THE DESCRIPTION OF THE SPECTRUM CHARACTERISTICS OF THE QUANTIZATION ERROR

Anna Domańska,
Institutue of Electronics and Telecommunications, Poznań University of Technology,
domanska@et.put.poznan.pl

Abstract – The present paper examines a deterministic approach to quantization in a-d conversion processes. Temporal and spectral representations of the quantization error are discussed, referring to conditions where the characteristics of the quantizer are ideal and a sinusoidal signal is converted. A formula is provided for estimating the bandwidth occupied by the quantization error. The spectrum model of the quantization error is discussed as a tool of the description and testing of A/D converters.

Keywords – quantization theory, quantization error spectrum, angle modulation.

1. INTRODUCTION

The purpose of the modelling of a-d conversion is to provide a formal description of the sequence of constituent operations. Quantization is a non-linear operation, and therefore is the most difficult to model.

Two approaches to the modelling and describing of quantization are possible: the statistical and the deterministic approach. The latter approach is applied in studies of the quantization of deterministic signals. In such cases, the descriptions of the signal after quantization and of the quantization error provide insights into their "structure" in the domains of time, frequency and energy.

The deterministic approach may be based on one of the two propositions, which differs by the ways in which they define the modelling of the static characteristics of the quantizer:

- the characteristics are modelled by the sum of the linear and the sawtooth functions [1, 2],
- the characteristics are modelled by the sum of the step functions [3,4].

In both cases, a relationship between the signals after and before quantization may be identified in the time domain as well as in the frequency domain.

The issue of the characterisation of the quantization error in the frequency domain is hereafter discussed based on the deterministic approach and the first proposition. Characterisation in the frequency domain refers to the quantization error of a sinusoidal signal. A/D converters are usually tested or described by applying sinusoidal signals to their inputs.

2. THE QUANTIZATION ERROR IN THE TIME DOMAIN

The following discussion pertains to an a-d conversion system with Nyquist rate A/D converters.

The quantizer is uniform and has *midtread* static characteristics. The product of a quantization performed by a quantizer having the characteristics $Q(x)$ may be equivalently presented as the sums of the products of conversions performed by two systems connected in parallel: one of these has linear characteristics $l(x)$ with a positive slope, and the other, sawtooth characteristics $g(x)$ with a negative of the slopes, featuring points of discontinuity; $Q(x) = l(x) + g(x)$.

The error of the operation of quantization is modelled by the sawtooth function, and may be presented as a Fourier series. As the function is periodic and odd, and its mean value amounts to 0, the specific form of the series may be expressed as follows:

$$e = g(x) = \frac{q}{\pi}\sum_{k=1}^{\infty}\frac{(-1)^k}{k}\sin\left(\frac{2\pi}{q}k \cdot x\right), \qquad (1)$$

where x is the input quantity of the system (the quantity before quantization) and q is the resolution of the quantizer. The quantization error amounts to the product of the conversion of the signal $x(t)$ by a system with sawtooth characteristics, i.e. $e(t)=g[x(t)]$. We emphasise that the error of the operation of quantization is a characteristic of the system, while the quantization error is a signal.

The quantization error is an infinite sum of angle modulated signals ($k \to \infty$); when a sinusoidal signal $x(t)=A\sin\omega_x t$ is converted, the quantization error is expressed by the dependence:

$$e(t)=\frac{q}{\pi}\sum_{k=1}^{\infty}\frac{(-1)^k}{k}\sin\left(\frac{2\pi}{q}k\cdot A\sin\omega_x t\right). \quad (2)$$

By applying the Jacobi equation, the above dependence is transformed into a function which is directly dependent on the time parameter

$$e(t)=\frac{2q}{\pi}\left\{\sum_{k=1}^{\infty}\frac{(-1)^k}{k}\sum_{n=1}^{\infty}J_{2n-1}(b_k)\sin(2n-1)\omega_x t\right\} \quad (3)$$

where $b_k=(2\pi/q)A\cdot k$ and J_{2n-1} is the Bessel function of the first kind of order *2n-1*.

The quantization error is the sum of odd-indexed harmonics of the signal *x(t)*. The amplitude of the harmonic component *2n-1* is given by $\frac{2q}{\pi}\sum_{k=1}^{\infty}\frac{(-1)^k}{k}J_{2n-1}(b_k)$.

3. THE QUANTIZATION ERROR IN THE FREQUENCY DOMAIN

The spectrum of the quantization error may be described based on Eq. (3) as the formula:

$$\mathcal{F}\{e(t)\}=j2q\left\{\sum_{k=1}^{\infty}\frac{(-1)^k}{k}\sum_{n=1}^{\infty}J_{2n-1}(b_k)\cdot[\delta(\omega+(2n-1)\omega_x)-\delta(\omega-(2n-1)\omega_x)]\right\} \quad (4)$$

❶ The spectrum (4) is the sum of an infinite set of elementary spectra ($k \to \infty$), representing an infinite set of angle modulated signals. Elementary spectra have the following characteristics:

- each elementary spectrum is unlimited ($n \to \infty$),
- in each elementary spectrum, the spectral lines appear at the same locations: $\pm(2n-1)\omega_x$,
- in each elementary spectrum, the spacing between spectral lines is constant, amounting to $2\omega_x$,
- the magnitude of the line *(2n – 1)* in the *k*-th elementary spectrum amounts to $2q\frac{(-1)^k}{k}J_{2n-1}(b_k)$. The magnitude of each spectral line can be determined using a table of the values of the Bessel function.

❷ The spectrum (4) consists of spectral lines appearing at the locations of odd-numbered multiples of the angular frequency ω_x of the sinusoidal signal before conversion. The spacing between consecutive lines is constant, amounting to twice the value of the angular frequency. The magnitude of the line *(2n – 1)* amounts to $2q\sum_{k=1}^{\infty}\frac{(-1)^k}{k}J_{2n-1}(b_k)$, or it is the sum of the magnitudes of the lines *(2n – 1)* in all the *k* elementary spectra.

❸ The bandwidth of the error signal is unlimited ($k \to \infty, n \to \infty$). Only an estimate value of the bandwidth can be expressed, assuming limitations of the number of the values of *k* and *n*. The number of *k* may be limited to $k_{\max} < \infty$ through assuming the accuracy of the Fourier series representation of the function *g*(*x*). The number of *n* may be limited to $n_{\max} < \infty$ by applying the rules of the calculation of the effective bandwidth of the angle-modulated signal. The effective bandwidth of the error signal is determined by the spacing between spectral lines and the assumed number of lines.

4. THE BANDWIDTH OF THE QUANTIZATION ERROR

4.1. The classic rules of the calculation of the effective bandwidth of the angle-modulated signal

Since the signal after angle modulation consists of a theoretically infinite number of spectral lines, therefore its bandwidth is also infinite. In practice, however, the number of spectral lines of significant magnitudes (according to the applied criterion) is finite. *Carson's rule* or *the universal curve method* (*Manaev's rule*) [5, 6] may be applied in order to calculate the effective bandwidth.

Carson's rule is used to establish the bandwidth B_C carrying 98% of the energy of the angle-modulated signal. If the modulating signal is a non-periodical deterministic signal $m(t)$, then the effective bandwidth is expressed by the formula:

$$B_C = 2(D+1)W, \quad (5)$$

where W is the bandwidth of the modulating signal and D is the deviation ratio:

$D_{PM} = k_p \max|m(t)|$, $D_{FM} = (k_f / W)\max|m(t)|$.

If the modulating signal is a periodical signal $x(t) = A\sin\omega_x t$ then the effective bandwidth is expressed by the formula:

$$B_C = 2(b+1) f_x, \quad (6a)$$

where $f_x = \omega/(2\pi)$ and b is the modulation index: $b_{PM} = k_p A$, $b_{FM} = k_f A / f_x$.

In such a case, the number of spectral lines of significant magnitudes, or the lines carrying 98% of the total energy of the angle-modulated signal, amounts to:

$$2n^C_{\max} = 2 \cdot ent(b+1). \quad (6b)$$

The universal curve method (*Manaev's rule*) makes it possible to establish the bandwidth B_M consisting of spectral lines whose magnitudes are not smaller than 1% of the magnitude of the line representing the non-modulated carrier signal. If the modulating signal is a periodical signal $x(t) = A\sin\omega_x t$ then the effective bandwidth is expressed by the formula:

$$B_M = 2\left(\sqrt{b} + b + 1\right) f_x, \quad (7a)$$

The number of spectral lines of significant magnitudes (according to the criterion of *Manaev's rule*) amounts to:

$$2n^M_{\max} = 2 \cdot ent\left(\sqrt{b} + b + 1\right). \quad (7b)$$

By comparing the formulae (6a) and (7a), we infer that the relationship between the two methods of calculation is $B_C < B_M$, or $n^C_{\max} < n^M_{\max}$.

The practical value of the two rules as methods of the calculation of the effective bandwidth of the quantization error may be summarized as follows:

- both methods may be applied to calculate the effective bandwidth of a single angle-modulated signal; in fact, however, the quantization error (2) is the sum of all such signals, and its spectrum (4) is the sum of the elementary spectra of angle-modulated signals, the k-th elementary spectrum having the form:

$$F\ \{e(t)\}\Big|_k = j2q\frac{(-1)^k}{k}\sum_{n=1}^{\infty} J_{2n-1}(b_k)\cdot \\ \cdot\left[\delta(\omega + (2n-1)\omega_x) - \delta(\omega - (2n-1)\omega_x)\right], \quad (8)$$

- the spectrum of the quantization error concentrates around the frequency f = 0, and therefore is organized like the spectrum of an angle-modulated signal without a carrier signal; since there is no spectral line representing the non-modulated carrier signal, *Manaev's rule* cannot be applied,
- the carrier frequency affects the location of the spectrum but not its bandwidth, since both the modulation index and the spacing between the lines are independent of it; *Carson's rule*, which bases on the criterion of energy, may be applied for the calculation of the effective bandwidths of all the elementary spectra.

4.2. The effective bandwidth of the quantization error

The effective bandwidth of the quantization error may be calculated upon assuming limitations of the infinite sets of the values of k and n.

The number of k may be limited to $k_{\max} < \infty$ through assuming the accuracy of the Fourier series representation of the function $e = g(x)$ (1). Thereupon, the spectrum of the quantization error (4) is defined as the sum of a finite set of elementary spectra.

By comparing the effective bandwidths of all the elementary spectra (calculated by means of *Carson's rule*), we notice that, as we have expected, the last significant spectrum is the widest. This is caused by the following relationship:

$$\begin{aligned} &1 < 2 < \ldots < k_{\max} \\ &b_1 < b_2 < \ldots < b_{\max} = (2\pi/q)A \cdot k_{\max} \\ &B_{C1} \subset B_{C2} \subset \ldots \subset B_{C\max} = 2(b_{\max}+1) f_x = \\ &\qquad = 2\cdot\left[(2\pi/q)A\cdot k_{\max} + 1\right] f_x \end{aligned}, \quad (9)$$

where B_{C_i} is the the effective bandwidth of the i-th elementary spectrum. Based on Equation (9) and the structure of the spectrum of the quantization error and of the elementary spectra, which has been discussed above, we infer that the bandwidth of the spectrum of the quantization error is not wider than that of the last significant elementary spectrum, or $B_C \approx B_{C\max}$.

Having taken all of the above findings into account, we describe characteristic ❸ (Sec. 3) of the spectrum of the quantization error as follows:

- the effective bandwidth of the quantization error is expressed by the formula: $B_C = 2(b_{\max} + 1) f_x$,
- the number of spectral lines of significant magnitudes amounts to $2n_{\max}^C = 2 \cdot ent(b_{\max} + 1)$,
- the number of the values of n (the upper limit of the second sum in (4)) is limited to $n_{\max}^C = ent(b_{\max} + 1)$.

5. THE SPECTRUM OF THE SIGNAL AFTER A-D CONVERSION

Given the model of the spectrum of the quantization error, the spectrum of the signal after quantization and the spectrum of the signal after a-d conversion (sampling and quantization) may be defined.

Since the relationship between signals after and before quantization in the time domain is the sum of $x_Q(t) = x(t) + e(t)$, the spectrum of the signal after quantization is expressed as:

$$X_Q(\omega) = X(\omega) + \mathcal{F}\{e(t)\}, \qquad (10)$$

where $X(\omega)$ is the spectrum of the signal before quantization
$X(\omega) = j\pi A[\delta(\omega + \omega_x) - \delta(\omega - \omega_x)]$ and
$\mathcal{F}\{e(t)\}$ is the spectrum of the quantization error (4).

The spectrum of the signal after a-d conversion is expressed as:

$$X_{SQ}(\omega) = \frac{1}{T_S}[X(\omega) + \mathcal{F}\{e(t)\}] * \delta_{\omega_S}(\omega), \qquad (11)$$

where T_S is the interval of sampling, $T_S = (2\pi)/\omega_S$ and $\delta_{\omega_S}(\omega)$ is the Dirac comb.

6. CONCLUSIONS

The quantization error of a sinusoidal signal is an angle-modulation signal.

The spectrum of the quantization error is organized as follows: it is a discrete spectrum; its composition is dependent on the nature of the quantizer, and for each type of quantizer (midtread, truncation or midrise) the quantization error is modeled as a different sawtooth function; the location of and spacing between the spectral lines are determined by the frequency of the sinusoidal signal before conversion, while the magnitudes of the lines are determined by the amplitude of the signal and the resolution of the A/D converter.

The effective bandwidth of the quantization error may be calculated by applying Carson's rule, upon assuming the accuracy of the Fourier series representation of the error of the operation of quantization.

The created model of the quantization error makes it possible to describe the spectrum characteristics of the results of the ideal quantization and the ideal a-d conversion. The model of the ideal quantization (conversion) in each case provides the reference for the assessment of the actual quantization (conversion). The assessment consists in comparing the result of a DTF test of the signal after conversion in the actual system with the result provided by the model.

References

[1] Blachman N., M., "The intermodulation and distortion due to quantization of sinusoids", IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. ASSP-33, pp. 1417-1426, December, 1985.

[2] Abuelma'atti M., T., "The intermodulation due to multicarrier quantization", IEEE Transactions on Communications, vol., COM-32, pp. 1211-1214, November, 1984.

[3] Bellan D., Brandolini A., Gandelli A., "Quantization theory-a deterministic approach", IEEE Transactions on Instrumentation and Measurement, vol. 48, pp. 18-25, February 1999.

[4] Bellan D., Brandolini A., Rienzo L., Gandelli A., "Deterministic quantization theory applied to ADC testing", Measurement, vol. 19, pp. 169-177, no. ¾, 1996.

[5] Haykin S., Communication Systems, New York, Wiley, 1994, 3.11.

[6] Proakis J., Salehi M., Communication Systems Engineering, NJ, Prentice Hall, 2002, 3.3.2, p. 104-107.