Validity of Uniform Quantization Error Model for Sinusoidal Signals Without and With Dither

MAHMOUD FAWZY WAGDY, SENIOR MEMBER, IEEE, AND WAI-MAN NG

Abstract—A technique is presented for determining the probability density function (PDF) and variance of the quantization error of a sinusoidal signal applied to a uniform quantizer, namely, an ideal A/D converter. The results are the basis for determining the validity of the uniform quantization noise model for this class of signals.

When dither is employed it tends to decorrelate quantization error and the input signal. Effect of added uniform dither on the PDF of the quantizer input is investigated, as well as the PDF of the quantization error corresponding to the dithered sinusoid. The results prove the validity of the uniform quantization model in this case.

I. Introduction

AMPLITUDE quantization has been extensively investigated in the past four decades. Widrow [1] constructed the probability density function (PDF) of quantization noise from the PDF of the quantizer input. He also [2] evaluated quantitatively the distortion resulting from rough quantization, and made a statistical analysis of amplitude quantized sampled data systems. More recently, Sripad and Snyder [3] gave a necessary and sufficient condition for quantization errors to be uniform and white. Since quantization noise constitutes a major source of errors in instrumentation, especially with A/D converters of 10 bits or less, it is important to develop more accurate quantization error models. Since the most common signal is the sinusoidal one, a major objective of this paper is to investigate the PDF and variance of the corresponding quantization error in terms of the amplitude of the sinusoid.

Dithering is an effective tool for reducing the effects of amplitude quantization. Schuchman [4] realized the conditions for a dither signal so that the quantizer noise becomes independent of the input signal. Vanderkooy and Lipshitz [5] used dither to resolve audio signals smaller than the quantizing step. The Japanese paper by Yamasaki [6] illustrated the effects of large amplitude dither on the spectrum of the quantizer output. Zames and Schneydor [7] showed that the behavior of a nonlinear system employing dither depends on the PDF of the dither signal. Lotto and Paglia [8] showed that dithering improves both differential and integral linearities of A/D converters. Again, due to the importance of the sinusoidal signal in instrumentation and measurement, another objective of this paper is to investigate the effect of different amplitudes of uniform dither on the PDF of the dithered sinusoid at the quantizer input. The PDF of the quantization error, in the presence of dither, is also obtained.

Manuscript received August 5, 1988; revised January 1, 1989.

The authors are with the Electrical Engineering Department, University of Lowell, Lowell, MA 01854.

IEEE Log Number 8926903.

II. PDF and Variance of Sinusoidal Signal Quantization Error

For the sinusoidal signal

$$x = A \sin \omega t \qquad (1)$$

A is the amplitude, ω is the radian frequency, and x lies in the dynamic range of the quantizer (bipolar A/D converter). x is regarded by the quantizer as a random variable. From (A-10) the PDF of x is given by

$$f_x(x) = \frac{1}{\pi\sqrt{A^2 - x^2}}, \qquad -A \leq x \leq A. \qquad (2)$$

For example, the PDF of x for two different sinusoids is shown in Fig. 1. A is a multiple of the quantizing step Δ, thus $A = 15.5$ on the figure means $A = 15.5\Delta$. Throughout the paper the quantizer input, x, is assumed to be a sequence of statistically independent random variables with PDF given by (2). However, x is actually obtained by sampling a sinusoidal function. It should be noted that these two methods give similar results only when the sampling frequency is not near to a harmonic or subharmonic of the sinusoid.

The PDF of the quantization error e is [3]

$$f_e(e) = \begin{cases} \dfrac{1}{\Delta} + \dfrac{1}{\Delta} \sum\limits_{n \neq 0} \phi_x\left(\dfrac{2\pi n}{\Delta}\right) \cdot \exp\left(\dfrac{-j2\pi ne}{\Delta}\right), \\ \qquad -\dfrac{\Delta}{2} \leq e \leq \dfrac{\Delta}{2} \\ 0, \qquad \text{otherwise} \end{cases} \qquad (3)$$

where $\phi_x(u)$ is the characteristic function of x and given by

$$\phi_x(u) = \int_{-\infty}^{\infty} f_x(x) \cdot \exp(jux)\, dx. \qquad (4)$$

Since $f_x(x)$ of (2) is an even function, then

$$\phi_x(u) = 2 \int_0^{\infty} \frac{\cos ux}{\pi\sqrt{A^2 - x^2}}\, dx. \qquad (5)$$

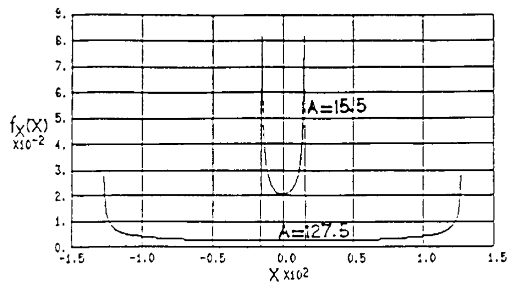

Fig. 1. PDF of two example sinusoids.

Using Fourier transform tables [9], it can be shown that

$$\phi_x(u) = J_0(Au) \tag{6}$$

where $J_0(z)$ is a Bessel function of zeroth order. By combining (3) and (6), and normalizing A and e with respect to Δ (which is equivalent to letting $\Delta = 1$), then

$$f_e(e) = 1 + \sum_{n \neq 0} J_0(2\pi n A) \cdot \exp(-j2\pi n e), \quad -\frac{1}{2} \leq e \leq \frac{1}{2}. \tag{7}$$

To simplify this expression, we consider the Bessel function of order v, given by [9]

$$J_v(z) = \sum_{m=1}^{\infty} \frac{(-1)^m \cdot (\frac{1}{2}z)^{v+2m}}{(m!)\Gamma(v+m+1)} \tag{8}$$

where Γ is the Gamma function, and $(m!)$ is factorial m. Consequently

$$J_0(2\pi n A) = \sum_{m=1}^{\infty} \frac{(-1)^m \cdot (\pi n A)^{2m}}{(m!)\Gamma(m+1)} \tag{9}$$

which indicates that $J_0(2\pi n A)$ is proportional to $(n)^{2m}$, and hence, an even function of n. Thus (7) can be rewritten as

$$f_e(e) = 1 + 2\sum_{n=1}^{\infty} J_0(2\pi n A) \cdot \cos(2\pi n e) \quad -\frac{1}{2} \leq e \leq \frac{1}{2}. \tag{10}$$

Using (10) the PDF of e is shown in Fig. 2 for $A = 1.5$, 3.5, 7.5, and 15.5. Computations revealed inflection points at $|e| = 0.475$ and 0.485. However, it is obvious that increasing A flattens the PDF of e and makes it closer to the rectangular-shaped quantization error model found in most of the relevant literature.

To determine the variance σ_e^2 of e we utilize the fact that $f_e(e)$ of (10) is an even function, and hence, has a zero mean, thus

$$\sigma_e^2 = \int_{-1/2}^{1/2} e^2 de + 2\sum_{n=1}^{\infty} J_0(2\pi n A) \cdot \int_{-1/2}^{1/2} e^2 \cdot \cos(2\pi n e)\, de. \tag{11}$$

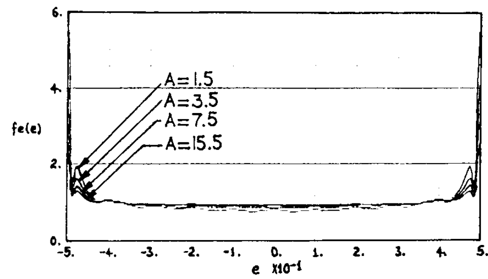

Fig. 2. PDF of quantization error for sinusoids using a new method based on Bessel functions.

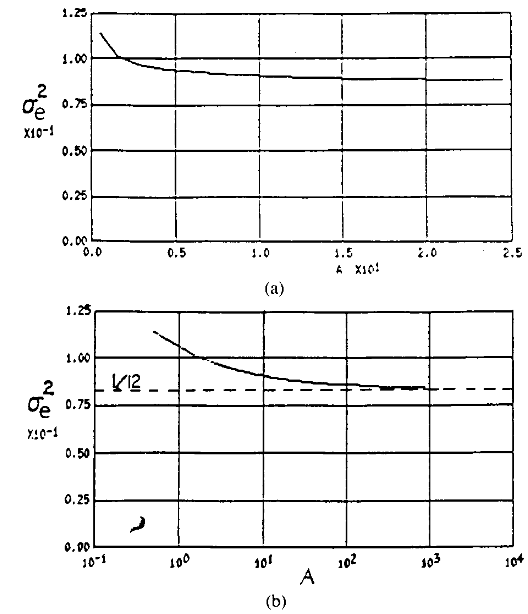

Fig. 3. Variance of quantization error versus sinusoid amplitudes.

Using a handbook of mathematical tables and formulas [10], it can be shown that

$$\int_{-1/2}^{1/2} e^2 \cdot \cos(2\pi n e)\, de = (-1)^n / 2\pi^2 n^2. \tag{12}$$

From (11) and (12) it follows that:

$$\sigma_e^2 = \frac{1}{12} + \frac{1}{\pi^2}\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cdot J_0(2\pi n A). \tag{13}$$

The variation of σ_e^2 versus A is shown in Fig. 3. It is obvious that as A increases, σ_e^2 approaches a normalized value of $1/12$. Given an error bound on σ_e^2 we can easily use (13) to determine the minimum amplitude of a sinusoidal signal in terms of the A/D converter quantizing step, Δ. Also, the propagation of quantization noise power in subsequent hardware stages, or software processing algorithms would be much more accurate for small sinusoids if (13) is used as opposed to the well-known value of $\Delta^2/12$.

III. Effect of Dithering on Quantizer Input and Output in Case of a Sinusoid

Dithering decreases the signal-dependent noise at the expense of increasing the signal-independent noise. A block diagram of the quantizer with added dithering is shown in Fig. 4. The PDF of the dithered signal, y, is related to that of the undithered signal, x, as follows:

$$f_y(y) = f_x(x) * f_d(d) \tag{14}$$

where $*$ denotes convolution. The PDF of x is given by (2), whereas the PDF of the dither d is shown in Fig. 5.

When the dither amplitude is smaller than the sinusoidal signal amplitude, i.e., $c < A$, it can be shown that (14) yields

$$f_y(y) = \begin{cases} \dfrac{1}{2\pi c}\left[\sin^{-1}\left(\dfrac{y+c}{A}\right) - \sin^{-1}\left(\dfrac{y-c}{A}\right)\right], \\ \qquad |y| \le A - c \\ \dfrac{1}{2\pi c}\cos^{-1}\left(\dfrac{|y| - c}{A}\right), \\ \qquad A - c \le |y| \le A + c \\ 0, \quad |y| > A + c. \end{cases} \tag{15}$$

The effect of changing c on the PDF of the quantizer input is shown in Fig. 6 for the case $A = 15.5\Delta$, for example. It is obvious that as c increases, $f_y(y)$ becomes less nonlinear and more flat; an advantage of larger amplitude dither.

When the dither amplitude is greater than the sinusoidal signal amplitude, i.e., $c > A$, it can be shown that (14) yields:

$$f_y(y) = \begin{cases} \dfrac{1}{2c}, \quad |y| \le c - A \\ \dfrac{1}{2\pi c}\cos^{-1}\left(\dfrac{|y| - c}{A}\right), \\ \qquad c - A \le |y| \le c + A \\ 0, \quad |y| > c + A. \end{cases} \tag{16}$$

The effect of changing c on the PDF of the quantizer input is shown in Fig. 7, for the case $A = 0.5\Delta$, for example. Again, $f_y(y)$ becomes more flat as c increases.

Let us now obtain the PDF of the quantization error in the presence of dither. To determine $f_e(e)$ using the same technique that led to (10), which is a closed form solution, would be rather difficult. Instead, we resort to Widrow's method [1] to derive the PDF of quantization noise via an additive linear process. $f_e(e)$ is the sum of the distributions of noise corresponding to the constituents that are added to get $f_y(y)$, the PDF of the quantizer input. This can be formulated as

$$f_e(e) = \sum_{n=-k}^{k} f_y(n\Delta + e), \quad \frac{-\Delta}{2} \le e \le \frac{\Delta}{2} \tag{17}$$

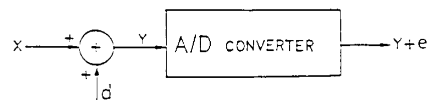

Fig. 4. The quantizer with added dither.

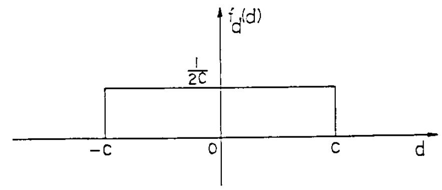

Fig. 5. PDF of the uniform dither signal.

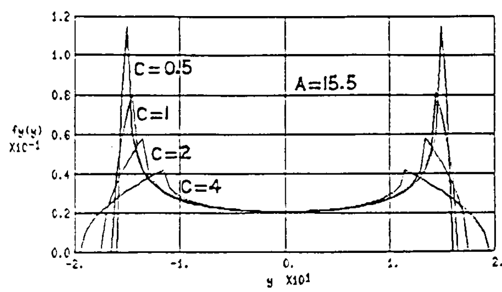

Fig. 6. PDF of the quantizer input for different dither amplitudes ($c < A$).

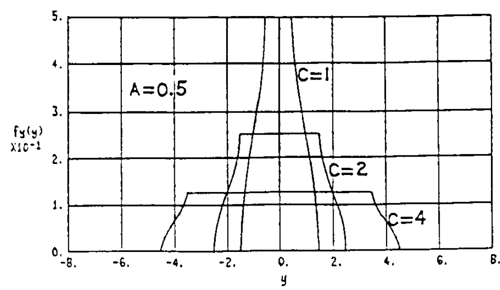

Fig. 7. PDF of the quantizer input for different dither amplitudes ($A < c$).

where $2k + 1$ is the number of quantization slots that can represent y.

With no dither $f_y(y)$ is given by (2) and $f_e(e)$, calculated via (17), is shown in Fig. 8 for $A = 1.5$, 3.5, 7.5, and 15.5. The PDF of e is obviously not flat-topped since $f_y(y)$ does not satisfy Widrow's quantization theorem [2], or the weaker condition obtained by Sripad and Snyder [3]. $f_e(e)$ of Fig. 8 is monotonic for $|e| > 0$, and is little different from Fig. 2 due to the difference in the methods.

When dither is present, $f_y(y)$ is given by (15) or (16), thus

$$k = \left(A + c - \frac{\Delta}{2}\right)\Big/\Delta. \tag{18}$$

It is obvious that if $2(A + c)$ is an odd number of Δ's, then y is represented by all the $2k + 1$ quantization slots,

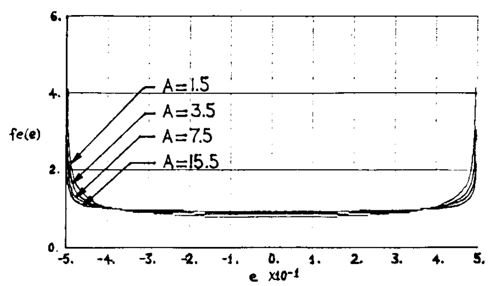

Fig. 8. PDF of quantization error for sinusoids using Widrow's method.

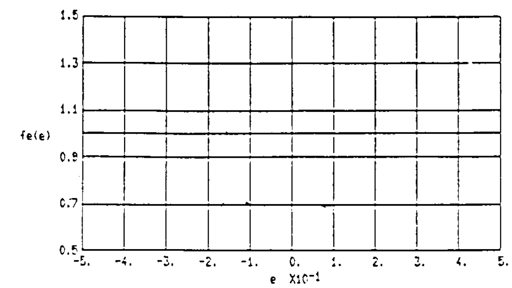

Fig. 9. PDF of quantization error for dithered sinusoids.

otherwise the highest quantization slots on both sides of the quantizer characteristics (bipolar A/D converter) will be partially filled. Several cases were investigated with the following normalized A and c: 1) $A = 2, 4, 8, 16$ with $C = 1/2$, 2) $A = 1\text{-}1/2, 2.2, 3\text{-}1/2, 7\text{-}1/2, 15\text{-}1/2$ with $C = 1$, and 3) $A = 3\text{-}1/2, 7\text{-}1/2, 15\text{-}1/2$ with $C = 2$. In these cases we consistently obtained a flat-topped PDF for the quantization error, as shown in Fig. 9. For the cases: 4) $A = 1.75$, $C = 0.25$, and 5) $A = 2.25$, $C = 0.75$, $f_e(e)$ is not flat-topped even though it is finite at $|e| = 0.5$. Cases 1)–5) indicate that for a flat-topped $f_e(e)$, the peak-to-peak dither magnitude should be an integral multiple, M, of the quantization slot, Δ, i.e.,

$$C = M\frac{\Delta}{2}. \tag{19}$$

This indicates that proper uniform dithering allows the well-known quantization error uniform model to be always valid.

IV. Conclusions

It is intuitively known that as the amplitude of a sinusoid increases the uniform quantization error model becomes more valid. Equation (10) is a closed-form formula which quantifies this intuitive observation. It is also known that the quantization noise variance $\Delta^2/12$ is more suitable for sinusoids when amplitudes are sufficiently large. Equation (13) is another closed-form formula which accurately quantifies noise variance for any value of the signal amplitude. This allows very accurate analysis and design for a wide variety of applications, especially those involving very small sinusoids.

Although dithering is a known concept, there are still some gaps in the literature on the subject. Equations (15) and (16) give more insight into the effect of uniform dither on the quantizer input. The computer simulations based on (17) truly verify that uniform dithering, of proper magnitude, renders the uniform quantization error model perfectly suitable for any sinusoid regardless of its amplitude.

Appendix
Determining the PDF of a Signal

The determination of the PDF of a signal is known by many researchers. However, it is worth presenting the PDF general formula along with a formal proof and an example application. To avoid generalities, we assume that the equation

$$x = g(t) \tag{A-1}$$

has three roots as in Fig. 10. The roots are denoted by t_i, $i = 1, 2, 3$, i.e.,

$$x = g(t_1) = g(t_2) = g(t_3). \tag{A-2}$$

Our objective is to find the set of values t, such that $x \le g(t) \le x + dx$, and the probability of this set in an arbitrary duration T. As we see from Fig. 10 this set consists of the following three intervals: $t_i \le t \le t_i + dt_i$, $i = 1, 2, 3$ where $dt_1 > 0$ and $dt_3 > 0$, but $dt_2 < 0$. It follows that:

$$\begin{aligned} P(x \le X \le x + dx) = {} & P(t_1 \le t \le t_1 + dt_1) \\ & + P(t_2 \le t \le t_2 + dt_2) \\ & + P(t_3 \le t \le t_3 + dt_3). \end{aligned} \tag{A-3}$$

Thus

$$P(x \le X \le x + dx) = \frac{dt_1}{T} + \frac{|dt_2|}{T} + \frac{dt_3}{T} \tag{A-4}$$

and since [11]

$$P(x \le X \le x + dx) = f_x(x) \cdot dx \tag{A-5}$$

where $f_x(x)$ is the PDF of x, then from (A-4) and (A-5) we deduce the following general theorem:

$$f_x(x) = \frac{1}{T}\sum_{i=1}^{n}\left|\frac{dt_i}{dx}\right| \tag{A-6}$$

where n is the number of real roots of the inverse function $t = \psi(x)$ derived from (A-1).

As an example, consider the sinusoid:

$$x = A \sin \omega t, \qquad 0 \le \omega t \le 2\pi \tag{A-7}$$

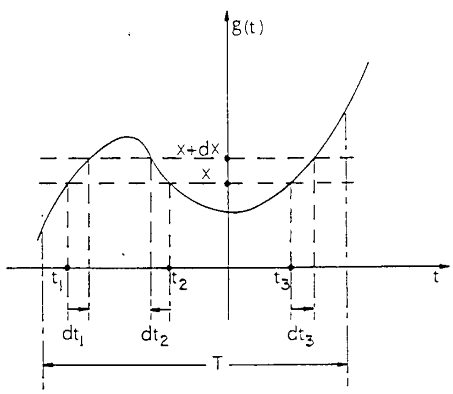

Fig. 10. The signal $x = g(t)$ in the duration T.

Denoting ωt by θ, then

$$\theta_{1,2} = \begin{cases} \sin^{-1}\left(\dfrac{x}{A}\right), & 0 \le \theta \le \pi \\ \pi + \sin^{-1}\left(\dfrac{-x}{A}\right), & \pi \le \theta \le 2\pi \end{cases} \quad \text{(A-8)}$$

$$\therefore \left|\frac{d\theta_i}{dx}\right|_{i=1,2} = \frac{1}{y\sqrt{1-(x/y)^2}} \text{ (for both } \theta \text{ ranges).} \quad \text{(A-9)}$$

Using (A-6)–(A-9) it follows that:

$$f_x(x) = \frac{1}{\pi\sqrt{A^2 - x^2}}, \qquad -A \le x \le A. \quad \text{(A-10)}$$

References

[1] B. Widrow, "A study of rough amplitude quantization by means of Nyquist sampling theory," *IRE Trans. Circuit Theory*, vol. CT-4, pp. 266–276, Dec. 1956.

[2] ——, "Statistical analysis of amplitude quantized sampled-data systems," *AIEE Trans. (Appl. Indust.)*, vol. 81, pp. 555–568, Jan. 1961.

[3] A. B. Sripad and D. L. Snyder, "A necessary and sufficient condition for quantization errors to be uniform and white," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-25, pp. 442–448, Oct. 1977.

[4] L. Schuchman, "Dither signals and their effects on quantization noise," *IEEE Trans. Commun. Technol.*, vol. COM-12, pp. 162–165, 1964.

[5] J. Vanderkooy and S. P. Lipshitz, "Resolution below the least significant bit in digital systems with dither," *J. Audio Eng. Soc.*, vol. 32, no. 3, pp. 106–113, Mar. 1984.

[6] Y. Yamasaki, "The application of large amplitude dither to the quantization of wide range audio signals," *J. Acoustic Soc. Japan*, vol. 39, no. 7, pp. 452–462, July 1983.

[7] G. Zames and N. A. Shneydor, "Dither in nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 660–667, Oct. 1976.

[8] I. D. Lotto and G. E. Paglia, "Dithering improves A/D converter linearity," *IEEE Trans. Instrum. Meas.*, vol. IM-35, pp. 170–177, June 1986.

[9] F. Oberhettinger, *Tabellen zur Fourier Transformation*. Berlin, Germany: Springer-Verlag, 1957.

[10] R. S. Burington, *Handbook of Mathematical Tables and Formulas*. 3rd ed., Sandusky, OH: Handbook Publishers, 1949.

[11] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. 2nd ed, New York: McGraw-Hill, 1984.