ON THE QUANTIZATION OF NOISY SIGNALS IN BINARY REGENERATION SYSTEMS

M.A. MATAR* , G. MARAL**

ABSTRACT

This paper deals with the problem of quantizating (A/Dconverting) binary signals contaminated by the background gaussian noise. It answers the question of determination of the uniform binary quantizer such that the statistics of the input signal are uniquely determined by those of its quantized version. Moreover, we study the properties of the quantization noise: mean, variance, autocorrelation function, power spectrum and its crosscorrelation with the input signal. Computer simulation results are given and analyzed.

* Avionics Chair, M.T.C., Cairo, EGYPT.

** Professor, Dept. of Electronics and Comm., ENST, ENSAE and UPS, Toulouse, France.

INTRODUCTION

The never-preceded advance in digital integrated circuits (VLSI and VHSIC) has its impact on the digital signal processing technique and the ever-increasing tendency towards digital implementation of systems. This is favoured especially in AVIONICS where high reliability and integrability of systems is an ever-forcing requirement against the hard environment where, sometimes, no human correcting action against failure or drift is feasible. To be able to profit of advantages of digital processing capabilities, the signals should be A/D converted and it is the objective of this paper to study this problem in binary regeneration systems appearing in communication, radar, data links, electronic warfare application and/or even when a digital storage or replica of the signal is required (for instance for non-real time processing).

Systems dealing with the regeneration of random binary signals burried in a background gaussian noise generally involve a statistical treatment which induces the structure of the system functional blocks (PLL, bit detectors, lock indicators, acquisition aids... etc) commonly based on a matched filtering (or correlation) operation. In such systems, the informations to be reconstructed are some of the statistics of the signal (generally up to the second order statistics as the input signal and noise are second order processes) rather than the signal wave-form itself. This suggests that whenever the signal is to be transformed in the regenerator, the statistical aspects of signal should remain recognizable. This is especially difficult when nonlinear transformations are involved.

Being a memoryless nonlinearity, the quantizer can be studied, at least in theory, using the characteristic function (or the transform) method [12] or the time domain generalized Fourier series expansion method [13]. Due to the complexity of these methods and the necessity of numerical determination of the required statistical quantities, this approach is not retained. Instead, short-cut answers based of the statistical quantization theory will be offered. These answers, even they are not complete, are quite useful. Whenever these answers are not sufficient, computer simulation is adopted and briefly outlined.

In section II, consideration of the problem in lights of the quantization theory for the infinite quantizer case is given. This is folowed, in section III, by the determination of alimited-number of levels quantizer derived under stated constraints. In section IV, computer simulation of the obtained quantizer is given and results are analyzed. This is followed by the main conclusions.

II. CONSIDERATIONS IN LIGHTS OF THE QUANTIZATION THEORY (INFINITE QUANTIZER CASE).

Let us consider Fig.1. showing schematically the situation in concern. A transmitted binary signal $S(t)$ is received after being contaminated by the noise $n(t)$ added by the transmission channel to form the signal $Z(t)$ which is then transformed into the quantized version $Z_Q(t)$ by a symmetric uniform step size (Q) memoryless quantizer with a finite (2N+1) number of levels. Formally $Z(t) = S(t) + n(t)$ (1.a)

The binary signal $S(t)$ is a random binary stream of anticorrelated equiprobable statistically independent symbols occuring at the rate $F = \frac{1}{T}$ (T is the symbol duration) and of amplitude A and symbol energy A^2T. The noise $n(t)$ is a zero mean stationary gaussian noise with constant spectral density No inside a given band (O, F_c), hence its variance σ^2 is N_oF_c. The signal $S(t)$ and the noise $n(t)$ are second order processes. As shown in Fig. 1, the quantized signal $Z_Q(t)$ is modelled as : $Z_Q(t) = Z(t) + n_Q(t)$

$$Z_Q(t) = Z(t) + n_Q(t) \tag{1.b}$$

Before analyzing the situation at hand, let us recall that the quantization theorem [1-4] may be stated as :

"The moments (if exist) $E[Z^m]$ are completely determined by $E[Z_Q^m]$ and the quantization noise is uniformly distributed over $(-\frac{Q}{2}, \frac{Q}{2})$ if the characteristic function is horizon-limited. A similar statement is valid for the joint moments if the joint characteristic function is horizon-limited. In this latter case, the quantization noise samples are uniformly distributed and statistically independent". For other details see [14].

<u>The noise-free case</u>: $n(t) = 0$

$$Z(t) = S(t) \Rightarrow p(S) = \frac{1}{2} \cdot [\delta(S-A) + \delta(S+A)] \tag{2}$$

where $\delta(.)$ is the delta-Dirac function

Then $\emptyset(\lambda) = \cos(\lambda A)$; $\emptyset(.)$ being the characteristic function

This function is not horizon limited but really a wide-ranging one.

Thus S(t) does not satisfy the quantization theorem. Neverthless, S(t) can be handled without distortion with the quantizer shown in Fig. 1 as far as its amplitude is larger than the value of the first two symmetrical transitions , i.e.

$$A \geqslant Q/2 \tag{3}$$

In such conditions, the quantizer functions as an amplifier with variable gain and the quantization noise is merely a replica of S(t) whose amplitude ranges from 0 to $Q-\varepsilon$, $\varepsilon > 0$, and thus is completely correlated with the signal.

Hence, one can consider that n_Q is zero without loss of generality. Concluding, the condition (3) must be imposed even in the presence of noise otherwise the quantizer suppresses the signal. Narrowband filtered signal will be discussed later on.

The noisy signal case

$$Z(t) = S(t) + n(t)$$

with

$$p_n(n) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp(-n^2/2\sigma^2)$$

Hence $\quad \varnothing_z(\lambda) = \varnothing_s(\lambda) . \exp(-\sigma^2 \lambda^2/2)$

$$\leq \exp(-\sigma^2 \lambda^2/2) \tag{4.a}$$

Similarly:

$$\varnothing_z(\lambda_1, \lambda_2) \quad \exp\left[-\frac{\sigma^2}{2}(\lambda_1^2 + 2\rho(\tau)\lambda_1\lambda_2 + \lambda_2^2)\right] \tag{4.b}$$

where $\rho(\tau)$ is the correlation coefficient of the gaussian noise.

It is thus clear that Z(t) has a characteristic function that is always less than that of the noise process n(t).

Thus, inequalities (4) lead to an interesting result : one may consider only the gaussian noise component when one adresses the problem of satisfaction of the conditions of the quantization theorem by noisy signals like Z(t). This facilitates the task as the gaussian noise is known to satisfy, with good approximations, the conditions of the quantization theorem [1-4] as far as $Q \leq \sigma$. Consequently, the process Z(t) satisfies the quantization theorem with good approximations and the condition:

$$Q \leq \sigma \tag{5}$$

will be retained for quantizing Z(t). Hence, the following results predicted by the above-mentioned theorem are valid for Z(t) :

$$E[Z_Q] = E[Z]$$

$$E[Z_Q^2] = E[Z^2] + Q^2/12$$

$$= A^2 + Q^2/12 + \sigma^2$$

$$R_{Z_Q}(\tau) = \begin{cases} R_Z(\tau) & \tau \neq 0 \\ E[Z^2] + Q^2/12 & \tau = 0 \end{cases}$$

III. DETERMINATION OF A LIMITED M-BIT QUANTIZER

In practice, we dispose quantizers with finite, and generally small, number of output levels. With such quantizers, a clipping effect is observed for too large input signal amplitudes. The conditions will now be given under which this practical quantizer approximates conveniently the theoretical quantizer considered before. It will be verified by simulation that these conditions satisfy this approximation.

3.1. Imposed constraints (criteria)

Let A_m be the upper limit in the positive dynamic range of the quantizer. The relation of A_m to the signal and noise standard deviation is known as the loading rule. We search such a loading rule for an M-bit quantizer under the following constraints :

1) The probability of clipping is less than 1%. This should reduce the contribution of the clipping noise to a level such that all happens as if a theoretical quantizer were used.
2) The quantization step Q is related to the noise standard deviation by :

   $$Q \leq \sigma$$

   as claimed for by the quantization theorem (condition (5)).
3) No signal suppression : $A \geq Q/2$ (condition (3)).
4) Conservation of the input signal-to-noise ratio with maximum loss of 0.5 dB in the range of (E/N) from 0 dB up to 15 dB.

3.2. Parameters of the quantizer

As the signal plus noise process is not gaussian but tends to be for low signal-to-noise ratio (SNR), we examine firstly the low SNR or the gaussian case and subsequently the signal plus noise case.

<u>The low SNR (Nearly Gaussian) Case</u>

i- Constraint (1) implies that $A_m \simeq 2.6\,\sigma$

ii- Constraint (2) implies, with $A_m = 2.6\,\sigma$, that: (Since $Q = 2A_m/(2^M - 1)$)

$$6.2 \lesssim 2^M \quad \text{or} \quad M \gtrsim 2.63 \quad \text{i.e.} \quad M \gtrsim 3$$

since only integer values for M are considered

The question is : what is the optimum (in the mean square error sense) number of quantizing bits M of such a loading $\frac{A_m}{\sigma} \simeq 2.6$?. Max [5] and Gray and Zeoli [6] have given answers to the inverse problem fixing the number of levels what is the optimum loading factor corresponding to a minimum variance (mean square error) of the total noise contributed by the quantization noise and the clipping noise.

From [6, table I] we reproduce the following table to guide the choice of M:

M	3	4	5	6
$\frac{A_m}{\sigma}$	1.9 ± 0.19	2.5 ± 0.25	2.9 ± 0.29	3.26 ± 0.326

This table indicates that with the clipping probability fixed above the 4 bit quantizer offers the optimum choice for low SNR conditions

<u>The signal plus noise case</u>

Passing now to the signal plus noise case, we propose to examine the question : is the 4 bit quantizer still convenient?. The answer will be given through surveying the preceding constraints:

i- Constraint (1) is respected if A_m is chosen such that : $A_m = A+2.6$

ii- With $A_m = A + 2.6\,\sigma$, the 4- bit quantizer step size Q is given by :

$$Q = \frac{2(A + 2.6\,\sigma)}{15} = \frac{A}{7.5} + \frac{5.2}{15}\,\sigma \qquad (6)$$

$$\frac{A}{\sigma} \leq 4.9 \qquad \text{or} \qquad \left(\frac{A}{\sigma}\right)_{dB} \leq 13.8 \text{ dB} \tag{7}$$

iii- For no signal suppression and satisfaction of relation (6) one obtains:

$$\frac{A}{\sigma} \geq 0.186 \qquad \text{or} \qquad \left(\frac{A}{\sigma}\right)_{dB} \geq -14.6\text{dB} \tag{8}$$

Combining (7) and (8) one obtains :

$$-14.6 \leq \left(\frac{A}{\sigma}\right)_{dB} \leq 13.8 \tag{9}$$

iv- In the SNR range fixed by (9), one guarantes :

$$P_c \leq 1\ \%$$

$$Q \leq \sigma$$

Consequently, as shown by the simulation results presented later on:

$$E\left[Z_Q^2\right] = A^2 + \sigma^2 + Q^2/12$$

Hence : $$(SNR)_{out} = \frac{A^2}{\sigma^2 + \frac{Q^2}{12}} = (SNR)_{in} \frac{1}{1+(Q^2/\sigma^2)/12}$$

i.e for the worst case $Q = \sigma$

$$(SNR)_{out} = (SNR)_{in} - 0.35 \quad \text{dB}$$

Which satisfies the constraint (4) as shown in [14] for the noisy Manchester coded data signal filtered by Butterworth filters with order $L \geq 2$ and cut-off frequency $F_c \in (F, 10F)$.

The question is now : Is it interesting to increase the number of bit, i.e. M = 5 or even higher?

If the same procedure is repeated for M = 5 and $P_c \leq 1\%$ one can show that the $\left(\frac{E}{N_o}\right)$ range is enlarged while the gain in $(SNR)_{out}$ over the required range (0-15) dB is trivial ($\lesssim 0.2$ dB) relative to the case where M = 4 bit. This gain is evidently obtained by reducing the quantization step for M=5. This trivial gain does not justify the increased complexity of the circuits succeeding the quantizer when M is increased to 5 bit. Therefore the 4 bit quantizer will be adopted.

IV. COMPUTER SIMULATION AND ANALYSIS OF RESULTS

The proposed 4-bit uniform quantizer loaded according to the linear rule is simulated on the Desktop computer HP-9845. The simulation scheme is given in Fig. 2. The binary Manchester signal is generated from a NRZ-L signal by modulation while the white noise sequence is generated using the Box-Muller Algorithm [10] . Then the noisy signal is filtered by Butterworth filters of various orders L and the filtered signal is then sampled to generate $Z(iT_s)$ upon which the quantizer logic is applied to generate $Z_Q(iT_s)$. The quantization noise is obtained as $n_Q(iT_s) = Z_Q(iT_s) - Z(iT_s)$. The sequences $n_Q(iT_s)$ and $Z(iT_s)$ were processed using a Fast Fourier Transform Algorithm (FFT) to generate the quantization noise ACF: $R_n(iT_s)$ (Hamming window is used) and the crosscorrelation function $R_{ZnQ}(iT_s)$. Two examples of results are given in Fig. 3 and 4 for $(E/N_o) = 0$ and for $F_s = 32$ F and $F_c = 4F$ and F respectively using a 4-pole LP Butterworth filter. The use of this high sampling rate was intended to examine the influence of high input correlation on the properties of quantizer output.

From then figures it can be seen that the quantization noise is practically "white" and practically uncorrelated with the input signal as the value for the autocorrelation and crosscorrelation coefficients are less that 5% (for $i \neq 0$ in the figures concerning the autocorrelation coefficient). This confirms that the chosen clipping probability is low enough to ensure that the practical quantizer behaves as the theoretical one. It is believed that a longer record length than the 1024 sample values would give better results in the sense of decreasing the correlation coefficient coefficient values. It must be underlined that an input samples autocorrelation coefficient $\rho_Z(\tau = T_s)$ given by :

$$\rho_Z(\tau) = \rho_s(\tau) . \frac{\sigma_s^2}{\sigma_s^2 + \sigma^2} + \rho_n(\tau) \frac{\sigma^2}{\sigma_s^2 + \sigma^2} \tag{10}$$

of 96 % was used (Fig. 3). Therefore we may give as a rule of thumb : The quantization noise is pseudowhite as far as the input samples are not almost completely correlated (values of correlation coefficient of about 0,96 are tolerable). This value (0.96) is clearly compatible with the 0.9 for gaussian input and $Q = \sigma$ given in [1] and 0.9999 for M = 7 given in [8,9].

In conclusion, we can write

$$R_{ZQ}(\tau) = R_Z(\tau) + R_{nQ}(\tau) + 2R_{ZnQ}(\tau)$$
$$\cong R_Z(\tau) + R_{nQ}(\tau)$$

and

$$R_{nQ}(\tau) \cong R_{nQ}(0)\ \delta(\tau)$$

This is for 2nd joint moment relations of the quantization theorem. As to the first moment relations, the density function of the quantization noise is verified by hestograms to be uniform. But for the total output noise in the considered model, it should be kept in mind that, even if the quantization noise is practically dominated by the additive noise, one can infer nothing concerning the probability density of the output additive noise of the quantizer. For sure, this noise is not gaussian as the quantizer performs a nonlinear operation on its input. Hence, the output densities should be calculated, if necessary, using the standard method of random variable transformations (see for example [11], or the same reasoning leading to the quantization theory (see [1]).

As a final point, it is interesting to put the used linear loading rule in contrast to the famous "4 σ - loading rule derived from voice quantization studies. The proposed linear rule is thought to allow a good noise-free performance and good tolerance to noisy environment. This is claimed for by the fact that the "4 σ - loading" ($\frac{A_m}{\sigma} = 4$) implies that for a given RMS signal value σ_Z the quantizer step size should be $Q = \frac{8\sigma_Z}{2^{M-1}}$. This means that for a filtered noiseless signal (with $F_c \cong F$) only a fraction of the quantizer range will be used as $\sigma_Z = A/\sqrt{3}$.

Then $A_m = 4\frac{A}{\sqrt{3}}$ and consequently $A = 0.43\ A_m$ (for nonfiltered signal $A = 0.5\ A_m$). Pointing out that the amplitude distribution of this filtered signal is nearly uniform (between $+A$, $-A$) and that the uniform quantizer is best adapted to this situation, one sees that the "4σ" loading utilizes only 43% of the quantizer range and therefore is not convenient. On the contrary, the proposed linear design rule leads, in this same situation, to a full utilization of the quantizer range, thus ensuring a good noise free performance as a lower quantization noise is produced.

V. CONCLUSIONS

It has been shown that a 4 bit uniform step size quantizer loaded according to the proposed linear rule presents a good compromise between the implementation complexity (growing with the increase in the number of bits) and the accuracy with the increase in the number of bits) and the accuracy with which the statistics of the signal is recovered from the quantized version. This has been performed using the quantization theorem which does not apply to the noise free binary biphase-L signal whereas it has been shown that it applies, with good approximations, when a gaussian noise is added to the signal.

A computer simulation for the specified quantizer shows that the quantization noise is practically pseudowhite and noncorrelated with the nonquantized signal as far as the input samples autocorrelation coefficient is of he order of 96%.

REFERENCES

1 KORN G.A. Random Process Simulation And Measurements" McGraw-Hill, 1966, Chap. 6.

2 WATTS, D.G. "A General Theory of Amplitude Quantization with application to Correlation Determination" Proc. IEEE, Vol. 109 C, pp. 209-218, 1962.

3 ROSS D.T."Sampling and Quantizing" in "Notes on Analog-Digital Conversion Techniques" M.I.T. Press 1957. Edited by A.K. Susskind.

4 SPIRAD A.B. and DRYDER D.L. "A Necessary and Sufficient Condition for Quantization Noise to be Uniform and White " IEEE Trans. Acoust Speech, Signal Processing, Vol. ASSP-25 pp 442-448, Oct. 1977.

5 Max. J. "Quantization for minimum Distortion" IEEE Trans. Inform Theory, Vol. IT-6, pp 7-12, March 1960.

6 GRAY G.A. and ZEOLI G.W. "Quantization and Saturation noise due to Analog-to-digital Conversion" IEEE Trans. Aerosp. Electron. Syst. Vol. AES-7, pp 222-223, January 1971.

7 GALLAGHER N.C. and BUCKIEW J.A. "Some Recent Developments in Quantization Theory" IEEE 12th Annual Southeastern Symposium an System Theory, pp. 295-301, Virginia Beach 19-20 May 1980.

8 BENNET, W.R. "Spectra of Quantized Signals" B.S.T.J, Vol. 27, No 3, pp. 446-472, July 1948.

9 CATERMOLE K.W. "Principles of Pulse Code Modulation", Iliffe Books, 1969 Chap. 3.

In conclusion, we can write

$$R_{ZQ}(\tau) = R_Z(\tau) + R_{nQ}(\tau) + 2R_{ZnQ}(\tau)$$
$$\cong R_Z(\tau) + R_{nQ}(\tau)$$

and

$$R_{nQ}(\tau) \cong R_{nQ}(0)\ \delta(\tau)$$

This is for 2^{nd} joint moment relations of the quantization theorem. As to the first moment relations, the density function of the quantization noise is verified by hestograms to be uniform. But for the total output noise in the considered model, it should be kept in mind that, even if the quantization noise is practically dominated by the additive noise, one can infer nothing concerning the probability density of the output additive noise of the quantizer. For sure, this noise is not gaussian as the quantizer performs a nonlinear operation on its input. Hence, the output densities should be calculated, if necessary, using the standard method of random variable transformations (see for example [11], or the same reasoning leading to the quantization theory (see [1]).

As a final point, it is interesting to put the used linear loading rule in contrast to the famous "4 σ - loading rule derived from voice quantization studies. The proposed linear rule is thought to allow a good noise-free performance and good tolerance to noisy environment. This is claimed for by the fact that the "4 σ - loading" ($\frac{A_m}{\sigma} = 4$) implies that for a given RMS signal value σ_Z the quantizer step size should be $Q = \frac{8\sigma_Z}{2^{M-1}}$. This means that for a filtered noiseless signal (with $F_c \cong F$) only a fraction of the quantizer range will be used as $\sigma_Z = A/\sqrt{3}$.

Then $A_m = 4\frac{A}{\sqrt{3}}$ and consequently $A = 0.43\ A_m$ (for nonfiltered signal $A = 0.5\ A_m$). Pointing out that the amplitude distribution of this filtered signal is nearly uniform (between $+A$, $-A$) and that the uniform quantizer is best adapted to this situation, one sees that the "4σ" loading utilizes only 43% of the quantizer range and therefore is not convenient. On the contrary, the proposed linear design rule leads, in this same situation, to a full utilization of the quantizer range, thus ensuring a good noise free performance as a lower quantization noise is produced.

V. CONCLUSIONS

It has been shown that a 4 bit uniform step size quantizer loaded according to the proposed linear rule presents a good compromise between the implementation complexity (growing with the increase in the number of bits) and the accuracy with the increase in the number of bits) and the accuracy with which the statistics of the signal is recovered from the quantized version. This has been performed using the quantization theorem which does not apply to the noise free binary biphase-L signal whereas it has been shown that it applies, with good approximations, when a gaussian noise is added to the signal.

A computer simulation for the specified quantizer shows that the quantization noise is practically pseudowhite and noncorrelated with the nonquantized signal as far as the input samples autocorrelation coefficient is of he order of 96%.

REFERENCES

1 KORN G.A. Random Process Simulation And Measurements" McGraw-Hill, 1966, Chap. 6.

2 WATTS, D.G. "A General Theory of Amplitude Quantization with application to Correlation Determination" Proc. IEEE, Vol. 109 C, pp. 209-218, 1962.

3 ROSS D.T."Sampling and Quantizing" in "Notes on Analog-Digital Conversion Techniques" M.I.T. Press 1957. Edited by A.K. Susskind.

4 SPIRAD A.B. and DRYDER D.L. "A Necessary and Sufficient Condition for Quantization Noise to be Uniform and White " IEEE Trans. Acoust Speech, Signal Processing, Vol. ASSP-25 pp 442-448, Oct. 1977.

5 Max. J. "Quantization for minimum Distortion" IEEE Trans. Inform Theory, Vol. IT-6, pp 7-12, March 1960.

6 GRAY G.A. and ZEOLI G.W. "Quantization and Saturation noise due to Analog-to-digital Conversion" IEEE Trans. Aerosp. Electron. Syst. Vol. AES-7, pp 222-223, January 1971.

7 GALLAGHER N.C. and BUCKIEW J.A. "Some Recent Developments in Quantization Theory" IEEE 12th Annual Southeastern Symposium an System Theory, pp. 295-301, Virginia Beach 19-20 May 1980.

8 BENNET, W.R. "Spectra of Quantized Signals" B.S.T.J, Vol. 27, No 3, pp. 446-472, July 1948.

9 CATERMOLE K.W. "Principles of Pulse Code Modulation", Iliffe Books, 1969 Chap. 3.

10 TAKOWITZ S.J. "Computational Probability and Simulation, Adison-Wesley, pp. 53-59, 1977.

11 PAPOULIS A. "Probability, Random Variables and RandomProcesses" McGraw-Hill, pp. 118-138, 1965.

12 THOMAS, J.B. "Introduction to statistical Communication Theory" 1969 (Chap. 6)

13 BLACHMAN, N.M "The uncorrelated output Components of a Nonlinearity" IEEE Trans. Information Theory, Vol. IT-14, pp. 250-255, March 1968.

14 MATAR, M.A. "Etude et Realisation d'un Synchronisateur Detecteur Numerique pour Signaux Biphase - L" (App.3) These Docteur- Ingenieur, ENSAE, Toulouse, FRANCE, 1981.

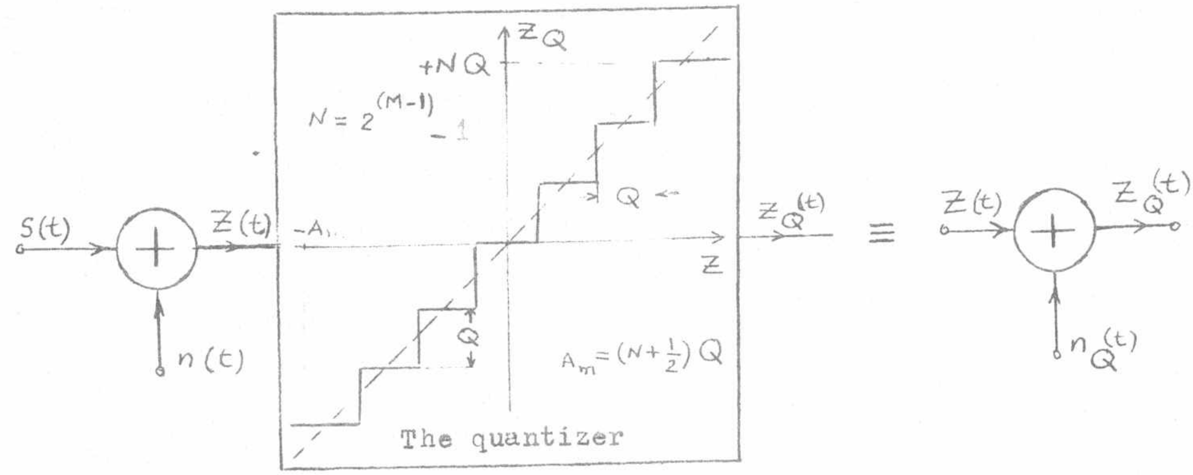

Fig.1. Uniform step-size quantizer: model cosidered.

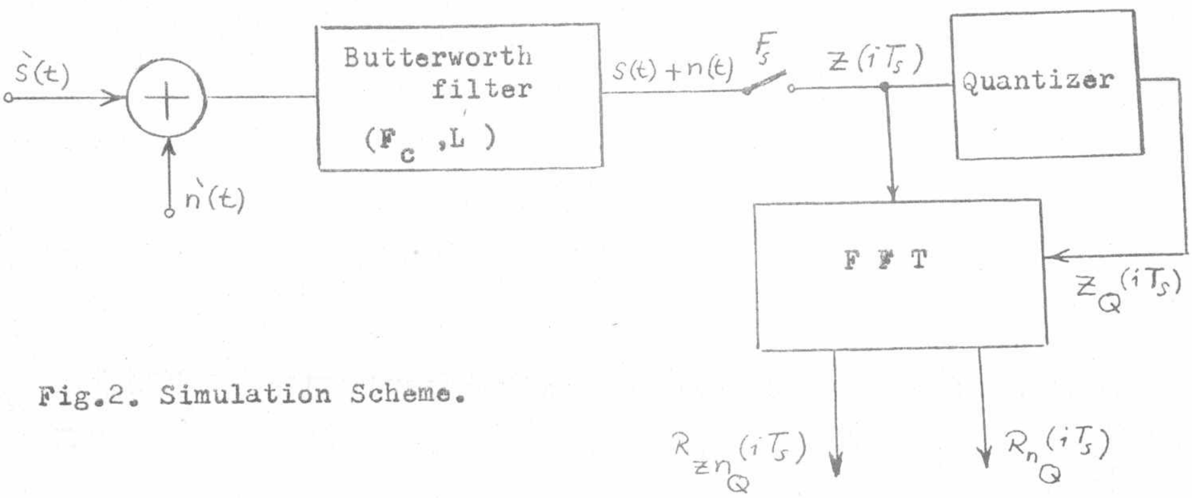

Fig.2. Simulation Scheme.

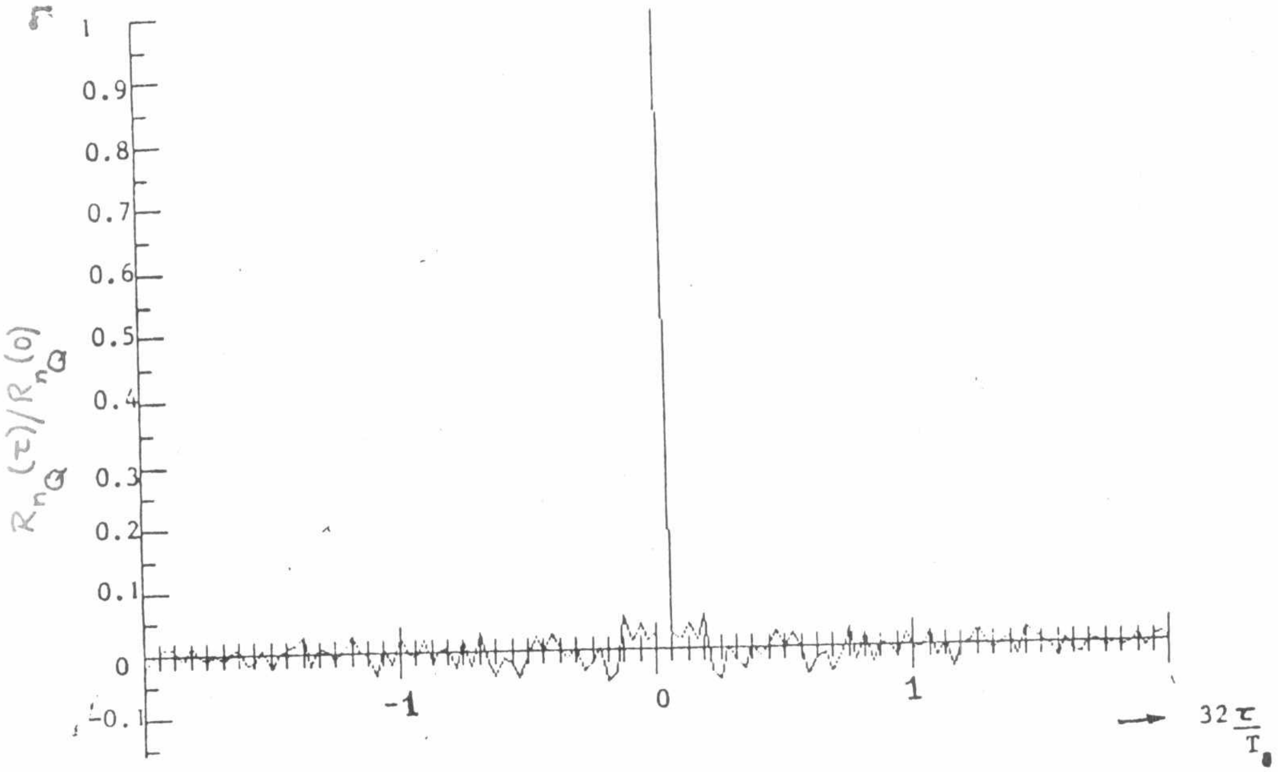

Fig 3.a. Autocorrelation coefficient $R_{n_Q}(\tau)/R_{n_Q}(0)$ $(F_c=4F)$.

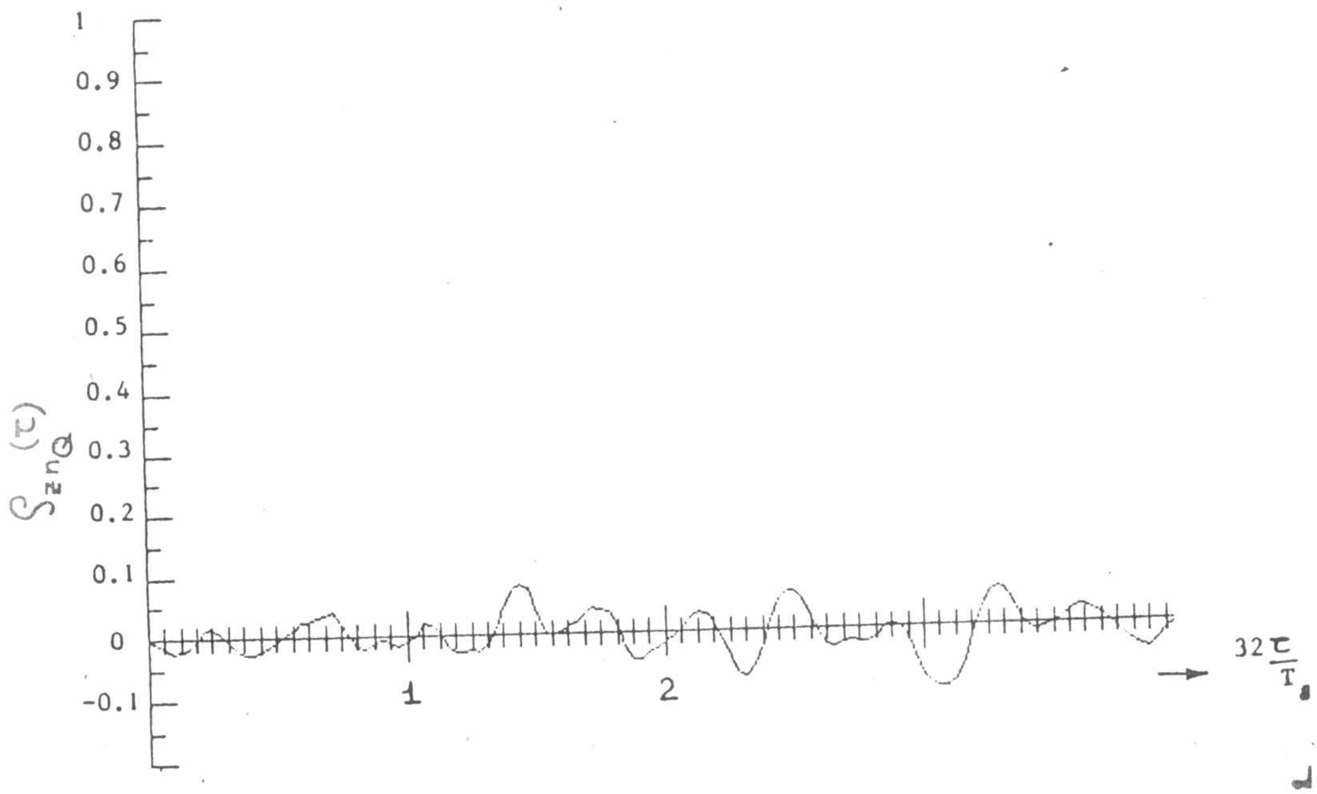

Fig 3.b. Crosscorrelation coefficient $\rho_{zn_Q}(\tau)$ $(F_c=4F)$.

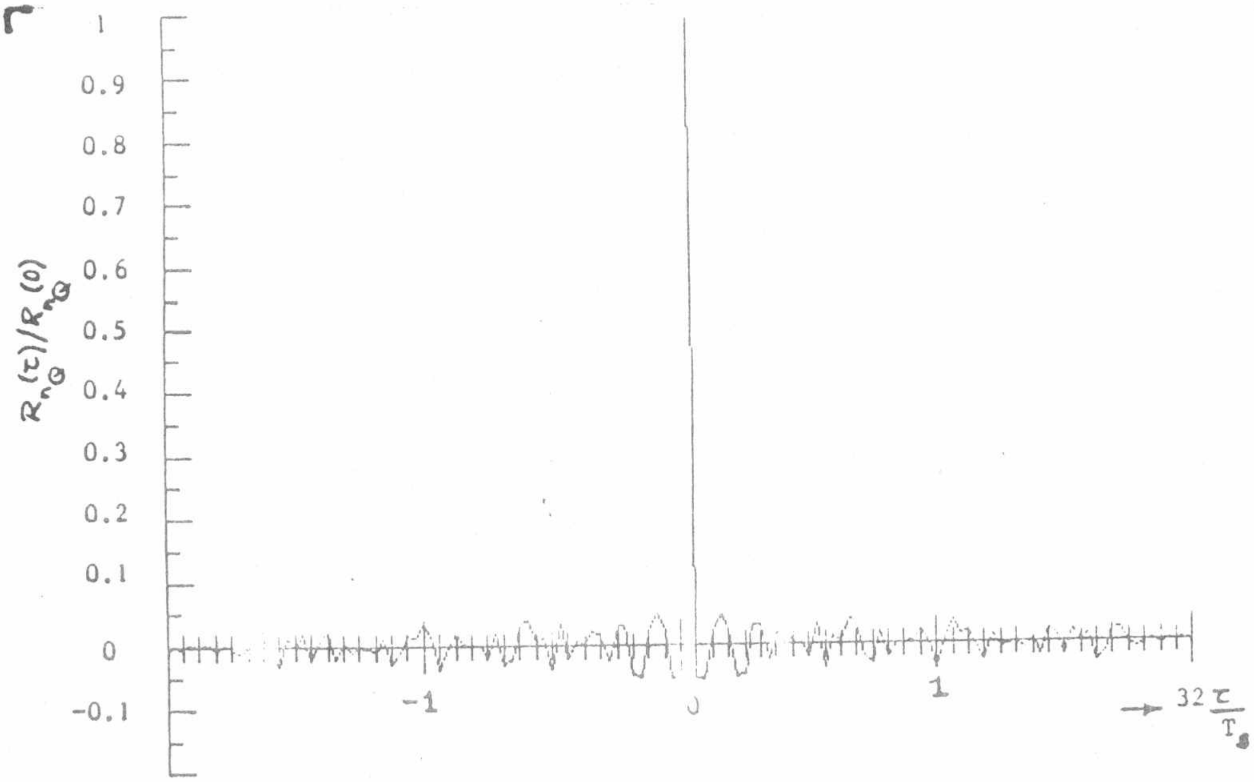

Fig 4.a. Autocorrelation coefficient $R_{n_Q}(\tau)/R_{n_Q}(0)$ $(F_c=F)$.

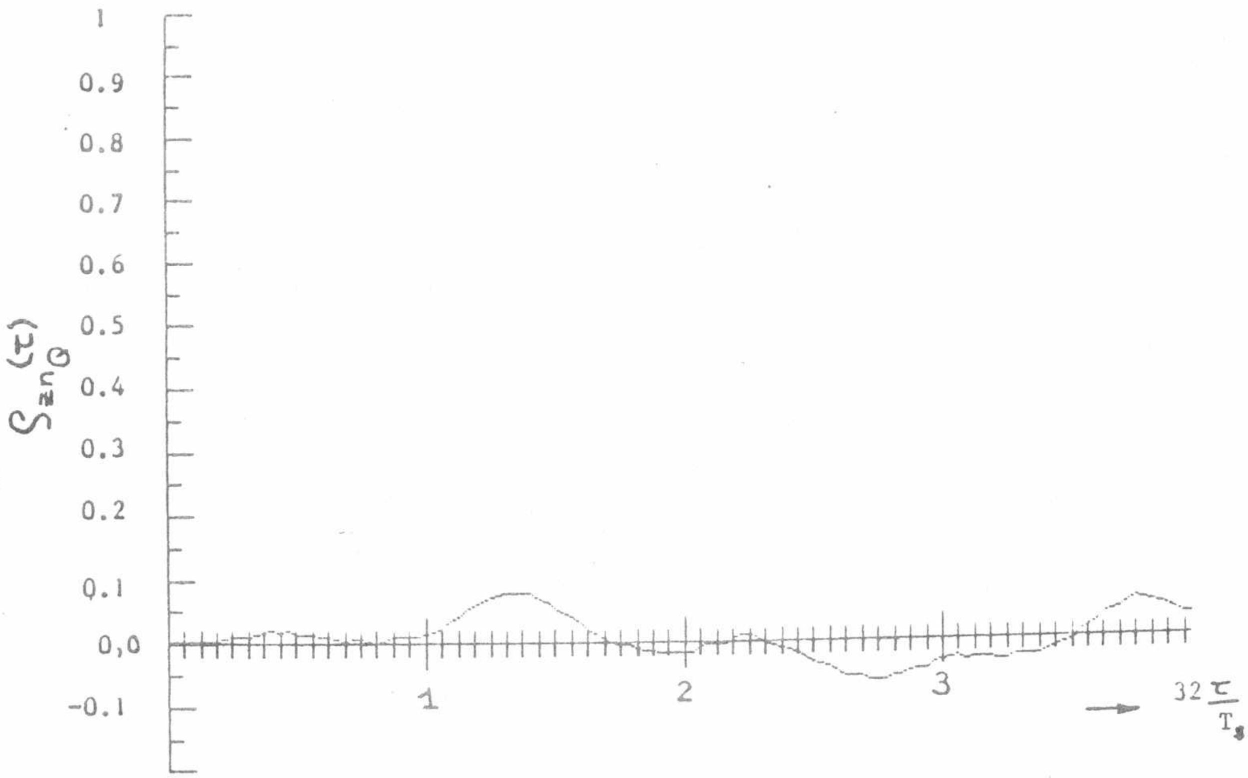

Fig 4.b. Crosscorrelation coefficient $\rho_{zn_Q}(\tau)$. $(F_c=F)$