The Noise Model of Quantization*

István Kollár†

Abstract

Quantization error can often be modeled by an independent, uniformly distributed, additive white noise. In the paper the conditions of white quantization error spectrum are summarized, and a new, easy-to-check condition is formulated. An intuitive explanation of the condition is provided.
Keywords: Statistical theory of quantization, noise model, quantization error, spectrum of quantized signals.

1 Introduction

The so-called noise model of quantization is a very effective tool to describe the effect of uniform quantizers (to be more exact, the noise model can be used for a memoryless scalar quantizer, working in open-loop [1] – a usual and useful approximation for A/D conversion in electronic measurement systems [2]–[4]). The noise model means that the quantization error,

$$e(t) = x_q(t) - x(t) \ , \tag{1}$$

is assumed as being

- independent of (or at least uncorrelated with) the original signal, $x(t)$;
- uniformly distributed in $(-q/2, q/2)$;
- white

Validity conditions of the above noise model in terms of the characteristic function can be found in the literature [5], [6], [2], [3]. However, they are rather difficult to check in practical cases. This Gordian knot can be cut by adding an auxiliary signal (the so-called dither) to the original signal just before quantization. If this (independent) dither fulfills the above conditions, the noise model may be applied. Nevertheless this must be checked again; although the dither is in the hands of the designer of the instrument, there is a need for easy-to-check, visual conditions.

This paper deals with the conditions of the white spectrum. After having summarized the previous results, a new, simple condition is developed.

*Published in Proc. 1st IMEKO TC4 Symposium 'Noise in Electrical Measurements', Como (Italy), June 19-21, 1986, pp. 125-129. OMIKK-Technoinform, Budapest, 1987.

†Technical University of Budapest, Department of Measurement and Instrument Engineering
H-1521 Budapest, Műegyetem rkp. 9. Hungary. Phone: +36 1 463-1774, fax: +36 1 463-4112
email: kollar@mmt.bme.hu, URL: http://www.mmt.bme.hu/~kollar

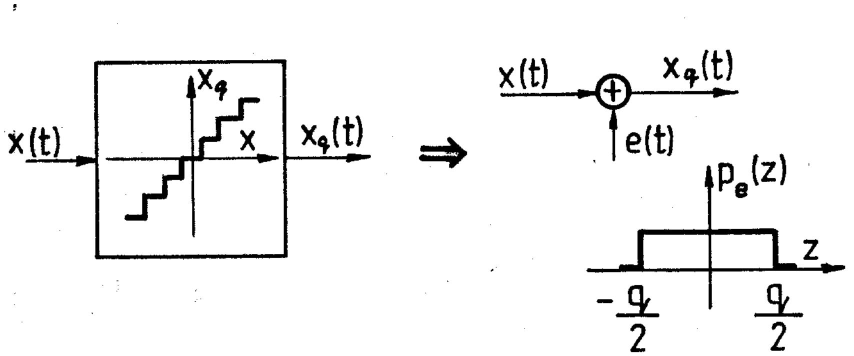

Figure 1: The noise model

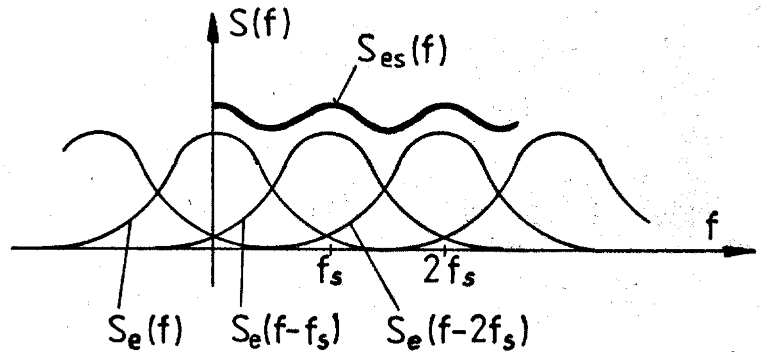

Figure 2: The effect of sampling in frequency domain

2 The Meaning of White Spectrum

First, it must be mentioned that the quantization error, *Eq.* (1), considered as a continuous-time stochastic process, cannot have white spectrum, since its variance is limited. However, in practical cases this spectrum is generally rather flat in a relatively broad band [7], [9], and if only this band is of interest, the spectrum may be considered as being white.

Moreover, quantization is usually performed in connection with sampling or on already sampled data. The effect of sampling is the repetition of the original spectrum. From *Fig. 2* it is obvious that in the case of smooth spectra and significant overlapping (sufficiently small f_s values) the resulting spectrum is more or less white. Thus, to provide white spectrum, an upper bound for f_s must be given.

Concerning the time domain, white spectrum means uncorrelated samples. This feature is rather easy to check, and will be useful in understanding the following results; on the other hand, this explains why whiteness of the spectrum is advantageous when the effect of quantization is investigated in measurement systems.

3 The Gaussian Case

In the literature some results are to be found concerning the features of the quantization noise. We are going to formulate conditions for the white spectrum, on the basis of these features.

a.) *Fig. 3* presents some of the results of Bennett [9], concerning the quantization error of band-limited white Gaussian noise. It can be observed that the spectra are rather smooth, and decreasing the quantum size, they broaden and flatten, as it is expected. According to *Fig. 2*, it seems to be reasonable to choose the sampling frequency not greater than the double of the 3dB (half-power) points, that is

$$f_s < 2 \cdot 150B \frac{\sigma_x}{q} \frac{(4+4)}{2^8} \approx 9.4 \frac{\sigma_x}{q} B \, , \tag{2}$$

because this provides significant overlapping.

b.) On the basis of computer evaluation of a series expansion of the quantization error spectra Robertson [8] gave a simple rule of thumb: provided that $\sigma_x > q$, the error spectra will be white even for colored signal spectra, if the sampling frequency is smaller than 2–3 times the Nyquist rate. For band-limited white noise, considering that in the case of $\sigma_x \gg q$ the limit is to be increased by the factor σ_x/q (see Section 5), the resulting condition,

$$f_s < 3 \cdot 2B \cdot \frac{\sigma_x}{q} = 6B \frac{\sigma_x}{q} \, , \tag{3}$$

is in good agreement with *Eq.* (2). It is self-evident that *Eq.* (3) is stricter, since it is derived from a condition valid for colored spectra as well.

Robertson's results can be used in the case of narrow-band Gaussian noise, too (see *Fig. 5*):

$$f_s < 3 \cdot 2 f_0 \frac{\sigma_x}{q} = 6 f_0 \frac{\sigma_x}{q} \, . \tag{4}$$

c.) Katzenelson [10] derived an approximate expression of the correlation coefficient of quantization error samples. If the correlation between the signal's samples is tight ($\rho_x \approx 1$),

$$|\rho_e| \approx e^{-4\pi^2 (1-|\rho_x|) \frac{\sigma_x^2}{q^2}} \, . \tag{5}$$

from the result of Katzenelson, supposing $\sigma_x = q$, it follows that the inequality

$$|\rho_e| < 0.1 \quad \text{if} \quad |\rho_x| < 0.94 \tag{6}$$

is valid. Using the expression of the correlation coefficient of the band-limited white noise:

$$\rho_x(\tau) = \frac{\sin(2\pi B \tau)}{2\pi B \tau} \, , \tag{7}$$

$$|\rho_x| < 0.94 \quad \text{if} \quad |\tau| > \frac{0.2}{2B} \, . \tag{8}$$

From *Eqs.* (6) and (8) it follows that uncorrelatedness ($|\rho_e| < 0.1$) of the error samples of a band-limited white Gaussian noise is provided if

$$f_s < \frac{2B}{0.2} \, . \tag{9}$$

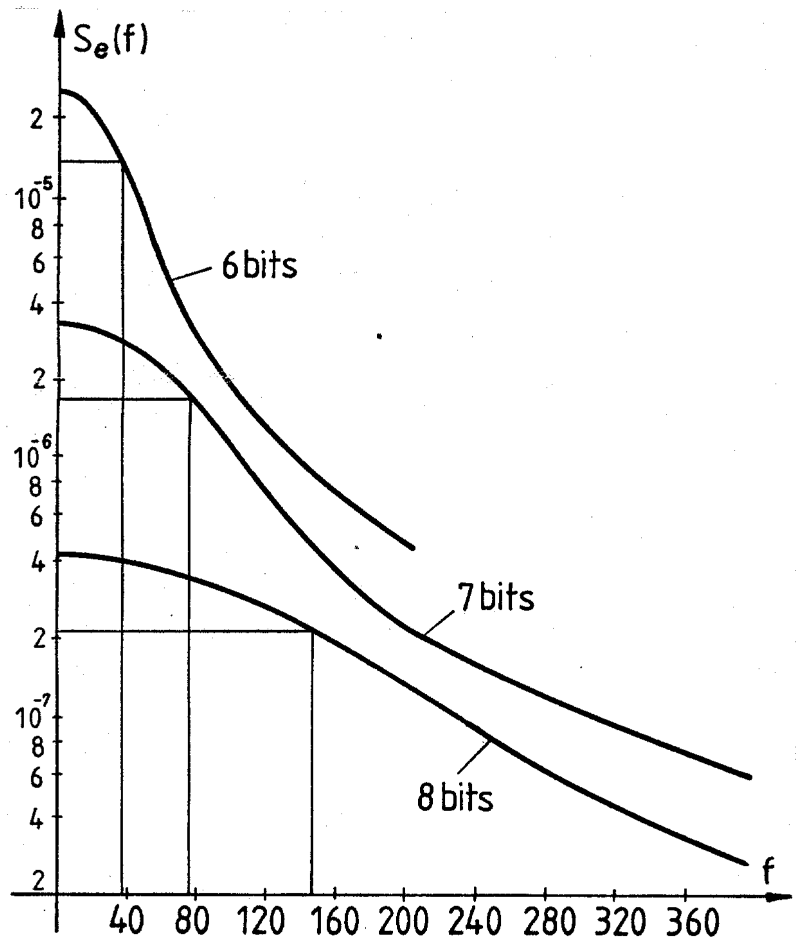

Figure 3: Quantization error spectra of band-limited Gaussian noise
Frequency unit: Bandwidth B of the original signal
Power unit: mean signal power (σ_x^2)
The interval $(-4\sigma_x, 4\sigma_x)$ is equal to the input range of the A/D converter

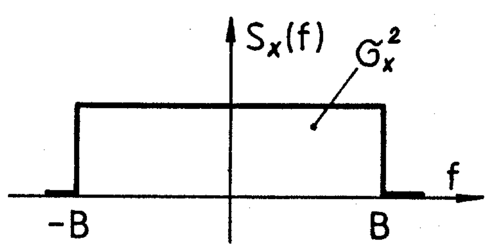

Figure 4: Spectrum of band-limited white noise

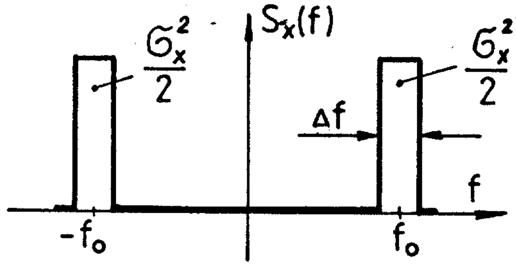

Figure 5: Spectrum of narrow-band noise: $\Delta f \ll f_0$

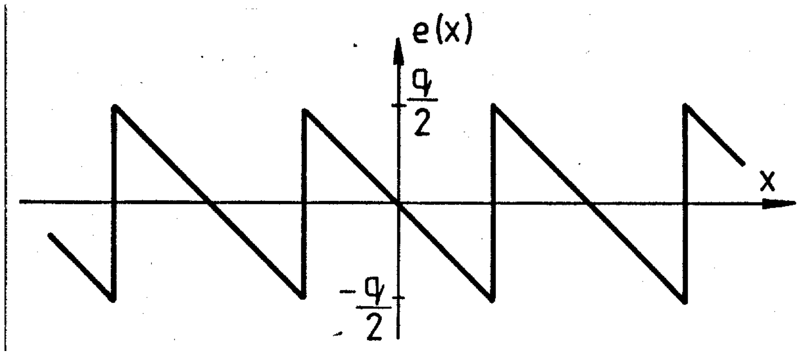

Figure 6: The quantization error as a function of the signal amplitude

A more detailed investigation, based on the Taylor series expansion of *Eq.* (7), shows that in the general case of $\sigma_x \neq q$ the condition

$$f_s < \frac{2B}{0.2}\frac{\sigma_x}{q} = 10B\frac{\sigma_x}{q} \tag{10}$$

ensures uncorrelatedness. *Eq.* (10) agrees well with (2).

4 Quantization Error of a Sine Wave

The case of a sine wave was investigated by CLAASEN and JONGEPIER [7]. Their starting point was that the quantization error as a function of the signal value x (*Fig. 6*) can be developed into a Fourier series:

$$e(x) = q\sum_{k=1}^{\infty}(-1)^k\frac{\sin(2\pi kx/q)}{k}\ . \tag{11}$$

If $x(t)$ is a sine wave:

$$x(t) = X\sin(2\pi f_1 t + \varphi)\ ,$$

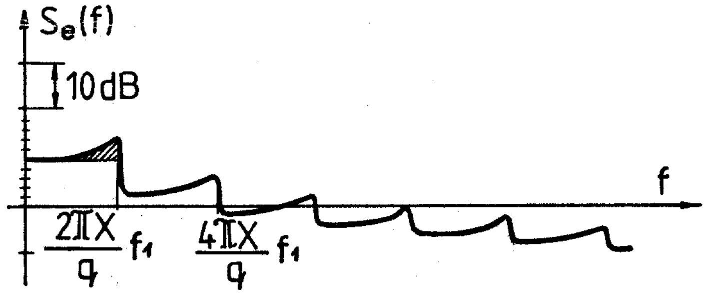

Figure 7: Quantization error spectrum of a sine wave

Eq. (11) is a sum of phase modulated signals, and the spectrum is approximately [7]:

$$S_e(f) \approx \frac{q^3}{2\pi^2} \sum_{k=1}^{\infty} \cdot \frac{f_{\dot{x}}\left(\frac{fq}{k}\right)}{k^3}\,, \tag{12}$$

where

$$f_{\dot{x}}(z) = \frac{1}{2\pi^2 f_1 x} \frac{1}{\sqrt{1 - \left(\frac{z}{2\pi f_1 X}\right)^2}}\,, \quad |z| < 2\pi f_1 X \tag{13}$$

is the probability density function of the first derivative of $x(t)$.

Equation (12) is clearly an approximation, since the quantization error is surely periodic with fundamental frequency f_1, and as such it has discrete spectrum. However, (12) was successfully verified by measurements with a spectrum analyzer. The character of a measured spectrum is shown in *Fig. 7* [7]. The infinite peaks of (13) are smoothed by the finite resolution of the analyzer, given in [7] as

$$\Delta f \approx \frac{1}{20} \frac{2\pi X}{q} f_1\,.$$

In the case of *Fig. 7* the amplitude X is equal to $15.5q$, that is, $\Delta f \approx 4.9 f_1$, and this explains why the discrete nature of the spectrum did not prevent verification.
From *Fig. 7* it is clear that $S_e(f)$ is flat for

$$|f| < \frac{\pi X}{q} f_1\,. \tag{14}$$

JONGEPIER and CLAASEN suggest this or an even smaller upper bound for the sampling frequency. Considering that the first peak contains ca. 10% of the total signal power, e.g. the condition

$$f_s < \frac{1}{2} \frac{\pi X}{q} f_1 \tag{15}$$

may be suggested. The resulting spectrum will contain some ripples for any value of f_s, however, the power associated to them will be small in comparison to the total power of the quantization error (ca. 1% to each peak in the case of *Eq.* (15).

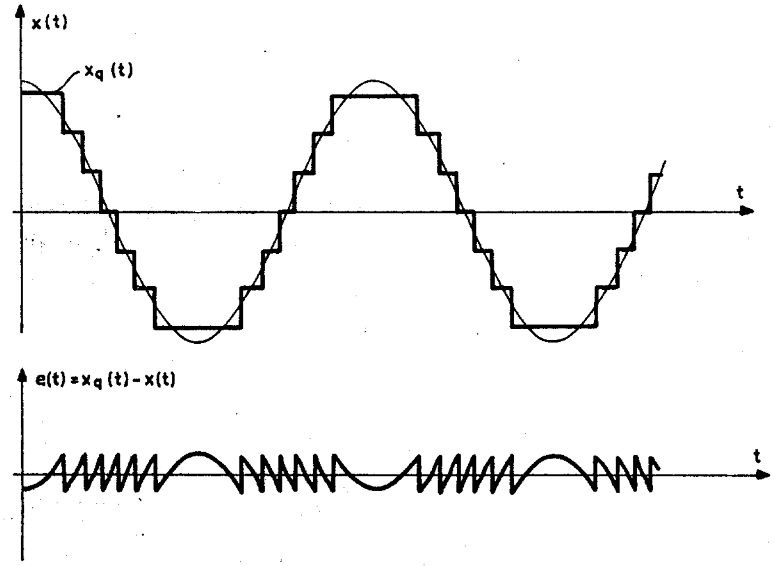

Figure 8: Quantization of a sine wave

5 A Uniform Condition

In the previous sections some completely different techniques were used to derive upper bounds for the sampling frequency. In this section an attempt is made to formulate a uniform condition, independent of concrete signal parameters [11].

In *Fig. 8* the quantization of a sine wave is illustrated. Let us notice that the waveform of the quantization error is in most time intervals very similar to a 'saw-tooth' signal. It seems to be quite natural to assume that the samples of the quantization error are uncorrelated, if not more than 1–2 samples are taken from a 'period'. The average length of these periods depends on the average slope of the signal, and can be expressed as follows:

$$T_p = \frac{q}{E\{|\dot{x}(t)|\}}\,, \tag{16}$$

where $\dot{x}(t)$ is the first derivative of the signal, and $x(t)$ is assumed to be stationary. Equation (16) explains why the factor σ_x/q was introduced e.g. in *Eq.* (3). From *Eq.* (16) an upper bound can be given for the sampling frequency

$$f_s < K\frac{E\{|\dot{x}(t)|\}}{q}\,, \tag{17}$$

where the constant K may be equal to 1..2.

Let us check this condition in the cases treated in Sections 3 and 4. For zero mean Gaussian variables

$$E\{|x|\} = \sqrt{\frac{2}{\pi}}\sigma_x \tag{18}$$

is valid. Using

$$\sigma_{\dot{x}}^2 = \int_{-\infty}^{\infty} |j2\pi f|^2 S_x(f)dt\,, \tag{19}$$

in the case of band-limited white noise

$$E\{|\dot{x}(t)|\} = \sqrt{\frac{2}{\pi}}\sigma_{\dot{x}} = \sqrt{\frac{8\pi}{3}}B\sigma_x \, , \tag{20}$$

and for narrow-band white noise

$$E\{|\dot{x}(t)|\} = \sqrt{\frac{2}{\pi}}\sigma_{\dot{x}} = \sqrt{8\pi}\sigma_x f_0 \tag{21}$$

are obtained.

For the sine wave

$$E\{|\dot{x}(t)|\} = 4f_1 X \, . \tag{22}$$

Comparing *Eqs.* (20)–(22) with (2), (4), and (15), it is obvious that in each case (17) may be used, with $K = 3.2, 1.2, 0.4$, respectively. One can observe that K is in the same order of magnitude for each type of signal (though its value varies more than it was supposed at the beginning of Section 5). It is not surprising that the 'worse' the signal behaves ('bad' means in our interpretation strongly deterministic, or being approximately constant in some time intervals), the smaller the value of K must be chosen.

On the basis of the above results (17) seems to be very useful, especially, because usually $E\{|\dot{x}(t)|\}$ can directly be measured. The appropriate value of K depends slightly on the waveform, nevertheless since (17) is an inequality, K may be chosen to be sufficiently small in order to have a safe upper bound.

6 Conclusion

It is shown in the paper that (17) may be used for two types of Gaussian signals and for sinusoidal signals, too, to provide white quantization error spectra. There are further examples [12], in which quantization errors behave similarly. However, it is still an open question, how the value of K must be chosen for a class of signals. Intuitively it is clear that if a signal does not have broad nearly-constant sections or nearly-constant-slope sections (the square wave and the triangle wave are counterexamples) (17) may be used e.g. with $K = 0.4$, or with greater values of K if the signal is of stochastic nature. Further investigations may provide an exact mathematical formulation of this restriction.

References

[1] Dobrowiecki, T. Modeling a Quantizer – Models and Possible Approaches. *Periodica Polytechnica Ser. Electrical Engineering*, Vol. 28, Nos. 2–3, pp. 159–172, 1984.

[2] Kollár, I. Statistical Theory of Quantization: Results and Limits. *Periodica Polytechnica Ser. Electrical Engineering*, Vol. 28, Nos. 2–3, pp. 173–190, 1984.

[3] Kollár, I. Errors Introduced by Digital Power/Energy Measurements. *Xth IMEKO World Congress*, May 22-26, 1985, Prague; *Acta IMEKO* 1985, Vol. 5, pp. 343–349.

[4] Kollár, I. Dynamic Range of Digital Spectrum Analyzers. *5th IMEKO Symposium, 'Intelligent Measurement'*, Jena, GDR, June 10–13, 1986. pp. 123-128.

[5] Widrow, B. A Study of Rough Amplitude Quantization by Means of Nyquist Sampling Theory. ScD. Dissertation, M.I.T. Elec. Eng. Dept., Electronics Systems Lab., Cambridge, Mass., 1956.

[6] Sripad, A. B. – Snyder, D. L. A Necessary and Sufficient Condition for Quantization Errors to Be Uniform and White. *IEEE Trans. on Acoustics, Speech and Signal Processing*, Vol. ASSP-25, No. 5, pp. 442–448, 1977.

[7] Claasen, T. A. C. M. and Jongepier, A. Model for the Power Spectral Density of Quantization Noise. *IEEE Trans. on Acoustics, Speech and Signal Processing*, Vol. ASSP–29, No. 4, pp. 914–917, 1981.

[8] Robertson, G. H. Computer Study of Quantizer Output Spectra. *Bell System Technical Journal*, Vol. 48, No. 5, pp. 2391–2403, 1969.

[9] Bennett, W. R. Spectra of Quantized Signals. *Bell System Technical Journal*, Vol. 27, No. 4, pp. 446–472, 1948.

[10] Katzenelson, J. On Errors Introduced by Combined Sampling and Quantization. *IRE Trans. on Automatic Control*, Vol. AC–7, pp. 58–68, 1962.

[11] Kollár, I. Measurement Theoretical Investigation of Fourier Analyzers. Cand. Sci. Dissertation, Hungarian Academy of Sciences, Budapest, 1984. (In Hungarian)

[12] Kollár, I. Steady-State Value Measurements with Quantized Data. *Proc. 2nd IMEKO TC8 Symposium, 'Theoretical and Practical Limits of Measurement Accuracy'*, Budapest, May 10–12 1983, pp. 92–101. IMEKO Secretariat, 1982.