions

The contributions of this paper are to present the complete process of a high-fidelity inverter for audio applications in the context of power electronics, and to extend the measured performance in prior work to an unambiguous high-fidelity level at significant output power levels. As part of the work, the spectral behavior of UPWM and NPWM is presented for arbitrary modulating signals. A complete digital inverter that begins with a digital audio signal and ends with an inductive load driven by a MOSFET bridge is developed. The signal processing steps begin with an upsampling block to convert relatively slow audio samples to the high switching frequency needed for a practical inverter. The fast samples are then converted with an interpolation approach into a NPWM data stream. The NPWM samples are processed through a noise-shaping filter to move much of the quantization noise out of the audio band, while at the same time allowing a moderate pulse-width resolution to support the full signal fidelity. The noise shaper output is then interpreted as pulse widths and used to generate gate drive signals for the inverter. Gate drive generation addresses dead time. A fourth-order passive low-pass filter augments a loudspeaker

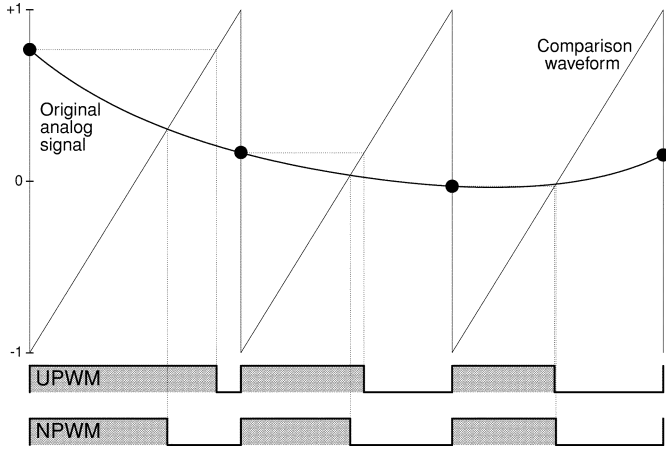


Fig. 2. Uniform versus natural PWM (trailing edge).

load to avoid high-frequency losses in the speaker coil. Experimental results indicate that the approach reaches the distortion and noise performance levels of some of the best available commercial audio amplifiers.

## II. NPWM AND UPWM SPECTRAL CHARACTERISTICS FOR ARBITRARY SIGNALS

Fig. 2 shows the difference between UPWM and NPWM. This is *trailing edge* modulation since the waveform results in switch turn-on at the beginning of each cycle, with modulation applying to the turn-off time later in the cycle. The sawtooth shown represents a hypothetical comparison waveform, needed for the NPWM case. If a sawtooth is used as a comparison waveform, the resulting modulation is called *single-edge* NPWM. If a triangular waveform is used instead, the result is called *double-edge* NPWM since both edges will carry information. UPWM is easy to generate. The sample is available, and a direct conversion from amplitude to time is the only step. For NPWM, an intersection point must be located.

### A. The General Spectrum

Consider a PWM process with arbitrary input signal  $x(t)$ , limited in range to  $|x(t)| < 1$ . The analysis given here is summarized from [16]. While any PWM approach can be modeled, we will consider trailing-edge, single-sided modulation (TEPWM) in this paper. The results are extended to leading-edge and double-edge cases in [16]. If  $x(t)$  were identically zero, the pulse-width modulator would produce  $sq(t)$ , a 50% duty-cycle square wave that takes on value +1 for  $kT \leq t < (k + (1/2))T$  and value -1 for  $(k + (1/2))T \leq t < (k + 1)T$ . The PWM signal can be represented as a sum of pulses over all time, in terms of the Heaviside step function  $u(t)$ , as

$$pwm(t) = sq(t) + 2 \sum_{k=-\infty}^{\infty} u(t - (k + \frac{1}{2})T) - u(t - kT - \tau_k) \quad (1)$$

where the  $k$ th pulse width  $\tau_k$  depends on  $x(t)$ . The sum in (1) is the *signal-dependent* part  $p(t)$ , and consists of a sequence of positive or negative pulses of duration  $|\tau_k - T/2|$ . If  $\tau_k > T/2$ , the  $k$ th pulse is positive, and begins at  $t = (k + (1/2))T$  while

if  $\tau_k < T/2$ , the  $k$ th pulse is negative, and ends at  $t = (k + (1/2))T$ . Regardless of the value of  $\tau_k$ , the fixed pulse edges in  $p(t)$  always occur at  $t = (k + (1/2))T$ .

The signal  $pwm(t) = sq(t) + p(t)$  is a finite power signal and thus its Fourier transform  $PWM(f)$  exists in the generalized sense. Let  $SQ(f)$  and  $P(f)$  denote the Fourier transforms of  $sq(t)$  and  $p(t)$  respectively. Since the square-wave Fourier series and transform are well known, knowledge of the signal-dependent spectrum  $P(f)$  will lead to  $PWM(f) = SQ(f) + P(f)$ . The pulse  $2[u(t - (k + (1/2))T) - u(t - kT - \tau_k)]$  has Fourier transform

$$(j\pi f)^{-1} [\exp(-j2\pi f(k + \frac{1}{2})T) - \exp(-j2\pi f(kT + \tau_k))]$$

and hence the spectrum of the modulated pulse train is given by

$$P(f) = \frac{1}{j\pi f} \sum_{k=-\infty}^{\infty} \exp\left(-j2\pi f\left(k + \frac{1}{2}\right)T\right) - \exp(-j2\pi f(kT + \tau_k)). \quad (2)$$

The values  $\tau_k$  can be considered for the UPWM and NPWM cases to determine the overall spectral behavior.

### B. Spectra of UPWM Signals

In a uniform-sampling TEPWM (UPWM) scheme, the  $k$ th pulse width  $\tau_k$  equals  $T[1 + x(kT)]/2$ . Let us use  $p(t)$  to denote the signal-dependent part of the UPWM signal, and  $P(f)$  to denote its Fourier transform. From (2) we get

$$\begin{aligned} P(f) &= \frac{\exp(-j\pi fT)}{j\pi f} \sum_{k=-\infty}^{\infty} \exp(-j2\pi f kT) \\ &\quad \cdot [1 - \exp(-j\pi fT x(kT))] \\ &= \exp(-j\pi fT) \sum_{k=-\infty}^{\infty} \sum_{n=1}^{\infty} \frac{(-j\pi fT)^{n-1}}{n!} S_n(f - kf_c) \end{aligned} \quad (3)$$

where  $S_n(f)$  is the Fourier transform of  $x^n(t)$  and we have used the discrete-time Fourier transform (DTFT) formula [17] to obtain (4). The overall spectrum  $PWM(f)$  of a uniform-sampling UPWM signal is thus  $SQ(f) + P(f)$  where  $P(f)$  is given in (4). The spectrum of a square wave consists of the carrier signal of frequency  $f_c = 1/T$  and all its odd-order harmonics with amplitudes decreasing inversely with order. However, as shown below,  $PWM(f) = SQ(f) + P(f)$  contains *all* the carrier harmonics with amplitudes decreasing inversely with order.

We consider  $P(f)$  and  $p(t)$  in more detail. Let

$$P(f) = Y(f) + \sum_{k=1}^{\infty} Y_k(f)$$

where

$$\begin{aligned} Y(f) &= \exp(-j\pi fT) \sum_{n=1}^{\infty} \frac{(-j\pi fT)^{n-1}}{n!} S_n(f) \\ &= \exp(-j\pi fT) \left[ X(f) + \sum_{n=2}^{\infty} \frac{(-j\pi fT)^{n-1}}{n!} S_n(f) \right], \end{aligned} \quad (5)$$

and

$$Y_k(f) = \exp(-j\pi fT) \sum_{n=1}^{\infty} \frac{(-j\pi fT)^{n-1}}{n!} \cdot [S_n(f - kf_c) + S_n(f + kf_c)]. \quad (6)$$

Thus  $p(t)$  includes a signal  $y(t)$  with spectrum  $Y(f)$  given by (5). It is now apparent that

$$y(t) = x(t - T/2) + \sum_{n=2}^{\infty} \frac{1}{n!} \left(\frac{-T}{2}\right)^{n-1} \frac{d^{n-1}}{dt^{n-1}} x^n(t - T/2). \quad (7)$$

This shows that the UPWM signal contains the modulating signal  $x(t)$  delayed by  $T/2$  as well as distortion arising from the derivatives of the powers of  $x(t - T/2)$ . The distortion terms are reduced in amplitude as  $n$  increases, and also when period  $T$  decreases. In addition to derivatives of the powers of  $x(t - T/2)$ , the UPWM signal also includes signal terms  $y_k(t)$ ,  $k > 0$ , whose spectra  $Y_k(f)$  are given by (6). Note that

$$y_k(t) = 2(-1)^k \sum_{n=1}^{\infty} \frac{1}{n!} \left(\frac{-T}{2}\right)^{n-1} \frac{d^{n-1}}{dt^{n-1}} \cdot [x^n(t - T/2) \cos(2\pi k f_c t)] \quad (8)$$

is a signal at frequency  $kf_c$ .

The above results are general and require only that  $|x(t)| \leq 1$ . More realistically,  $x(t)$  is strictly band-limited to the *signal bandwidth*  $f_s$ . The carrier frequency will be chosen such that  $f_c \geq 2f_s$ . This condition is required in order to recover  $x(t)$  from the PWM signal without aliasing effects. Notice that requiring that  $f_c \geq 2f_s$  does not guarantee perfect reconstruction—some residual nonlinear distortion is always present. The distortion performance is better when the value of the *switching ratio*  $f_c/f_s$  is large. In [16], it is shown that if the magnitude of the derivative of  $x(t)$  is smaller than  $2f_c$  [which holds whenever  $x(t)$  is band-limited to  $f_s$  where  $f_s < f_c/\pi$ ], then the signal  $y_k(t)$  in (8) can be expressed as

$$y_k(t) = \frac{2(-1)^k}{k\pi} [\sin(2\pi k f_c t) - \sin(2\pi k f_c t - k\pi y(t))] \quad (9)$$

which can be recognized as the difference of the  $k$ th carrier harmonic and the *phase modulation* of the  $k$ th carrier harmonic by  $y(t)$ . From (9), the UPWM signal  $pwm(t)$  obtained by uniform sampling of  $x(t)$  can be written as

$$pwm(t) = y(t) + \sum_{k=1}^{\infty} \frac{2}{k\pi} [\sin(2\pi k f_c t) - (-1)^k \cdot \sin(2\pi k f_c t - k\pi y(t))] \quad (10)$$

where  $y(t)$  is given by (7). Notice that  $pwm(t)$  contains every carrier harmonic with amplitude decreasing inversely with order, and also the signal  $y(t)$  in the baseband together with  $y(t)$  phase-modulated onto every carrier harmonic.

### C. Demodulation of UPWM

Consider now the demodulation process. When  $x(t)$  is a low-pass signal of bandwidth  $f_s$ , much of the distortion represented by the signals  $y_k(t)$  can be eliminated by low-pass filtering. That is, a simple demodulator for a PWM signal

consists of an ideal low-pass filter with cut-off frequency  $f_L$ . In this case, the distortion is dominated by the sum in (7). Take the demodulated signal to be  $z(t)$ . The filter output  $Z(f)$  is approximately

$$Z(f) \approx \exp(-j\pi fT)X(f) - \exp(-j\pi fT) \left(\frac{j\pi fT}{2}\right) S_2(f), \quad |f| \leq f_L \quad (11)$$

in the passband. It follows that, ignoring any truncation of the  $S_2(f)$  caused by the ideal low-pass filter, the output signal  $z(t)$  of the demodulator for a uniform-sampling TEPWM signal becomes approximately

$$z(t) \approx x(t - T/2) - \frac{T}{2} x(t - T/2) \frac{d}{dt} x(t - T/2). \quad (12)$$

The approximations indicate that the distortion reduces in amplitude inversely with the carrier frequency  $f_c$ . In particular, the power or energy in the distortion signal should reduce at the rate of approximately 6 dB per octave as  $f_c$  increases.

Now consider NPWM in which derivative distortion does not occur.

### D. The Spectra of NPWM Signals

In natural-sampling NPWM, the  $k$ th pulse width  $\tau_k$  is given by  $\tau_k = t_k - kT$  where  $t_k$  denotes the solution to the  $k$ th crossing-point equation:

$$t_k = kT + \frac{T[1 + x(t_k)]}{2}, \quad t_k \in [kT, (k+1)T].$$

A useful viewpoint is to consider that the signal resulting from natural sampling of  $x(t)$  is the same as UPWM sampling of  $\hat{x}(t)$ , where  $\hat{x}(t)$  is a distorted form of  $x(t)$ . The UPWM results can be applied to  $\hat{x}(t)$  to obtain the spectrum of  $\hat{p}(t)$ , where the caret distinguishes this NPWM signal from the UPWM signal  $p(t)$  studied above.

In order to determine  $\hat{x}(t)$ , a solution  $\hat{t}$  can be defined for the *generalized crossing-point equation*

$$\hat{t} = t + \frac{T}{2} x(\hat{t}) \quad \hat{t} \in [t - T/2, t + T/2] \quad (13)$$

where  $t$  is any fixed real number. Clearly, for each choice of  $t$ , there must be at least one solution to (13), and if the derivative of  $x(\hat{t})$  is smaller than  $2/T = 2f_c$ , the solution is unique. To have (13) define a one-to-one correspondence between  $t$  and  $\hat{t}$ ,  $x(t)$  must be such that

$$\left| \frac{dx(\hat{t})}{d\hat{t}} \right| < \frac{2}{T} = 2f_c, \quad -\infty < \hat{t} < \infty. \quad (14)$$

When  $x(t)$  is a low-pass signal of bandwidth  $f_s$ , a theorem of Bernstein (cf. [18, Theorem 6]) gives that

$$\max \left| \frac{dx(t)}{dt} \right| \leq 2\pi f_s \max |x(t)| = 2\pi f_s. \quad (15)$$

Thus, (14) holds for all possible choices of low-pass functions  $x(t)$  if it is required that  $f_c > \pi f_s$ . For some low-pass signals, a carrier frequency as small as the Nyquist frequency  $2f_s$  will suffice to ensure that (14) is satisfied, but more generally  $f_c > \pi f_s$  is needed for PWM.

With  $\hat{t}$  related to  $t$  uniquely via (13), the function  $\hat{x}(t)$  can be defined as

$$\hat{x}(t) = x(\hat{t}), \quad -\infty < t < \infty. \quad (16)$$

Notice that the  $k$ th natural-sampling pulse width  $\tau_k = T[1 + x(t_k)]/2$  for the signal  $x(t)$  equals  $T[1 + \hat{x}((k + (1/2))T)]/2$ , that is, the  $k$ th *uniform-sampling pulse width* for the signal  $\hat{x}(t + T/2)$ . Also,  $|\hat{x}(t)| = |x(\hat{t})| \leq 1$ , and thus  $\widehat{pw\hat{m}}(t)$ , the NPWM signal based on  $x(t)$ , is the same as  $pw\hat{m}(t, \hat{x}(t + T/2))$ , the UPWM signal based on  $\hat{x}(t + T/2)$ . Therefore, (7) leads to the conclusion that  $\widehat{pw\hat{m}}(t)$  includes the signal

$$\begin{aligned} \hat{y}(t) &= y(t, \hat{x}(t + T/2)) \\ &= \hat{x}(t) + \sum_{n=1}^{\infty} \frac{1}{(n+1)!} \left(\frac{-T}{2}\right)^n \frac{d^n}{dt^n} \hat{x}^{n+1}(t) \end{aligned} \quad (17)$$

with spectrum

$$\hat{Y}(f) = \sum_{n=1}^{\infty} \frac{(-j\pi f T)^{n-1}}{n!} \hat{S}_n(f) \quad (18)$$

where  $\hat{S}_n(f)$  is the Fourier transform of  $\hat{x}^n(t)$ . More generally, the signal-dependent portion  $\hat{p}(t)$  of the natural-sampling TEPWM signal has spectrum

$$\hat{P}(f) = \hat{Y}(f) + \sum_{k=1}^{\infty} \hat{Y}_k(f) \quad (19)$$

where  $\hat{Y}(f)$  is given in (18), while for  $k \geq 1$

$$\hat{Y}_k(f) = (-1)^k \sum_{n=1}^{\infty} \frac{(-j\pi f T)^{n-1}}{n!} [\hat{S}_n(f - kf_c) + \hat{S}_n(f + kf_c)]. \quad (20)$$

Equations (17)–(20) are unsatisfactory in that the baseband signal  $\hat{y}(t)$  as well as the spectra  $\hat{Y}_k(f)$  and  $\widehat{PW\hat{M}}(f)$  are specified in terms of  $\hat{x}(t)$  and its powers instead of  $x(t)$  and its powers. The problem can be remedied by showing that, in fact,  $\hat{y}(t)$  as given in (17) is *precisely*  $x(t)$ : when NPWM is used, there is *no harmonic distortion* (or  $T/2$  delay) in the baseband. The proof that  $\hat{y}(t) = x(t)$  is based on a result due to Lagrange (cf. [19, Section 7.32]). Details are provided in [16], where it is shown that

$$x(t) = \hat{x}(t) + \sum_{n=1}^{\infty} \frac{1}{(n+1)!} \left(\frac{-T}{2}\right)^n \frac{d^n}{dt^n} \hat{x}^{n+1}(t). \quad (21)$$

The right side of (21) is the same as the right side of (17) and hence,  $\hat{y}(t) = x(t)$  which is the result sought. Thus it has been shown that when natural sampling is used, the modulating signal  $x(t)$  appears *without harmonic distortion or delay* in the PWM modulator output. Previously, this result had been proved only for single-tone and double-tone signals.

From this analysis, (19) becomes

$$\begin{aligned} \widehat{PW\hat{M}}(f) &= X(f) + \sum_{k=1}^{\infty} (-1)^k \sum_{n=1}^{\infty} \frac{(jk\pi)^{n-1}}{n!} \\ &\cdot [S_n(f + kf_c) + (-1)^{n-1} S_n(f - kf_c)]. \end{aligned} \quad (22)$$

We can write the NPWM signal  $\widehat{pw\hat{m}}(t)$  as

$$\begin{aligned} \widehat{pw\hat{m}}(t) &= sq(t) + \hat{p}(t) \\ &= x(t) + \sum_{k=1}^{\infty} \frac{2}{k\pi} [\sin(2\pi k f_c t) - (-1)^k \\ &\quad \cdot \sin(2\pi k f_c t - k\pi x(t))]. \end{aligned} \quad (23)$$

Note that  $\widehat{pw\hat{m}}(t)$  consists of the modulating signal  $x(t)$  together with every carrier harmonic, and of  $x(t)$  phase-modulated onto each carrier harmonic, with amplitudes decreasing inversely with  $k$ . Thus there is an equivalence between a PWM process and a phase modulation process.

To our surprise, we have not found the representation of NPWM shown in (23) in the literature, even though it seems reasonable to assume that such a result must be quite well known. After all, it is well known that phase-modulation of a carrier by a sinusoid produces a spectrum involving Bessel functions, and, of course, so does NPWM with sinusoidal input. Surely someone must already have compared the spectra and discovered that they were identical, and thus deduced that (23) holds for the special case when  $x(t)$  is a sinusoid.

#### E. Demodulation of NPWM

Consider demodulation in the case of NPWM. As with UPWM, when  $x(t)$  is a low-pass signal of bandwidth  $f_s$ , an ideal low-pass filter with cut-off frequency  $f_L \geq f_s$  can be used to demodulate the natural-sampling TEPWM signal. Most of the statements made in Section II-C are applicable here. In particular, the spectrum of the demodulated signal  $\hat{z}(t)$  is

$$\begin{aligned} \hat{Z}(f) &= X(f) + \sum_{n=m}^{\infty} \sum_{k=1}^{N(n)} \frac{(jk\pi)^{n-1}}{n!} \\ &\cdot [S_n(f + kf_c) + (-1)^{n-1} S_n(f - kf_c)], \quad |f| \leq f_L \end{aligned} \quad (24)$$

where

$$N(n) = \left\lfloor \frac{nf_s + f_L}{f_c} \right\rfloor. \quad (25)$$

In contrast to UPWM, there is no distortion in the baseband except for the signal components corresponding to  $S_n(f \pm kf_c)$ ,  $k > 0$  that alias into the filter passband. We can write that

$$\hat{Z}(f) \approx X(f), \quad |f| \leq f_L \quad (26)$$

which suggests that the NPWM signal can be demodulated without any distortion. This is not strictly true, of course. In fact, it can be shown that the signal components aliasing into the filter passband are much larger in NPWM than in UPWM. In the latter, these components are considerably smaller than the baseband distortion described in (7) and hence do not contribute appreciably to the total distortion. With NPWM, however, these signal components are the only source of distortion in the baseband. Let  $m$  be the smallest positive integer for which  $N(m) = \lfloor (mf_s + f_L)/f_c \rfloor = 1$ . Then, neglecting the terms for  $k > 1$ ,  $n > m$  in (24), we can write

$$\begin{aligned} \hat{Z}(f) &\approx X(f) + \frac{(j\pi)^{m-1}}{m!} [(-1)^{m-1} S_m(f - f_c) + S_m(f + f_c)] \\ &\quad |f| \leq f_L. \end{aligned} \quad (27)$$

TABLE II  
RTD IN dB FOR TEPWM,  $x(t) = \text{sinc}^2(100t)$

$f_c/f_s$	2	3	4	6	8
Uniform Sampling	-3.1	-6.0	-8.5	-12.0	-14.5
Natural Sampling	-16.9	-39.7	-67.0	-132.8	-207.3

Since  $\pi^{m-1}/m!$  is a decreasing function of  $m$  for  $m > 3$ , the distortion aliasing into the baseband can be reduced by increasing  $m$ , that is, by choosing  $f_c \gg f_L, f_s$ . In fact, since  $m! \sim \exp(O(m \ln m))$ , and  $S_m(f \pm f_c)$  is small for  $|f| \leq f_L$ , the magnitude  $|\hat{Z}(f) - X(f)|$  of the distortion spectrum should decrease as  $\exp(-O(m \ln m))$  as  $m$  increases. Numerical results given in the next subsection seem to indicate that asymptotically the distortion might be decreasing as fast as  $\exp(O(m^2))$ .

#### F. Example

For UPWM, a *distortion signal* can be defined as  $z(t) - x(t - T/2)$  where  $z(t)$  is the inverse Fourier transform of  $Z(f)$  as given in (11). For NPWM, a distortion signal is defined as  $\hat{z}(t) - x(t)$  where  $\hat{z}(t)$  is the inverse Fourier transform of  $\hat{Z}(f)$  as given in (24). The total distortion (TD) is the energy or power of the distortion signal, and the relative total distortion (RTD) is defined as the ratio of the total distortion to the energy or power of the input signal in the filter passband, that is

$$\text{RTD} = \frac{\text{TD}}{\int_{-f_L}^{f_L} \Phi_x(f) df} \quad (28)$$

where  $\Phi_x(f)$  is the energy or power spectral density of  $x(t)$ , and the integral in the denominator of (28) gives the total signal energy or power in the demodulating filter output.

As an example, consider the finite energy modulating signal  $x(t) = \text{sinc}^2(100t)$  which has bandwidth  $f_s = 100$  Hz. Choose  $f_L = 100$  Hz also, and note that this particular input signal satisfies (14) if  $f_c \geq 2f_s$ . Table II shows the RTD in decibels for different switching ratios  $f_c/f_s$  for UPWM and NPWM. These results were obtained by numerical integration of (28). For UPWM, when the switching ratio doubles, the RTD drops by about 6 dB. This supports the conclusion from (12). In comparison, the RTD of NPWM decreases much faster. For NPWM, the values of the RTD in dB shown in Table II can be approximated by the function  $-1.525(f_c/f_s)^2 - 16.6(f_c/f_s) + 22.75$ , that is, the distortion is decreasing as  $\exp(-O(m^2))$ . For a switching frequency more than about eight times the signal bandlimit, distortion in the NPWM case is below the 24-b equivalent level in Table I.

#### G. PCM-to-PWM Conversion

So far, it has been demonstrated that NPWM, in the context of a sufficiently high switching frequency, can in principle lead to a high-fidelity result. Now we wish to determine how NPWM can be used in the context of a signal that is already in digital form. The nature of a digital PWM process begins with conversion from the incoming data stream into pulse-width values. Consider the conventional CDA format. The audio information in CDA is stored as two channels of 16-b pulse-code modulation (PCM). Each digital word represents the amplitude of

the audio signal at a uniform sampling rate of 44.1 kHz (intervals of 22.7  $\mu$ s). Other digital formats all are fundamentally PCM systems, perhaps with higher rates or longer word lengths. A digital PWM process must convert the uniformly-sampled amplitude information into time information to establish the switching instants.

A straightforward conversion to PWM would be to translate amplitude variation directly into pulse width variation. The  $2^{16}$  possible heights correspond to  $2^{16} = 65\,536$  possible pulse widths. In principle, we could build a PWM signal that switches at 44.1 kHz and has any of these possible widths. This direct approach is not feasible for three reasons. First, Table II shows that the switching frequency is not far enough above the 20 kHz signal band to avoid baseband distortion. Second, this would be a UPWM process, since the waveform information is uniformly sampled. The output would contain harmonic distortion. Third, digital generation of these PWM pulses requires time resolution of  $(22.7 \mu\text{s})/2^{16} = 0.346$  ns, which is difficult to achieve.

Since direct amplitude-to-width conversion is not a valid solution, it is more logical to first use a higher switching frequency, then use a process that involves NPWM. Presumably a higher switching frequency would reduce the number of necessary pulse widths without deteriorating the baseband information. A typical inverter might begin with a  $10\times$  switching frequency (at least 200 kHz in this context). For more convenient data processing, it would be useful to use a multiple  $2^n$  of the CDA sampling rate as the switching frequency. A choice of  $2^2 \times 44.1$  kHz = 176.6 kHz is adequate but a bit low for convenient filtering. A choice of  $2^3 \times 44.1$  kHz = 352.8 kHz should be possible. Higher rates will push the performance of available power MOSFETs. Unfortunately, a higher switching frequency exacerbates the time resolution problem. At 352.8 kHz, time resolution for  $2^{16}$  pulse widths is just 43 ps. A noise-shaping process will be used to reduce the number of distinct pulse widths and bring the time resolution to a practical value.

Given a switching frequency that is a multiple of the sampling frequency, an upsampling process [20] can be used to generate a faster data stream suitable for the higher rate. Upsampling is the process of generating intermediate sample values by interpolation or by sample insertion and filtering. In a typical process, extra zero-value samples are inserted in the original data stream. Since the audio signal is bandlimited, conventional digital filtering can remove the high harmonics and provide extra samples that should track the original waveform. Upsampling is a common step in the signal reconstruction process in commercial digital audio systems. Our implementation starts with an  $8\times$  upsampling process, which provides a PCM audio data stream at the target 352.8 kHz switching frequency.

Once the upsampled data are available, it would be straightforward to generate UPWM directly from the samples, but again harmonic distortion would be produced. NPWM generation is computationally more complex, since signal crossings, as in Fig. 3, relative to a 352.8 kHz sawtooth or triangle must be determined. Our implementation converts from uniform to natural sampling in real time with an interpolation process. To retain high fidelity, a precise interpolation is essential. The issue is illustrated in Fig. 4, which shows an expanded view of two audio signals over a short time interval. Of interest is the actual value

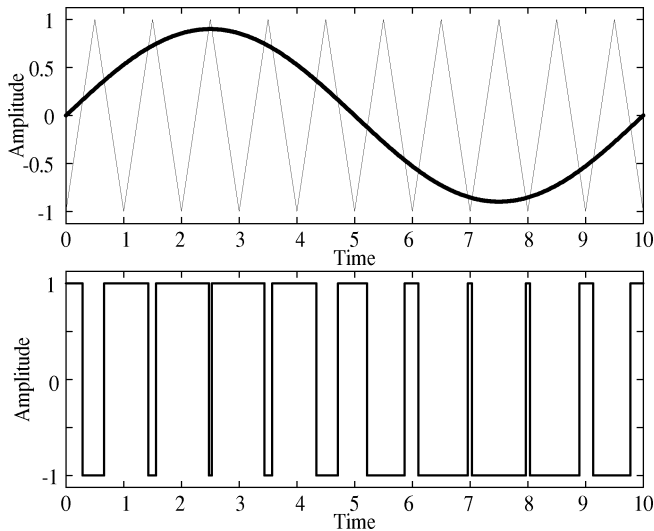


Fig. 3. Example of a naturally sampled PWM signal.

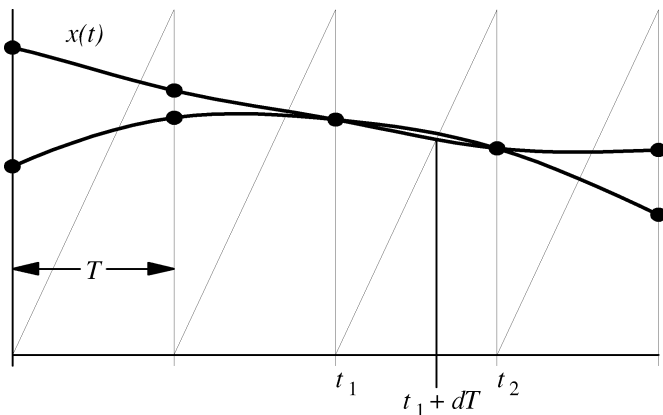


Fig. 4. Two audio signals over a few sample intervals.

of the signal  $x(t)$  at the crossing point,  $x(t_1 + dT)$ . Interpolation provides an estimate of this value. Consider an interpolation that uses just two sample values,  $x(t_1)$  and  $x(t_2)$ . The logical choice is linear interpolation, which will give the same estimate for both signals in this case. More generally, working from  $n$  sample values, a logical choice is Lagrange interpolation [21], in which a polynomial of degree  $n - 1$  is found to pass through the  $n$  points.

Lagrange interpolation requires division operations, which can be inconvenient in a DSP implementation, and can be computationally intensive. As a result, there have been many efforts to produce simplified interpolation methods for the purpose of PCM-to-NPWM conversion. Since the upsampled PCM stream can be interpreted directly as UPWM, the process can be considered a UPWM-to-NPWM conversion approach. In [9] and [22], corrections to linear interpolation are proposed. In [9], the approach attempts to adjust the sample process along a continuum between UPWM and NPWM, based on certain fitting parameters. The parameters are signal-dependent and must adapt throughout the waveform. In [22], it is proposed to add a curvature correction to linear interpolation. This can provide improvements for known single-frequency signals, but for random audio signals (as in Fig. 4) there is no consistent way to modify

a linear interpolation. In [23], a predictor–corrector approach is taken to the interpolation process. In [24], a Newton–Raphson method is applied iteratively to estimate the crossing times. The performance can be made high—if a mathematical representation of the signal is available. The need for such an expression and the iterative process leads to high computational complexity for the Newton–Raphson method.

Lagrange interpolation, used directly, does not provide an error estimate. A Taylor series analysis (which assumes ideal knowledge of derivatives as well as the samples themselves) for linear interpolation suggests that the error magnitude is bounded by about half the square of the sampling interval times the second derivative of the signal. For CDA, the interval is  $2.83 \mu\text{s}$ , and the second derivative magnitude can be as high as  $\omega^2$ . For a 20 kHz signal, the product  $\omega^2 T^2/2 = 0.065$ , and an error of several percent is possible. This is pessimistic, however, since audio signals rarely have strong high-frequency content and because harmonic distortion on this fast signal will not be audible. For a modulating frequency of 6.67 kHz (the highest one for which third-harmonic distortion might be audible), the factor  $\omega^2 T^2/2 = 0.007$ , or about  $-43$  dB relative to full scale. The interpolation algorithm introduced previously by our group [25] uses four neighboring uniform samples to estimate each natural sampling point. With  $n$  samples and Lagrange interpolation, the error should be bounded by  $\omega^n T^n/n!$ . For four samples and 6.67 kHz signals, this gives an error bound of  $8.3 \times 10^{-6}$ , or  $-102$  dB. This is a worst-case peak value. In simulation with a 6.67 kHz signal with 90% depth of modulation, a four-point Lagrange interpolation exhibited about  $-120$  dB of distortion [26].

Since  $-102$  dB is better than the CDA equivalent 16-b word length, some simplifications should be possible. The algorithm in [25] uses polynomial approximations to avoid divisions, and leads to a fast interpolation computation process compared to Lagrange interpolation. This is a good match to real-time implementation, at the expense of lower accuracy. This algorithm, as implemented in the experiments below, requires 14 multiplications, 11 additions, and 1 comparison per sample. The performance in simulations was about 6 dB worse than Lagrange interpolation, still providing an equivalent 16-b word length.

More recently, we have used an analysis based on Lagrange's Expansion Theorem [27], [19] to provide more accurate error estimates and improved algorithms. It can be shown [28] that the width of the  $k$ th pulse, under trailing-edge NPWM, is given by

$$\tau_k = \frac{T}{2} + \sum_{n=1}^{\infty} \left(\frac{T}{2}\right)^n \frac{1}{n!} \frac{d^{n-1}}{dt^{n-1}} x^n(t)|_{t=kT+T/2},$$

$$k = \dots, -1, 0, 1, 2, \dots \quad (29)$$

Since the derivatives are bounded, this series can be truncated to some number of terms to provide increasingly accurate estimates of the crossing point.

The truncation process also provides an error estimate. In this case, truncation to one term represents UPWM, while higher order terms require additional sample points to approximate derivatives. When four terms are used, for instance, derivatives through third-order are required, but no divisions are involved.



An algorithm that uses seven sample points to estimate three derivatives was able to interpolate the crossing point with 16 multiplications and 19 additions, and showed effective distortion of just  $-114$  dB for a 6.67 kHz signal. Algorithms based on the Lagrange expansion process offer excellent promise for future extension to precision digital audio.

### III. NOISE SHAPING

The PCM-to-NPWM conversion process preserves the resolution of the data stream if the interpolation is accurate. The natural sampling cross points are computed with the same number of bits as the PCM word length. For CDA signals, the converted results provide 16-b resolution for 352.8 kHz samples—corresponding to a time resolution of 43 ps. In order to avoid an excessive clock rate for PWM generation, the number of bits per pulse and, consequently, per sample, can be reduced. Simple truncation would reduce the word length but would also introduce significant quantization noise. The noise power would be uniformly spread across the spectrum, so much of it would lie inside the audio band and would be audible at the output.

There is a well known way to reduce the number of bits without compromising the signal-to-noise ratio (SNR) inside the audio band. Noise-shaping [15], [29], [30] introduces about the same total noise power as truncation, but the noise power spectral density is shaped as part of the process. If the quantization noise is made to be low inside the audio band and high above it, the number of bits can be reduced while retaining the same in-band SNR. Noise-shaping can be performed only if there is a frequency band to which the noise power can be redirected, out from the audio band. Since the signal has already been upsampled, the band to which the noise power can be redirected extends from 20 kHz up to 176.4 kHz (half the sampling rate).

From a power electronics perspective, noise-shaping can be understood in part in terms of fast averaging. Consider a dc–dc converter in which the duty ratio resolution is just 10%, but in which the desired duty ratio is 0.433. If the switching frequency is high, the duty ratio can alternate between 0.4 and 0.5 so as to deliver 40% duty about 2/3 of the time and 50% duty about 1/3 of the time. There will be additional ripple (at a frequency lower than  $f_c$ ), but a precise average can be generated if the extra ripple can be filtered or tolerated.

The noise-shaping technique is shown in block form in Fig. 5. Most of the prior work in noise-shaping deals with a 1-b sample data stream, but this is not a severe limitation. Here,  $\hat{x}[k] = \hat{x}(kT)$  is the 16-b NPWM information coming in at 352.8 kHz and  $v[k]$  is chosen to be an 8-b representation of the same information at this rate. In general, we need an optimal solution of  $H(z)$  such that the noise power is minimized in the audio band. To facilitate implementation, a simpler approach is to realize an  $n$ th-order FIR filter with integer coefficients, to provide the form

$$H(z) = (1 - z^{-1})^n.$$

It can be shown that  $n = 3$  supports the reduction to an 8-b word length while preserving equivalent 16-b word-length per-

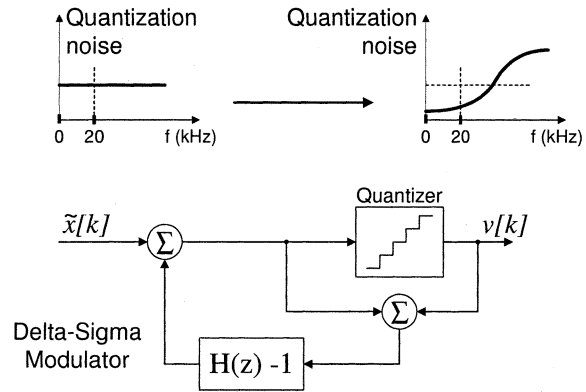


Fig. 5. Top: bit reduction by means of noise-shaping preserves in-band SNR. Bottom: noise-shaping can be implemented by means of delta-sigma modulators.

formance in the audio band, given 352.8 kHz switching. The value  $n = 4$ , with

$$H(z) = (1 - z^{-1})^4 = z^{-4} - 4z^{-3} + 6z^{-2} - 4z^{-1} + 1$$

is slightly (about 1 dB) better with a similar complexity level. Other choices of  $H(z)$  are possible [31].

One challenge is the effect of the noise-shaping process on the crossing-time estimates from the PCM-to-NPWM conversion. In effect, after noise-shaping the crossing times are “contaminated” by the output quantization noise from the noise shaper. We can consider the time estimates for NPWM to be a separate *correction function*,  $c(t)$ , with a bandwidth approaching half the switching frequency, to the upsampled UPWM data stream. Although the noise-shaper output does not interfere strongly in the audio band, it does have the potential to interfere with higher-frequency aspects of the correction function. In practice, we have found that noise-shaping filters with  $n > 4$  actually increase the total in-band noise. This is likely to result from aliasing with the correction function. In the literature, noise shapers with orders no higher than five are usually recommended, although no detailed investigations of the issues have been reported. Noise-shaping effects in PWM processes are a possible area of future study.

The output  $v[k]$  from the noise-shaping filter is a list of switching times that, thanks to the bit reduction, requires PWM pulses with only 256 different possible widths in each 2.83  $\mu$ s period. Inside the audio band, the 8-b and 16-b representations are almost equivalent, while extra quantization noise is present in the ultrasonic range. The time resolution needed to construct the PWM wave is now 2.83  $\mu$ s/256 = 11.1 ns. This corresponds to a clock frequency of 90.3 MHz, which at present can be found even in relatively low-cost circuits. So far, the fidelity still corresponds to 16-b word length, with a set of signal processing algorithms that are straightforward to perform in real time on a typical DSP chip. The extension to 24-b audio requires interpolation methods that use a few more sample values and noise-shaping that retains a few more bits for

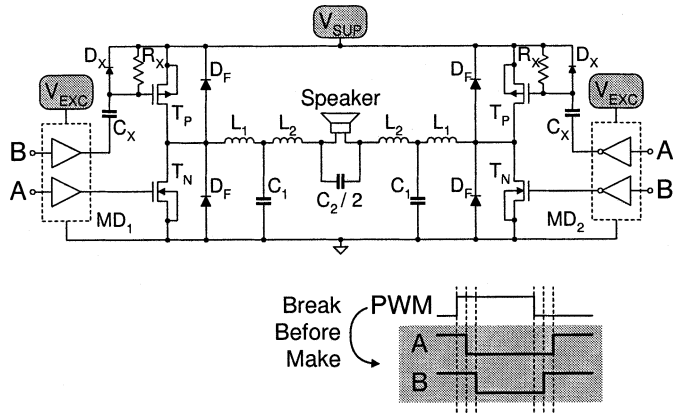


Fig. 6. Inverter circuit.

pulse-width resolution. Better time resolution can be supported by techniques such as those in [32], in which a digital process controls a much faster ring oscillator that provides finer pulse width values.

#### IV. PWM GENERATOR AND INVERTER

The 8-b signal  $v[k]$  in Fig. 5 can now be converted to a logic-level PWM waveform with a simple counter. To avoid PWM saturation, the counter inserts margins at each pulse edge in order to guarantee a minimum and a maximum pulse width. This reduces the dynamic range but minimizes the effect of any nonlinearities in the power stage. Thirty-two extra clock cycles are inserted as margins in each pulse period, and therefore the actual counter clock frequency is  $352.8 \text{ kHz} \times (256 + 32) = 101.6 \text{ MHz}$ . The logic counter also is used to generate dead times that ensure break-before-make (BBM) switch operation and avoid shoot-through currents [33] in the power stage.

Fig. 6 shows the PWM inverter power stage. It is a conventional full-bridge complementary MOSFET inverter followed by a passive low-pass filter. Level-shifting circuitry is used for the high-side gate drivers. There is nothing special about the devices—our experimental test unit uses low-cost IRF520 parts for the  $N$ -channel devices. Because of the dead time, there are two separate signals, A and B, for each audio channel. The MOSFETs never create a direct path from the supply line to ground.

The passive filter is not essential to output performance, since the loudspeaker itself will provide a low-pass function, but the use of an additional filter smooths the speaker current and avoids excess energy loss in the loudspeaker coils and magnetics. It also avoids radiation of the fast PWM waveform. The filter design uses a four-pole Butterworth characteristic, with a corner frequency high enough to avoid both amplitude and phase distortion in the audio band.

The circuit in Fig. 6 does not provide power supply rejection, and indeed adjustment of the power supply rail provides a convenient volume control. Since the output signal is proportional to the inverter rail voltage, it is important that the rail should be free of in-band noise and interference. On the other hand, an adjustable switching power supply, linked to the 352.8 kHz PWM frequency, would not add distortion at the output.

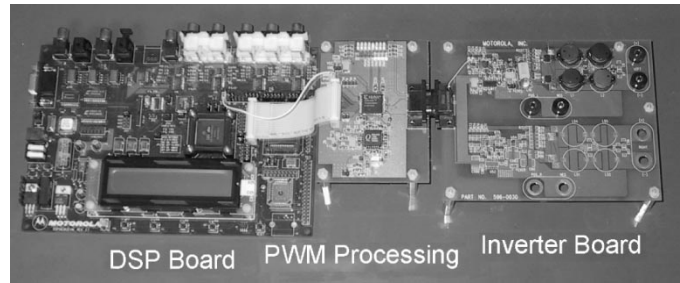


Fig. 7. Complete experimental system for 15 W amplifier.

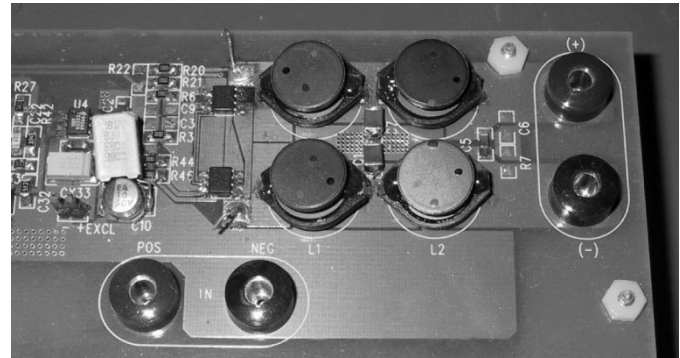


Fig. 8. Expanded view of inverter stage.

It might be expected that device forward-drop would lead to additional distortion. However, this is not the case when MOSFETs are used. Since the devices act as resistors when on, the circuit analysis shows that the drops are exactly equivalent to placing a resistor in series with the loudspeaker. This reduces the available output amplitude based on a voltage-divider effect, but does not introduce distortion. In principle, there would be no distortion at all with matched devices, but in practice even discrete devices give good performance. This effect can be tested by reducing the power supply rail voltage. In our experimental system, audio fidelity is maintained even at rail voltages below 1 V. The loudspeaker output sounds fine even with a rail voltage of 0.1 V or less!

It is possible to provide power supply rejection and performance improvement either with an augmented PWM process that takes into account a sensed supply voltage, or by using a special “feedback” approach. In essence, response to power supply variation or other nonlinearities can be viewed as a small correction to the pulse-width provided by the signal-processing algorithm. For signal fidelity, the question is whether the actual switching times observed at the inverter output match the intended switching times into the gate drives (plus some delay). For power supply rejection, the question is whether the volt-seconds delivered to the speaker match the commanded values. In either case, this is a compensation process that involves the “raw” PWM waveform, as opposed to a feedback process based on the intended analog output signal. Through a determination of whether the output should be higher or lower, the pulse widths can be adjusted one bit at a time until optimum performance is obtained. The process is similar in some ways to correction methods described in [13].

TABLE III  
POWER EFFICIENCY AND DISTORTION OF THE PUSH-PULL STAGE WITH THROUGH-HOLE COMPLEMENTARY POWER MOSFETS, DIGITAL BBM, AND NO FEEDBACK. THE INPUT SIGNAL IS A 1-KHz SINUSOID

Input signal (dBFS)	$V_{SUP}$ (V)	Output power (W)	Efficiency (%)	THD (%)
-1	43	55.0	87.97	0.22
0	43	68.9	89.04	0.20
0	48	84.1	87.62	0.40
0	50	90.1	87.07	0.56

TABLE IV  
RESULTS WITH FEEDBACK

Input level (dBFS)	$V_{SUP}$ (V)	$P_{OUT}$ (W)	Efficiency (%)	THD+N (%)
-1	43	50.6	81.72	0.022
0	43	63.8	83.40	0.056
-1	45	51.3	79.70	0.020
0	45	64.8	82.30	0.050
-1	48	54.4	79.88	0.025

## V. EXPERIMENTAL RESULTS

The hardware stage is shown in Figs. 7 and 8. Fig. 7 shows a commercial Motorola DSP development board which processes a digital optical signal to produce real-time PWM. An output conversion board comprises an FPGA that implements the PWM counter and creates the square waves. The output is then delivered to the inverter stage at the right. The inverter stage is shown more closely in Fig. 8. The MOSFETs are in surface-mount packages (two of the four are on the bottom of the board). The four inductors for the output low-pass filter are clearly visible. The banana jacks for input dc power and output audio power provide a scale. This small output circuit is rated to deliver up to about 15 W continuous, with no heat sinking other than the circuit board metallization. An inverter version with through holes for TO-220 MOSFET packages was able to deliver up to 90 W into a 6  $\Omega$  load, and was used for the output data given below. The layouts were arranged to avoid crosstalk among the gate drives, but there are no special precautions for shielding.

Table III shows measured distortion results for a 1 kHz sinusoidal test signal with the uncompensated system. These are purely open-loop results, with no power supply rejection and no input other than power and the digital signal. At 100% modulation with a 43 V supply, the inverter delivers almost 70 W into an 8  $\Omega$  load with 89% efficiency. The total distortion and noise of 0.2% represents a level of -54 dB, close to the target high-fidelity level, and comparable to the performance of our purchased comparison amplifier.

Table IV shows measured distortion for the complete system, including output compensation to eliminate power supply sensitivity. Distortion pulse noise levels of 0.02% harmonic distortion at 51 W output compare favorably with high-end commercial-grade audio amplifiers. The efficiency of 80% or more at these levels is a dramatic departure from commercial hardware. This distortion level demonstrates that a PWM inverter

can achieve high fidelity at a significant output power level. Additional work with improved interpolation and adjustments to the correction process has pushed the output distortion plus noise result down to about -80 dB at 50 W. Now that distortion levels better than 0.02% can be demonstrated, it is reasonable to expect PWM inverter technology to make significant inroads into amplifier applications over the next few years.

## VI. CONCLUSION

High-fidelity and PWM are not mutually exclusive. It has been shown in this paper that a signal-processing procedure that leads to 8-b resolution NPWM can produce audiophile fidelity at the output of a conventional full-bridge MOSFET inverter. It has also been shown that a fully digital processing sequence is possible, beginning from a digital audio data stream and ending at a PWM output into a filtered load. Prototype circuits have demonstrated 0.02% combined total harmonic distortion plus noise for a standard audio test signal at 51 W output and 82% efficiency, based on conventional low-cost TO-220 power MOSFETs. A smaller version, implemented with surface-mount MOSFETs and capable of about 15 W of audiophile output, was demonstrated at the 2000 IEEE Workshop on Computers in Power Electronics [34].

## REFERENCES

- [1] P. Z. Peebles, Jr, *Communication Systems Principles*. Reading, MA: Addison-Wesley, 1976, p. 333.
- [2] A. B. Carlson, *Communication Systems*, 3rd ed. New York: McGraw-Hill, 1986.
- [3] B. Wilson, Z. Ghassemlooy, and A. Lok, "Spectral structure of multitone pulse width modulation," *Electron. Lett.*, vol. 27, pp. 702-704, 1991.
- [4] P. H. Mellor, S. P. Leigh, and B. M. G. Cheetham, "Reduction of spectral distortion in class D amplifiers by an enhanced pulse width modulation sampling process," *Proc. Inst. Elect. Eng.*, vol. 138, pp. 441-448, Aug. 1991.
- [5] R. E. Ziemer and W. H. Tranter, *Principles of Communications*, 4th ed. New York: Houghton-Mifflin, 1995, p. 435.
- [6] H. B. Black, *Modulation Theory*. Princeton, NJ: Van Nostrand, 1953.
- [7] P. Wood, *Switching Power Converters*. Princeton, NJ: Van Nostrand, 1981.
- [8] P. Midya and P. T. Krein, "Noise properties of pulse-width modulated power converters: Open-loop effects," *IEEE Trans. Power Electron.*, pp. 1134-1143, Nov. 2000.
- [9] P. H. Mellor, S. P. Leigh, and B. M. G. Cheetham, "Digital sampling process for audio class D, pulse width modulated power amplifiers," *Electron. Lett.*, vol. 28, pp. 56-58, Jan. 1992.
- [10] J. M. Goldberg and M. B. Sandler, "New high accuracy pulse width modulation based digital-to-analogue convertor/power amplifier," in *IEE Proc.: Circuits, Devices, Syst.*, vol. 141, Aug. 1994, pp. 315-324.
- [11] K. M. Smith, K. M. Smedley, and Y. Ma, "Realization of a digital PWM power amplifier using noise and ripple shaping," in *Proc. IEEE Power Electron. Spec. Conf.*, 1995, pp. 96-102.
- [12] L. Risbo and T. Morch, "Performance of an all-digital power amplification system," in *Proc. 104th AES Convention*, May 1998, Preprint 4695.
- [13] K. Nielsen, "PEDEC—A novel pulse referenced control method for high quality digital PWM switching power amplification," in *Proc. IEEE Power Electron. Spec. Conf.*, 1998, pp. 200-207.
- [14] K. P. Sozanski, R. Strzelecki, and Z. Fedyczak, "Digital control circuit for class-D audio power amplifier," in *Proc. IEEE Power Electron. Spec. Conf.*, 2001, pp. 1245-1250.
- [15] R. E. Hiorns, J. M. Goldberg, and M. B. Sandler, "Design limitations for digital audio power amplification," in *Proc. IEE Colloq. Digital Audio Signal Processing*, 1991, pp. 4/1-4/4.
- [16] Z. Song and D. V. Sarwate, "The frequency spectrum of pulsewidth modulated signals," *Signal Processing*, 2002, submitted for publication.
- [17] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-Time Signal Processing*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 1999.

- [18] G. C. Temes, V. Barcion, and F. C. Marshall, III, "The optimization of bandlimited systems," *Proc. IEEE*, vol. 61, pp. 196–234, Feb. 1973.
- [19] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, 4th ed. Cambridge: Cambridge University Press, 1952.
- [20] T. J. Cavicchi, "DFT time-domain interpolation," in *Proc. Inst. Elect. Eng.*, vol. 139, June 1992, pp. 207–211.
- [21] E. Weisstein, "Lagrange interpolating polynomial," *World of Mathematics*, [Online]. Available: <http://mathworld.wolfram.com/LagrangeInterpolatingPolynomial.html>, May 28, 2001.
- [22] M. Johansen and K. Nielsen, "A review and comparison of digital PWM methods for digital pulse modulation amplifier (PMA) system," in *Proc. 107th AES Convention*, Sept. 1999, Preprint 5039.
- [23] P. Midya, B. Roeckner, P. Rakers, and P. Wagh, "Prediction correction algorithm for natural pulse width modulation," in *Proc. 109th AES Convention*, Sept. 2000, Preprint 5194.
- [24] J. M. Goldberg and M. B. Sandler, "Pseudo-natural pulse width modulation for high-accuracy digital-to-analogue conversion," *Electron. Lett.*, vol. 27, pp. 1491–1492, Aug. 1991.
- [25] C. Pascual and B. Roeckner, "Computationally efficient conversion from pulse-code-modulation to naturally-sampled pulse-width-modulation," in *Proc. 109th AES Convention*, Sept. 2000, Preprint 5198.
- [26] C. Pascual, "All-digital audio amplifier," Ph.D. dissertation, Univ. of Illinois, Urbana, IL, 2000.
- [27] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. Washington, DC: U. S. Govt. Printing Office, 1968.
- [28] Z. Song, "Digital pulse width modulation: Analysis, algorithms, and applications," Ph.D. dissertation, Univ. of Illinois, Urbana, IL, 2001.
- [29] A. Paul and M. Sandler, "Design issues for a 20-b D/A converter based on pulse width modulation and noise-shaping," in *Proc. IEE Colloq. Adv. A-D D-A Conv. Techn. Applicat.*, 1993, pp. 4/1–4/4.
- [30] S. R. Norsworthy, R. Schreier, and G. C. Temes, *Delta-Sigma Data Converters*. New York: IEEE Press, 1996.
- [31] P. Midya, M. Miller, and M. Sandler, "Integral noise-shaping for quantization of pulse width modulation," in *Proc. 109th AES Convention*, 2000, Preprint 5193.
- [32] J. Xio, A. Peterchev, and S. R. Sanders, "Architecture and IC implementation of a digital PWM controller," in *Proc. IEEE Power Electron. Spec. Conf.*, 2001, pp. 465–471.
- [33] P. T. Krein, *Elements of Power Electronics*. New York: Oxford University Press, 1998, p. 556.
- [34] C. Pascual, P. T. Krein, P. Midya, and B. Roeckner, "High-fidelity PWM inverter for audio amplification based on real-time DSP," in *Proc. IEEE Workshop Comput. Power Electron.*, 2000, pp. 227–232.



**César Pascual** was born in Teruel, Spain, in 1968. He received the B.S. degree in telecommunications engineering from the Polytechnic University of Catalonia, Spain, in 1992 and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana-Champaign (UIUC), in 1997 and 2000, respectively.

He was awarded a fellowship by the Spanish financial institution "la Caixa" to extend his studies in the United States. His professional experience includes teaching and research at the university as well as working in industry. From 1994 to 1995, he worked for SEMA Group S.A.E., Barcelona, Spain. From 1997 to 2000, he did research for the UIUC and worked for Motorola Labs, Schaumburg, IL. During 2001, he was with Massana Technologies, Madrid, Spain, and is currently a Design Engineer at Sharp Electronics Spain. He is the author of several international publications and is the recipient or corecipient of three U.S. patents.



**Zukui Song** was born in Anqing, Anhui, China, in 1972. He received the B.S. degree from Nanjing University, Nanjing, China, in 1993, the M.S. degree from Peking University, Beijing, China, in 1996, and the Ph.D. degree from University of Illinois at Urbana-Champaign (UIUC), in 2001, all in electrical engineering.

He was a Technical Consultant with Andersen Consulting (Accenture), in 1997. He was a Teaching Assistant in the Department of Electrical and Computer Engineering, UIUC, from August 1998 to December 1999. From January 2000 to December 2001, he was a Research Assistant in the Coordinated Science Laboratory, University of Illinois. Since January 2002, he has been with Aware, Inc., Bedford, MA, where he is currently a Digital Signal Processing Engineer. His research interests are communication systems and digital signal processing.



**Philip T. Krein** (S'76–M'82–SM'93–F'00) received the B.S. degree in electrical engineering and the A.B. degree in economics and business from Lafayette College, Easton, PA, and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois, Urbana.

He was an Engineer at Tektronix, Beaverton, OR, with responsibilities in analog design and product development. He then returned to the University of Illinois where he is now the Grainger Director's Chair Professor of Electrical and Computer Engineering, and the Director of the Grainger Center for Electric Machinery and Electromechanics. At the University, he has introduced lecture and laboratory courses in power electronics and electric drives, and performed research in large-signal analysis, design, and control of power electronic circuits and systems. He is the author of an undergraduate textbook, *Elements of Power Electronics* (London, UK: Oxford University Press, 1998). From 1997 to 1998, he was a Senior Fulbright Scholar at the University of Surrey, Guildford, UK. He holds seven U.S. patents.

Dr. Krein received the IEEE Industry Applications Society First Prize Paper Award in 1990. He is a member of Phi Beta Kappa, Tau Beta Pi, Eta Kappa Nu, and Sigma Xi. He is Past President of the IEEE Power Electronics Society. He served as General Chair of the 1997 IEEE Power Electronics Specialists Conference. He was selected as a University Scholar at the University of Illinois for 2000–2002.



**Dilip V. Sarwate** (S'68–M'73–SM'78–F'90) received the B.S. degree in physics and mathematics from the University of Jabalpur, Jabalpur, India in 1965, the B.Eng. degree in electrical communication engineering from the Indian Institute of Science, Bangalore, in 1968, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 1973.

Since January 1973, he has been with the University of Illinois at Urbana-Champaign, where he is currently a Professor of electrical and computer engineering, and a Research Professor in the Coordinated Science Laboratory. His research interests are in the general areas of communication systems and information theory, with emphasis on multiuser communications, error-control coding, and signal design.

Dr. Sarwate has served as the Treasurer of the IEEE Information Theory Group, as an Associate Editor for Coding Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY, and as a member of the Editorial Board of the PROCEEDINGS OF THE IEEE. He was a Co-Chairman of the Eighteenth, Nineteenth, Thirty-First and Thirty-Second Annual Allerton Conferences on Communication, Control, and Computing held in 1980, 1981, 1993, and 1994. In 1985, he served as a Co-Chairman of the Army Research Office Workshop on Research Trends in Spread Spectrum Systems. He has also been a member of the program committees for the IEEE Symposia on Spread Spectrum Techniques and Their Applications (ISSSTA) and the 1998 and 2001 Conferences on Sequences and Their Applications (SETA), as well as of several advisory committees for international conferences.



**Pallab Midya** (SM'96) received the B.Tech. degree from the Indian Institute of Technology, Kharagpur, in 1988, the M.S. degree from Syracuse University, Syracuse, NY, in 1990, and the Ph.D. in electrical engineering from the University of Illinois at Urbana-Champaign in 1995.

Since 1995, he has been with Motorola Labs, Schaumburg, IL, where he is currently a Distinguished Member of the Technical Staff. He has 13 issued and 10 pending U.S. patents. He has authored 25 IEEE, IEE, AES papers. He introduced Supply Modulation as a method to increase RF transmitter efficiency. He was the chief architect of an IC that doubled the transmitter efficiency of Motorola's satellite phone. This technology has since been incorporated into Motorola cellular phones. He initiated a research program for Direct Digital RF Generation and Amplification to create modulated RF signal using a switching amplifier. He has been a technical lead in Motorola's Digital Switching Audio Amplifier program.

Dr. Midya received the Richard Bass Outstanding Young Power Electronics Engineer Award from IEEE and the Young Alumni Achievement Award from the University of Illinois in 2002. He is a member of Sigma Xi, HKN Electrical Engineering Honor Society, and the Audio Engineering Society.



**William (Bill) J. Roeckner** received the B.S. and M.S. degrees in electrical engineering from the University of Cincinnati, Cincinnati, OH, in 1983 and 1986, respectively.

In 1987, he joined the Gould Central Research Labs, Rolling Meadows, IL, where he was involved with GaAs IC process development for high-speed digital applications. Since 1988, he has been with Motorola Labs, Schaumburg, IL, where he has been involved in the development and design of numerous mixed-mode CMOS ICs for both automotive and cellular applications. Most recently, he has been focused on the development of high efficiency digital audio amplifier technology. He has eight U.S. patents.