General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

ESA SM-83
Agence Spatiale Européenne
QUANTIFICATEURS OPTIMUMS POUR UNE PROBABILITE D'ENTREE GAUSSIENNE ET POUR UNE MESURE DE DISTORSION EN VALEUR ABSOLUE
J.-C. Demaret (Université de Liège)
Juin 1975
iv + 23 pages

I. Demaret, J.-C. (Univ. de Liège)

II. ESA SM-83

III. Texte en anglais

Les paramètres de quantificateurs uniformes et non-uniformes, jusqu'à 10 bits de quantification, optimums pour une probabilité d'entrée gaussienne et pour une mesure de distorsion en valeur absolue sont déterminés. – On entend par quantificateurs optimums des quantificateurs à distorsion minimale. La méthode numérique d'optimalisation utilisée converge relativement vite. La comparaison entre quantificateurs optimums uniformes et non-uniformes est faite.

ESA SM-83
Agence Spatiale Européenne
QUANTIFICATEURS OPTIMUMS POUR UNE PROBABILITE D'ENTREE GAUSSIENNE ET POUR UNE MESURE DE DISTORSION EN VALEUR ABSOLUE
J.-C. Demaret (Université de Liège)
Juin 1975
iv + 23 pages

I. Demaret, J.-C. (Univ. de Liège)

II. ESA SM-83

III. Texte en anglais

Les paramètres de quantificateurs uniformes et non-uniformes, jusqu'à 10 bits de quantification, optimums pour une probabilité d'entrée gaussienne et pour une mesure de distorsion en valeur absolue sont déterminés. – On entend par quantificateurs optimums des quantificateurs à distorsion minimale. La méthode numérique d'optimalisation utilisée converge relativement vite. La comparaison entre quantificateurs optimums uniformes et non-uniformes est faite.

ESA SM-83
Agence Spatiale Européenne
QUANTIFICATEURS OPTIMUMS POUR UNE PROBABILITE D'ENTREE GAUSSIENNE ET POUR UNE MESURE DE DISTORSION EN VALEUR ABSOLUE
J.-C. Demaret (Université de Liège)
Juin 1975
iv + 23 pages

I. Demaret, J.-C. (Univ. de Liège)

II. ESA SM-83

III. Texte en anglais

Les paramètres de quantificateurs uniformes et non-uniformes, jusqu'à 10 bits de quantification, optimums pour une probabilité d'entrée gaussienne et pour une mesure de distorsion en valeur absolue sont déterminés. – On entend par quantificateurs optimums des quantificateurs à distorsion minimale. La méthode numérique d'optimalisation utilisée converge relativement vite. La comparaison entre quantificateurs optimums uniformes et non-uniformes est faite.

ESA SM-83
Agence Spatiale Européenne
QUANTIFICATEURS OPTIMUMS POUR UNE PROBABILITE D'ENTREE GAUSSIENNE ET POUR UNE MESURE DE DISTORSION EN VALEUR ABSOLUE
J.-C. Demaret (Université de Liège)
Juin 1975
iv + 23 pages

I. Demaret, J.-C. (Univ. de Liège)

II. ESA SM-83

III. Texte en anglais

Les paramètres de quantificateurs uniformes et non-uniformes, jusqu'à 10 bits de quantification, optimums pour une probabilité d'entrée gaussienne et pour une mesure de distorsion en valeur absolue sont déterminés. – On entend par quantificateurs optimums des quantificateurs à distorsion minimale. La méthode numérique d'optimalisation utilisée converge relativement vite. La comparaison entre quantificateurs optimums uniformes et non-uniformes est faite.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

ACKNOWLEDGEMENT

I am very grateful to the European Space Research Organisation and to the National Aeronautics and Space Administration from which I was awarded an International fellowship.

I thank Professors P. Bergmans and T. Berger for their guidance during my M.S. research.

OPTIMUM QUANTISERS FOR A GAUSSIAN INPUT PROBABILITY DENSITY AND FOR THE MAGNITUDE-ERROR DISTORTION MEASURE

ABSTRACT

The parameters of non-uniform and uniform quantisers up to ten bits of quantisation, optimum for a Gaussian input probability and for the magnitude-error distortion criterion are computed.

Optimum quantisers must be understood as quantisers with minimum distortion. The numerical method used for the optimisation converges relatively rapidly. The comparison between optimum non-uniform quantisers and optimum uniform quantisers is made.

PREFACE

This paper is part of a thesis presented to the Faculty of the Graduate School of Cornell University (Ithaca, N-Y, U.S.A.) for the Degree of Master of Science (January 1974). The thesis is entitled 'Digital methods for the estimation of the R.M.S. value of bandlimited singular Gaussian processes'.

We had to compute quantisers with minimum distortion when the input probability was a Gaussian distribution and when the magnitude-error criterion was taken. Although the algorithm used converged relatively rapidly, we hope in publishing the parameters of such quantisers, that other users will be saved computation time.

1. INTRODUCTION

The digitising process of physical phenomena (mostly continuous functions of time, denoted x(t)) consists of two parts: sampling and quantising (Figure 1).

The process of sampling consists of observing the continuous functions only at discrete times τ_i. The τ_i's may be distributed either uniformly or at random in $[\tau_a, \tau_b]$, where $[\tau_a, \tau_b]$ is the record length. It follows that a set of samples is denoted: $\{ x(\tau_i) \}$.

A set of predetermined values having been defined, the process of quantising consists of rounding off the value of each sample to the closest value of the set. This is performed by a quantiser.

Clearly, quantisation introduces an error. An attempt must be made to minimise this error by making the number of predetermined values as large as possible and, given this number, by choosing a suitable interval between these values.

This paper is concerned with the computation of such 'optimum' intervals.

The use of optimum quantisers is required in accurate digital measurement processes.

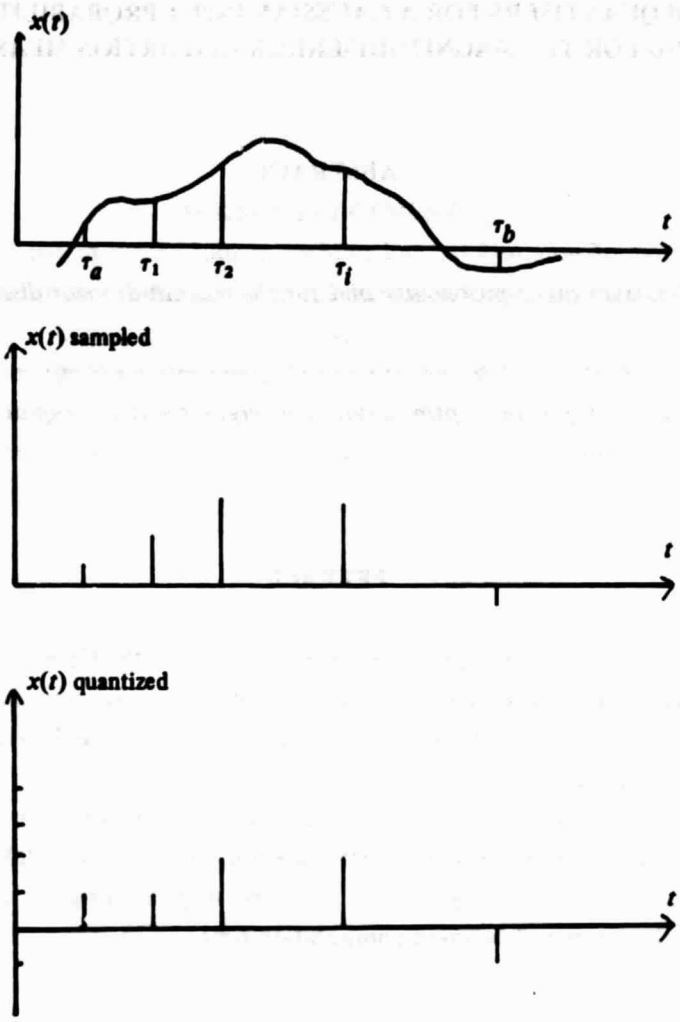

Figure 1. Continuous function digitising.

2. DEFINITIONS

A n_Q-level quantiser maps any real value of the input into one of n_Q real values at the output $x \in (T_i, T_{i+1}]$ is reproduced by the output L_i where the T_i's are the quantisation thresholds and the L_i's are the output levels (Figure 2).

The performance of a quantiser can be measured by the distortion introduced between input and output and by the output entropy. If $p(x)$ is the input probability density and $\rho(x,y)$* the distortion measure, the overall distortion D is:

---

* The cost function $\rho(x,y)$ is a non-negative function which specifies the penalty of reproducing x by y.

$$D = \sum_{i=1}^{n_l} \int_{T_i}^{T_{i+1}} \rho(x, L_i) \cdot p(x) \cdot dx \qquad (1)$$

The output entropy*, in nats, is given by:

$$H = -\sum_{i=1}^{n_l} Q_i \cdot \log_e (Q_i) \qquad (2)$$

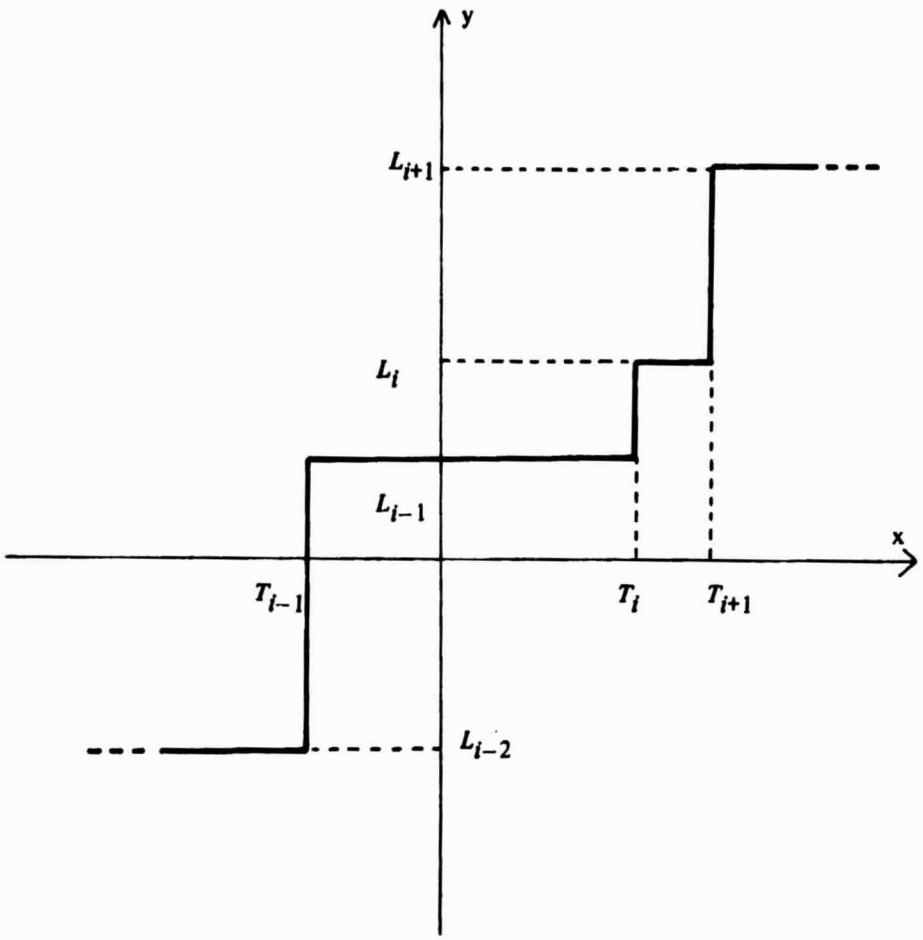

Figure 2. Quantiser configuration.

* The output entropy gives the average value of self-information over the output levels of the quantiser.

where:

$$Q_i = \text{Prob}\,(L_i) = \text{Prob}\,\{T_i < x \leqslant T_{i+1}\} = \int_{T_i}^{T_{i+1}} p(x)\,.\,dx$$

We define a n_b-bit uniform quantiser as a quantiser such that:

1) the thresholds are equally spaced:

$$|\,T_{i+1} - T_i\,| = \Delta T, \qquad i=1,\ \ldots,\ \qquad n_l = 2^{n_b},$$

(We shall call ΔT the threshold increment of the uniform quantiser);

2) the output levels are at the middle of the input ranges:

$$L_i = \frac{T_i + T_{i+1}}{2}\ ,\ i = 1\ ,\qquad \ldots\ ,\qquad n_l\ .$$

In data processing, one tries to minimise the distortion introduced by the quantisers. Thus, in what follows, optimising a quantiser will mean finding the T_i's and the L_i's such that D is minimum for the given $p(x)$ and $\rho(x,y)$.

We consider in what follows

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}}\ \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{3}$$

$$\rho(x,y) = |\,x - y\,| \tag{4}$$

The Gaussian probability density is found in many problems and the magnitude-error distortion is well suited to the concept of the absolute error of a measurement.

3. NUMERICAL COMPUTATION CONSIDERATIONS

The programs were run on the mini-computer DEC P.D.P. 11/40 of which the main characteristics are:

- 16-bit word,
- direct addressing of 16k-bit words (k = 1024),
- floating point arithmetic,
- ANSI-standard FORTRAN IV compiler.

A) As the input probability density, (3), is symmetric about the mean μ, optimum quantisers will be symmetric about μ too.

B) The mean of the Gaussian probability density is just a parameter which positions the density along the axis. We assume in what follows that the mean is zero; if not, we just translate the thresholds and the output of the quantiser.

C) We are interested, in fact, in n_b-bit quantisers. The number of levels is always even: $n_l = 2^{n_b}$.

D) Mathematically the support of the probability density is $(-\infty, +\infty)$, but because of the finite representation of numbers in the computer, the support will be finite, say $[-\alpha, +\alpha]$, i.e., Prob $\{x \notin [-\alpha, +\alpha]\} = 0$.

E) As consequence of remarks A, B, C and D:

1) a slightly different notation is used:
the set of thresholds is:

$$\left\{ T_i \;\; i = 1, \pm 2, \; \ldots \; , \pm \; \frac{n_l}{2} + 1 \right\} \;\; ;$$

the set of output levels is:

$$\left\{ L_i \, , \;\; i = \pm 1 \, , \;\; \ldots \; , \pm \; \frac{n_l}{2} \right\} ;$$

2) the midway threshold is always 0, $T_1 = 0$;

3) the extreme threshold is $T_{(n_l/2)+1} = \alpha$ (Figure 3). In fact we do not have to worry about the actual value of α (it actually depends on the computer itself), we just assume in the numerical computation that

$$\text{Prob} \left\{ x \notin [- T_{(n_l/2)+1}, + T_{(n_l/2)+1}] \right\} = 0;$$

4) we shall compute half quantisers, say for $x \geqslant 0$, i.e. compute the T_i's and the L_i's up to

$$i = \frac{n_l}{2} \;\; ;$$

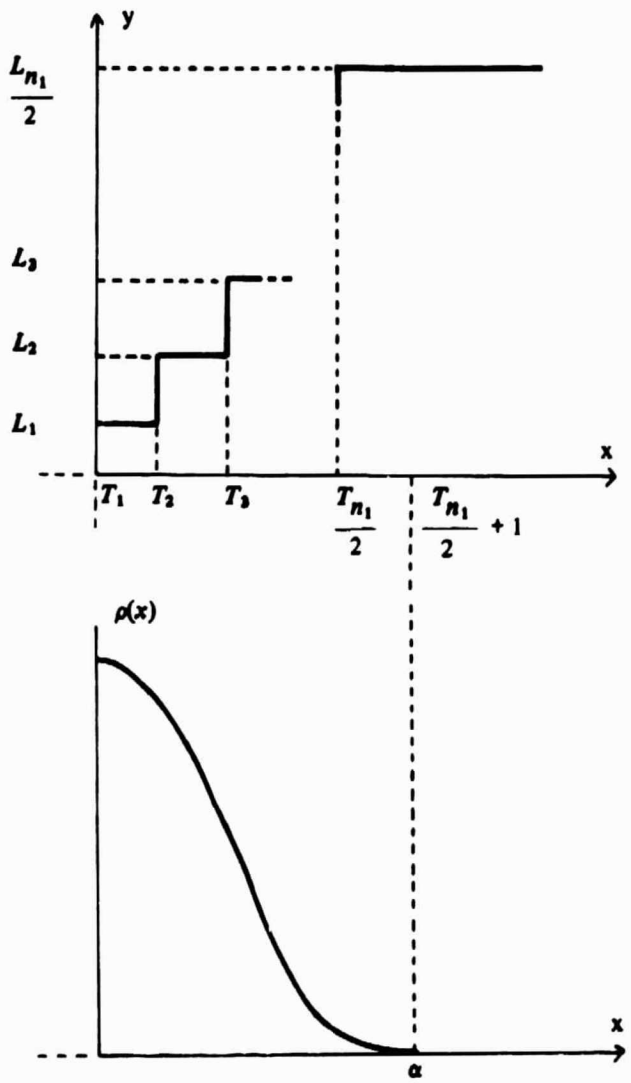

Figure 3. Numerical considerations on quantisers.

5) the practical expression for D is:

$$D = 2 \cdot \sum_{i=1}^{n_l/2} \int_{T_l}^{T_{l+1}} \rho(x, L_l) \cdot p(x) \cdot dx \qquad (5)$$

6) The output entropy is:

$$H = -2 \cdot \sum_{i=1}^{n_l/2} Q_l \cdot \log_e (Q_l) \qquad (6)$$

7) $T_l < L_l$, $l = 1, \quad \ldots, \quad n_l/2$

F) For the computation, we consider a unit variance; if the actual variance is σ^2, it suffices to multiply the T_l's, the L_l's and D by σ, the output entropy remaining the same.

This is justified by the two following equalities:

$$\frac{1}{\sigma\sqrt{2\pi}} \int_a^b \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = \frac{1}{\sqrt{2\pi}} \int_{\frac{a}{\sigma}}^{\frac{b}{\sigma}} \exp\left(-\frac{x^2}{2}\right) dx$$

$$\frac{1}{\sigma\sqrt{2\pi}} \int_a^b x \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = \frac{\sigma}{\sqrt{2\pi}} \int_{\frac{a}{\sigma}}^{\frac{b}{\sigma}} x \exp\left(-\frac{x^2}{2}\right) dx$$

4. OPTIMISATION

4.1. NON-UNIFORM QUANTISER

4.1.1. In view of what optimisation means in this context, the thresholds and the output levels must satisfy for

$$l = 1, \ldots, \frac{n_l}{2} : \frac{\partial D}{\partial T_l} = 0 \text{ and } \frac{\partial D}{\partial L_l} = 0$$

For the magnitude-error distortion measure, (4), we obtain for

$$l = 1, \ldots, \frac{n_l}{2} :$$

1. $$|T_l - L_{l-1}| = |T_l - L_l|$$

   $$\text{or } T_l = \frac{L_{l-1} + L_l}{2} \tag{7}$$

2. $$\int_{T_l}^{L_l} p(x) \cdot dx = \int_{L_l}^{T_{l+1}} p(x) \cdot dx$$

   $$\text{or } L_l = p^{-1}\left\{\frac{1}{2}[P(T_{l+1}) + P(T_l)]\right\} \tag{8}$$

For the $P(x)$ and the $P^{-1}(x)$ functions we refer to the Appendix. The second condition is known as the conditional median requirement for the L_l's.

It has been pointed out[2] that for some pathological input probability densities these conditions are only necessary but not sufficient. On the basis of our numerical results we conjecture that this is not our case and that the T_l's and L_l's satisfying (7) and (8) lead to an absolute minimum for D.

4.1.2. Computation

A. We use the relaxation method[3]. The main steps of this method are:

1. Initial partition of the T_l's:

$$\left\{ T_l : T_1 = 0,\ 0 < T_l < T_{l+1},\ l = 2, \ldots, \frac{n_l}{2} \right\},$$

The intial partition is arbitrary.

2. Finding the L_l's satisfying (8).
3. Computation of the T_l's satisfying (7).
4. Repetition of steps 2 and 3 until stabilisation occurs (as explained below).

B. When shall we stop the iterative process?

In numerical methods there exist several stopping criteria:

1. Absolute or relative accuracy in the unknowns.
2. A given number of steps.
3. Stabilisation of the solution.

Here we can stop when a new set $\{T_l\}$ or $\{L_l\}$ does not introduce a large variation in the value of D. Let us consider the actual value D_a of the distortion. Writing (5) for (4), we find that D_a is expressed by:

$$D_a = 2 \cdot \sum_{l=1}^{\frac{n_l}{2}} \int_{T_l}^{T_{l+1}} |x - L_l| \cdot p(x) \cdot dx$$

which can be written: $D_a = D_t + D_c$, with

$$D_t = 2 \cdot \sum_{l=1}^{\frac{n_l}{2}} \left[- \int_{T_l}^{L_l} x \cdot p(x) \cdot dx + \int_{L_l}^{T_{l+1}} x \cdot p(x) \cdot dx \right]$$

$$D_c = 2 \sum_{l=1}^{\frac{n_l}{2}} L_l \left[\int_{T_l}^{L_l} p(x) \cdot dx - \int_{L_l}^{T_{l+1}} p(x) \cdot dx \right]$$

D_t is the theoretical expression for the distortion and the term D_c is due to the fact that each L_l is not exactly the conditional median.

This offers us a stopping criterion which must occur after each complete iteration of the algorithm. We start with a set $\{T_l\}$, we compute the set $\{L_l\}$, and then the new set $\{T_l\}$. At this point one iteration is complete.

We shall stop the algorithm when D_c is a given fraction of D_t.

C. Algorithm.

Referring to Figure 4, we find the algorithm clearly presented in Figure 5. We define $n(x) \triangleq 1/\sqrt{2\pi} \exp(-x^2/2)$.
We also compute the output entropy H.

D. Results.

The results are given in Table 1 for $1 \leqslant n_b \leqslant 8$.

In Figure 7 we have plotted $\log D$ versus n_b and in Figure 8, H versus n_b.

It proved impossible to determine quantisers with $n_b \geqslant 9$, because the numerical approximation of the functions $P(x)$ and $P^{-1}(x)$ produced oscillations in the computation.

In Figure 9 we have plotted T_l versus l for the 4 –, 5 –, 6 – and 7-bit half quantisers.

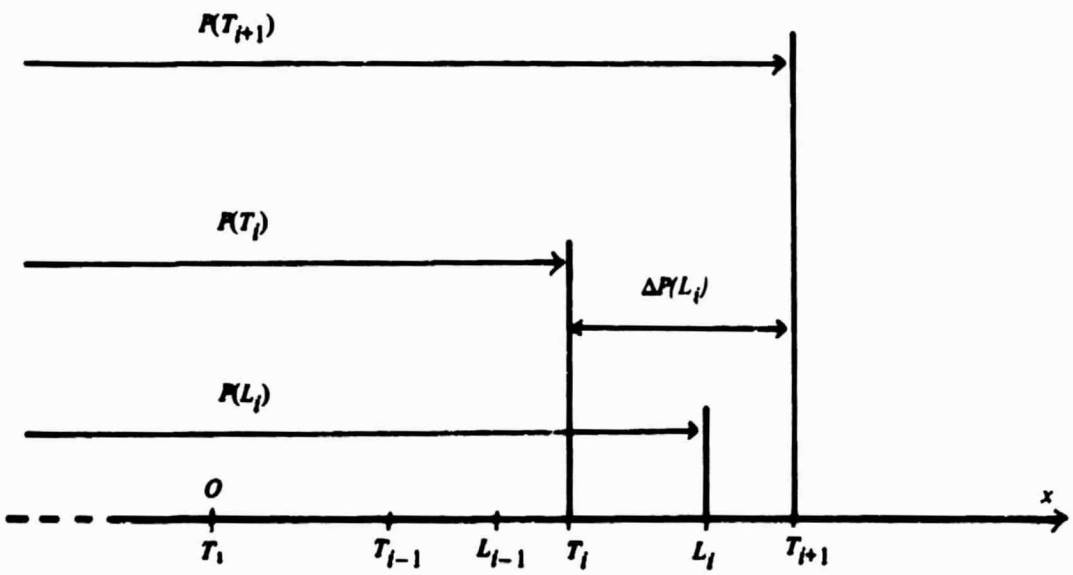

Figure 4. Computation notations for the optimisation.

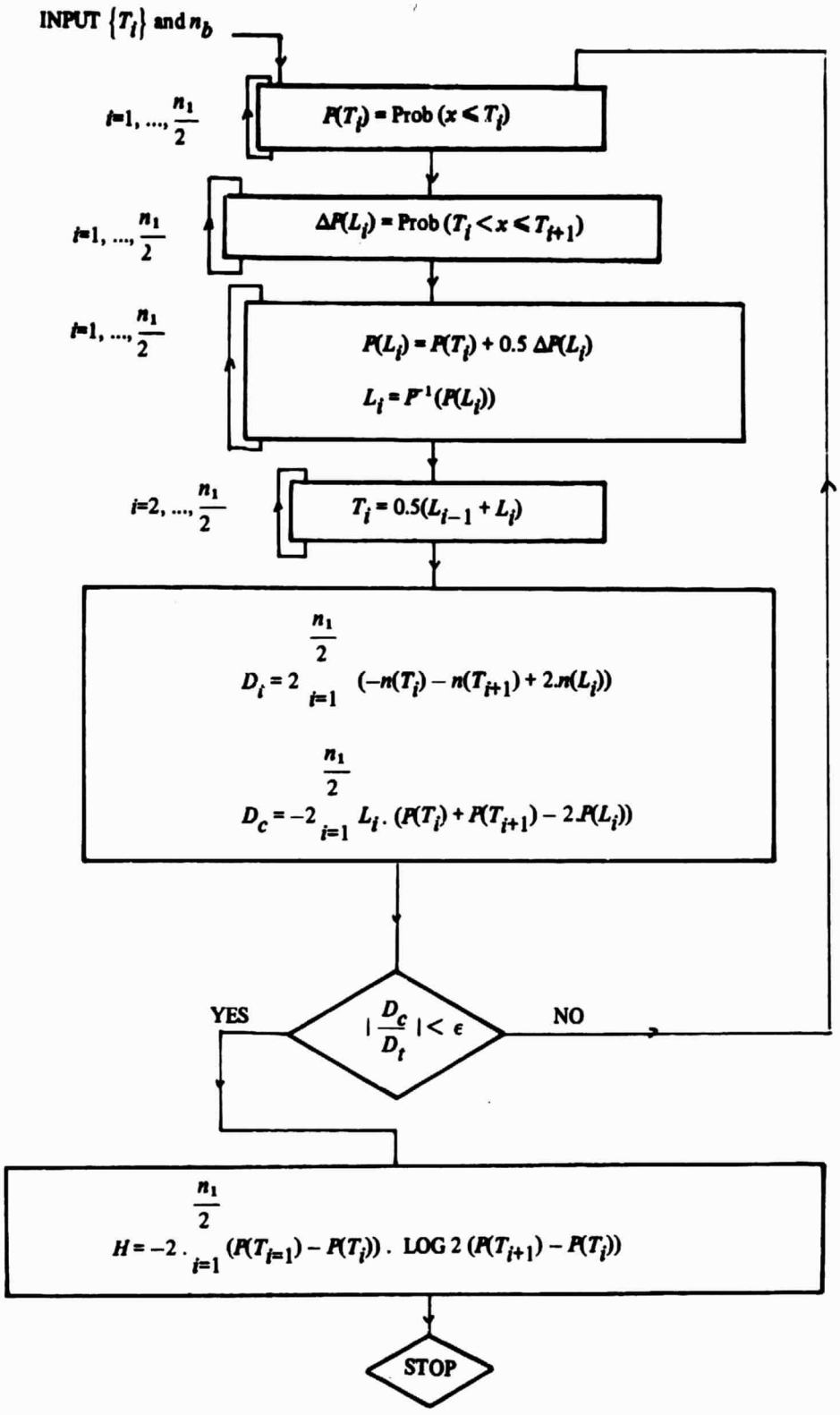

Figure 5. Optimisation algorithm of non-uniform quantisers.

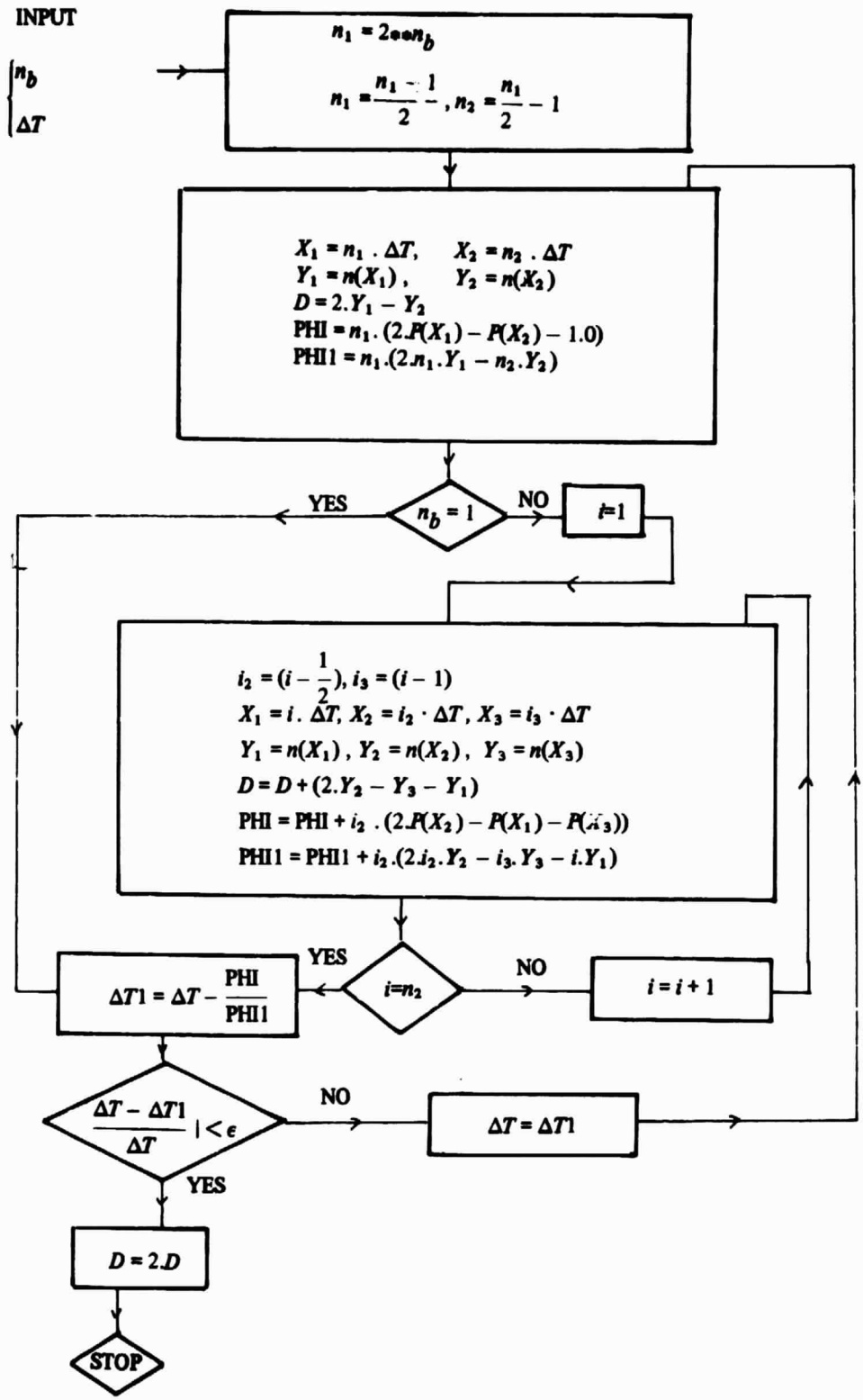

Figure 6. Optimisation of uniform quantisers.

TABLE 1

Parameters of the Optimum non-uniform Quantiser;
Input Probability Density: N(0, 1)
Distortion Measure: $\rho(x, y) = |x - y|$

NUMBER OF BITS: 1		
DISTORTION: 0.4735D 00 -0.2575D-03		ENTROPY: 0.1000D 01
	THRESHOLDS	LEVELS
1	0.0000D 00	0.7742D 00
NUMBER OF BITS: 2		
DISTORTION: 0.2657D 00 + 0.1945D-04		ENTROPY: 0.1977D 01
	THRESHOLDS	LEVELS
1	0.0000D 00	0.3773D 00
2	0.8216D 00	0.1266D 01
NUMBER OF BITS: 3		
DISTORTION: 0.1429D 00 + 0.8706D-05		ENTROPY: 0.2944D 01
	THRESHOLDS	LEVELS
1	0.0000D 00	0.2018D 00
2	0.4130D 00	0.6243D 00
3	0.8688D 00	0.1113D 01
4	0.1453D 01	0.1793D 01
NUMBER OF BITS: 4		
DISTORTION: 0.7455D-01 + 0.4128D-05		ENTROPY: 0.3915D 01
	THRESHOLDS	LEVELS
1	0.0000D 00	0.1041D 00
2	0.2097D 00	0.3154D 00
3	0.4256D 00	0.5358D 00
4	0.6537D 00	0.7717D 00
5	0.9019D 00	0.1032D 01
6	0.1182D 01	0.1333D 01
7	0.1521D 01	0.1709D 01
8	0.1988D 01	0.2267D 01
NUMBER OF BITS: 5		
DISTORTION: 03814D-01 + 0.3617D-04		ENTROPY: 0.4894D 01
	THRESHOLDS	LEVELS
1	0.0000D 00	0.5116D-01
2	0.1027D 00	0.1543D 00
3	0.2069D 00	0.2595D 00
4	0.3136D 00	0.3678D 00
5	0.4239D 00	0.4799D 00

NUMBER OF BITS: 5		
DISTORTION: 03814D-01 + 0.3617D-04		ENTROPY: 0.4894D 01
	THRESHOLDS	LEVELS
6	0.5384D 00	0.5968D 00
7	0.6580D 00	0.7192D 00
8	0.7837D 00	0.8482D 00
9	0.9167D 00	0.9851D 00
10	0.1059D 01	0.1132D 01
11	0.1212D 01	0.1292D 01
12	0.1380D 01	0.1468D 01
13	0.1569D 01	0.1670D 01
14	0.1791D 01	0.1911D 01
15	0.2067D 01	0.2223D 01
16	0.2462D 01	0.2702D 01
NUMBER OF BITS: 6		
DISTORTION: 0.1935D-01 – 0.1726D-04		ENTROPY: 0.5878D 01
	THRESHOLDS	LEVELS
1	0.0000D 00	0.2530D-01
2	0.5072D-01	0.7614D-01
3	0.1019D 00	0.1276D 00
4	0.1537D 00	0.1799D 00
5	0.2066D 00	0.2333D 00
6	0.2606D 00	0.2879D 00
7	0.3158D 00	0.3438D 00
8	0.3724D 00	0.4010D 00
9	0.4302D 00	0.4595D 00
10	0.4894D 00	0.5193D 00
11	0.5498D 00	0.5804D 00
12	0.6116D 00	0.6428D 00
13	0.6746D 00	0.7064D 00
14	0.7389D 00	0.7715D 00
15	0.8047D 00	0.8379D 00
16	0.8719D 00	0.9058D 00
17	0.9407D 00	0.9755D 00
18	0.1011D 01	0.1047D 01
19	0.1084D 01	0.1121D 01
20	0.1159D 01	0.1198D 01
21	0.1238D 01	0.1278D 01
22	0.1320D 01	0.1362D 01
23	0.1406D 01	0.1451D 01
24	0.1498D 01	0.1545D 01
25	0.1596D 01	0.1648D 01
26	0.1704D 01	0.1760D 01
27	0.1822D 01	0.1884D 01
28	0.1954D 01	0.2025D 01
29	0.2108D 01	0.2190D 01
30	0.2291D 01	0.2393D 01
31	0.2527D 01	0.2661D 01
32	0.2873D 01	0.3085D 01

NUMBER OF BITS: 7

DISTORTION: 0.9861D-02 – 0.5744D-05 | ENTROPY: 0.6864D 01

	THRESHOLDS	LEVELS		THRESHOLDS	LEVELS
1	0.0000D 00	0.9821D-02	2	0.1968D-01	0.2954D-01
3	0.3951D-01	0.4948D-01	4	0.5961D-01	0.6975D-01
5	0.8011D-01	0.9047D-01	6	0.1011D 00	0.1117D 00
7	0.1227D 00	0.1336D 00	8	0.1449D 00	0.1561D 00
9	0.1678D 00	0.1794D 00	10	0.1915D 00	0.2035D 00
11	0.2160D 00	0.2284D 00	12	0.2413D 00	0.2542D 00
13	0.2675D 00	0.2808D 00	14	0.2945D 00	0.3083D 00
15	0.3225D 00	0.3366D 00	16	0.3512D 00	0.3658D 00
17	0.3808D 00	0.3958D 00	18	0.4113D 00	0.4267D 00
19	0.4425D 00	0.4583D 00	20	0.4745D 00	0.4906D 00
21	0.5072D 00	0.5237D 00	22	0.5405D 00	0.5574D 00
23	0.5745D 00	0.5917D 00	24	0.6091D 00	0.6266D 00
25	0.6443D 00	0.6620D 00	26	0.6799D 00	0.6978D 00
27	0.7160D 00	0.7341D 00	28	0.7524D 00	0.7708D 00
29	0.7893D 00	0.8078D 00	30	0.8265D 00	0.8451D 00
31	0.8640D 00	0.8828D 00	32	0.9017D 00	0.9207D 00
33	0.9398D 00	0.9588D 00	34	0.9781D 00	0.9973D 00
35	0.1017D 01	0.1036D 01	36	0.1055D 01	0.1075D 01
37	0.1095D 01	0.1114D 01	38	0.1134D 01	0.1154D 01
39	0.1174D 01	0.1194D 01	40	0.1214D 01	0.1234D 01
41	0.1255D 01	0.1276D 01	42	0.1296D 01	0.1317D 01
43	0.1339D 01	0.1360D 01	44	0.1382D 01	0.1403D 01
45	0.1425D 01	0.1448D 01	46	0.1471D 01	0.1493D 01
47	0.1517D 01	0.1541D 01	48	0.1565D 01	0.1589D 01
49	0.1615D 01	0.1640D 01	50	0.1667D 01	0.1693D 01
51	0.1721D 01	0.1749D 01	52	0.1778D 01	0.1808D 01
53	0.1839D 01	0.1870D 01	54	0.1903D 01	0.1936D 01
55	0.1972D 01	0.2008D 01	56	0.2046D 01	0.2085D 01
57	0.2127D 01	0.2169D 01	58	0.2216D 01	0.2263D 01
59	0.2316D 01	0.2369D 01	60	0.2429D 01	0.2490D 01
61	0.2562D 01	0.2633D 01	62	0.2722D 01	0.2811D 01
63	0.2931D 01	0.3051D 01	64	0.3243D 01	0.3435D 01

NUMBER OF BITS: 8

DISTORTION: 0.5208D-02 -- 0.1927D-06 | ENTROPY: 0.7815D 01

	THRESHOLDS	LEVELS		THRESHOLDS	LEVELS
1	0.0000D 00	0.3086D-02	2	0.6199D-02	0.9303D-02
3	0.1243D-01	0.1556D-01	4	0.1872D-01	0.2188D-01
5	0.2509D-01	0.2830D-01	6	0.3158D-01	0.3485D-01
7	0.3819D-01	0.4154D-01	8	0.4497D-01	0.4840D-01
9	0.5193D-01	0.5545D-01	10	0.5908D-01	0.6271D-01
11	0.6646D-01	0.7020D-01	12	0.7407D-01	0.7794D-01
13	0.8194D-01	0.8594D-01	14	0.9009D-01	0.9423D-01
15	0.9852D-01	0.1028D 00	16	0.1073D 00	0.1117D 00
17	0.1163D 00	0.1209D 00	18	0.1257D 00	0.1305D 00

NUMBER OF BITS: 8					
DISTORTION: 0.5208D-02 – 0.1927D-06			ENTROPY: 0.7815D 01		
	THRESHOLDS	LEVELS		THRESHOLDS	LEVELS
19	0.1354D 00	0.1404D 00	20	0.1455D 00	0.1507D 00
21	0.1560D 00	0.1613D 00	22	0.1668D 00	0.1724D 00
23	0.1781D 00	0.1838D 00	24	0.1897D 00	0.1956D 00
25	0.2017D 00	0.2078D 00	26	0.2142D 00	0.2205D 00
27	0.2270D 00	0.2335D 00	28	0.2403D 00	0.2470D 00
29	0.2540D 00	0.2609D 00	30	0.2681D 00	0.2752D 00
31	0.2826D 00	0.2899D 00	32	0.2975D 00	0.3051D 00
33	0.3129D 00	0.3206D 00	34	0.3286D 00	0.3366D 00
35	0.3448D 00	0.3530D 00	36	0.3614D 00	0.3697D 00
37	0.3783D 00	0.3869D 00	38	0.3956D 00	0.4044D 00
39	0.4133D 00	0.4223D 00	40	0.4314D 00	0.4405D 00
41	0.4498D 00	0.4591D 00	42	0.4685D 00	0.4780D 00
43	0.4876D 00	0.4972D 00	44	0.5069D 00	0.5167D 00
45	0.5266D 00	0.5365D 00	46	0.5465D 00	0.5565D 00
47	0.5667D 00	0.5768D 00	48	0.5871D 00	0.5973D 00
49	0.6077D 00	0.6181D 00	50	0.6286D 00	0.6390D 00
51	0.6496D 00	0.6602D 00	52	0.6708D 00	0.6815D 00
53	0.6922D 00	0.7029D 00	54	0.7137D 00	0.7245D 00
55	0.7353D 00	0.7462D 00	56	0.7571D 00	0.7680D 00
57	0.7789D 00	0.7899D 00	58	0.8008D 00	0.8118D 00
59	0.8228D 00	0.8338D 00	60	0.8448D 00	0.8559D 00
61	0.8669D 00	0.8779D 00	62	0.8890D 00	0.9000D 00
63	0.9111D 00	0.9221D 00	64	0.9332D 00	0.9442D 00
65	0.9553D 00	0.9663D 00	66	0.9773D 00	0.9884D 00
67	0.9994D 00	0.1010D 01	68	0.1021D 01	0.1032D 01
69	0.1043D 01	0.1054D 01	70	0.1065D 01	0.1076D 01
71	0.1087D 01	0.1098D 01	72	0.1109D 01	0.1120D 01
73	0.1130D 01	0.1141D 01	74	0.1152D 01	0.1163D 01
75	0.1174D 01	0.1184D 01	76	0.1195D 01	0.1206D 01
77	0.1216D 01	0.1227D 01	78	0.1238D 01	0.1248D 01
79	0.1259D 01	0.1269D 01	80	0.1280D 01	0.1290D 01
81	0.1301D 01	0.1311D 01	82	0.1322D 01	0.1332D 01
83	0.1343D 01	0.1353D 01	84	0.1364D 01	0.1374D 01
85	0.1384D 01	0.1395D 01	86	0.1405D 01	0.1415D 01
87	0.1426D 01	0.1436D 01	88	0.1446D 01	0.1457D 01
89	0.1467D 01	0.1477D 01	90	0.1487D 01	0.1498D 01
91	0.1508D 01	0.1519D 01	92	0.1529D 01	0.1539D 01
93	0.1550D 01	0.1560D 01	94	0.1571D 01	0.1581D 01
95	0.1592D 01	0.1602D 01	96	0.1613D 01	0.1624D 01
97	0.1635D 01	0.1646D 01	98	0.1657D 01	0.1668D 01
99	0.1679D 01	0.1690D 01	100	0.1702D 01	0.1713D 01
101	0.1725D 01	0.1737D 01	102	0.1749D 01	0.1761D 01
103	0.1773D 01	0.1786D 01	104	0.1799D 01	0.1811D 01
105	0.1825D 01	0.1838D 01	106	0.1852D 01	0.1866D 01
107	0.1880D 01	0.1894D 01	108	0.1909D 01	0.1924D 01
109	0.1940D 01	0.1956D 01	110	0.1972D 01	0.1989D 01
111	0.2006D 01	0.2024D 01	112	0.2042D 01	0.2061D 01
113	0.2080D 01	0.2100D 01	114	0.2121D 01	0.2141D 01

NUMBER OF BITS: 8					
DISTORTION: 0.5208D-02 – 0.1927D-06			ENTROPY: 0.7815D 01		
	THRESHOLDS	LEVELS		THRESHOLDS	LEVELS
115	0.2164D 01	0.2186D 01	116	0.2210D 01	0.2234D 01
117	0.2259D 01	0.2285D 01	118	0.2313D 01	0.2340D 01
119	0.2371D 01	0.2401D 01	120	0.2434D 01	0.2467D 01
121	0.2504D 01	0.2541D 01	122	0.2582D 01	0.2623D 01
123	0.2670D 01	0.2717D 01	124	0.2771D 01	0.2826D 01
125	0.2891D 01	0.2956D 01	126	0.3037D 01	0.3119D 01
127	0.3229D 01	0.3340D 01	128	0.3519D 01	0.3699D 01

4.2. UNIFORM QUANTISER.

Uniform quantisers have been defined previously.

4.2.1. Again, for a symmetric input probability $p(x)$, the overall distortion D is expressed by:

$$D = 2\Big\{\sum_{i=1}^{\frac{n_l}{2}-1} \int_{(i-1).\Delta T}^{i.\Delta T} \rho[x, (i-\tfrac{1}{2}).\Delta T] \,.\, p(x) \,.\, dx + \int_{(\frac{n_l}{2}-1)\,\Delta T}^{\infty} \rho(x, \frac{n_l-1}{2}.\Delta T) \,.\, p(x) \,.\, dx\Big\} \tag{9}$$

The optimisation problem is much easier than previously because ΔT is the only parameter. It must satisfy: $\partial D/\partial(\Delta T) = 0$ which can be written for the absolute error distortion measure as follows:

$$\sum_{i=1}^{\frac{n_l}{2}-1} (i-\tfrac{1}{2}).\Big[\int_{(i-1).\Delta T}^{(i-1/2).\Delta T} p(x) \,.\, dx - \int_{(i-1/2).\Delta T}^{i.\Delta T} p(x) \,.\, dx\Big] + \frac{n_l-1}{2}\Big[\int_{(\frac{n_l}{2}-1)\,\Delta T}^{(\frac{n_l-1}{2})\,\Delta T} p(x) \,.\, dx - \int_{\frac{n_l-1}{2}\Delta T}^{\infty} p(x) \,.\, dx\Big] = 0 \tag{10}$$

4.2.2. Computation

A. We use the Newton-Raphson method.

Let Φ be equation (10).

The problem consists in finding $[\Delta T]^*$ such that

$\Phi([\Delta T]^*) = 0$

The Newton-Raphson iterations are defined by:

$$[\Delta T]^{k+1} = [\Delta T]^k - \frac{\Phi([\Delta T]^k)}{\Phi'([\Delta T]^k)}$$

The main characteristic of this method is its quadratic convergence, which means that, near the result $[\Delta T]^*$, the number of correct digits in $[\Delta T]^k$ doubles with each iteration. If this behaviour is not observed, we can be assured that the method has not been implemented correctly. Moreover, this offers us a stopping criterion; we stop the iteration process when the relative change in $[\Delta T]^k$ is negligible (less than 10^{-8}).

We again computed H, the quantiser output entropy, but its calculation does not appear in the program flow chart.

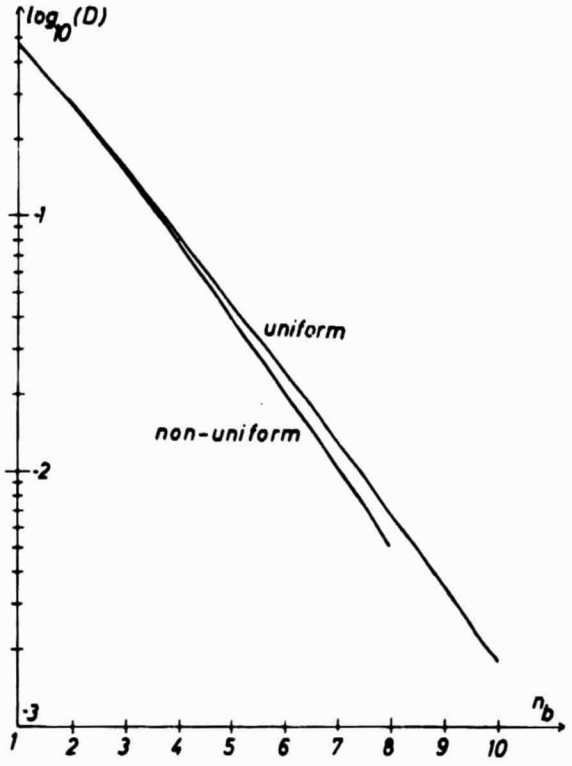

Figure 7. Distortion versus number of quantisation bits.

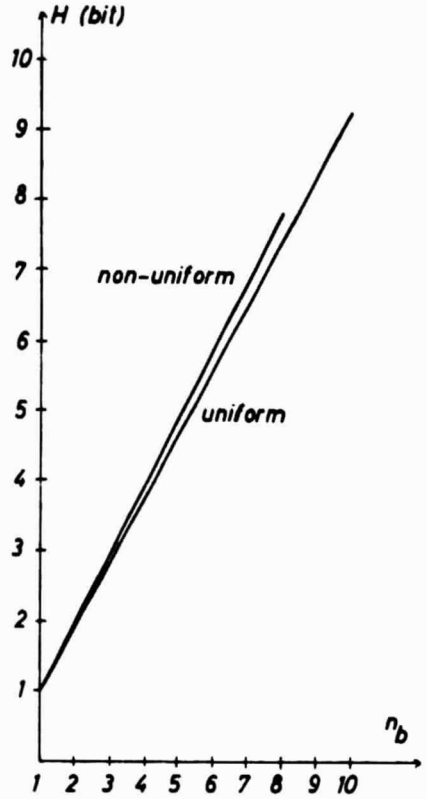

Figure 8. Output entropy versus number of quantisation bits.

TABLE 2.
Parameters of the Optimum Uniform Quantiser
Input Probability Density: N(0, 1),
Distortion Measure: $\rho(x, y) = |x - y|$

n_b	ΔT	D	H
1	0.1349D 01	0.4732D 00	0.9999
2	0.8350D 00	0.2669D 00	1.973
3	0.4875D 00	0.1474D 00	2.904
4	0.2759D 00	0.8042D-01	3.806
5	0.1531D 00	0.4350D-01	4.700
6	0.8368D-01	0.2335D-01	5.560
7	0.4524D-01	0.1245D-01	6.495
8	0.2424D-01	0.6600D-02	7.403
9	0.1289D-01	0.3480D-02	8.318
10	0.6814D-02	0.1826D-02	9.241

B. Flow chart of the program, see Figure 6.

C. Results.

The optimum ΔT and the corresponding D and H are given in Table 2 for $1 \leqslant n_b \leqslant 10$. We have plotted $\log(D)$ versus n_b in Figure 7 and H versus n_b in Figure 8.

5. COMPARISON OF NON-UNIFORM AND UNIFORM QUANTISER

For a given number of bits, we have:

1. The distortion of a uniform quantiser is slightly greater than that of the non-uniform quantiser (Figure 7).
2. The output entropy of the uniform quantiser is slightly smaller than that of the non-uniform quantiser (Figure 8).

In both cases, the output entropy is close to its maximum, $H_{\max} = n_b$ (bits), which means that the output levels are almost equiprobable.

In Figure 9 we have plotted the thresholds T_i versus their rank i for both types of the 4 –, 5 –, 6- and 7-bit half quantisers. For the non-uniform quantiser, $\Delta T_i = (T_{i+1} - T_i) \uparrow$ as $i \uparrow$ which explains why the probability of the output levels is quasi-uniform. In Figure 10, we have plotted the probabilities, Q_i, of the output levels L_i versus i for both types of the 7-bit half quantiser.

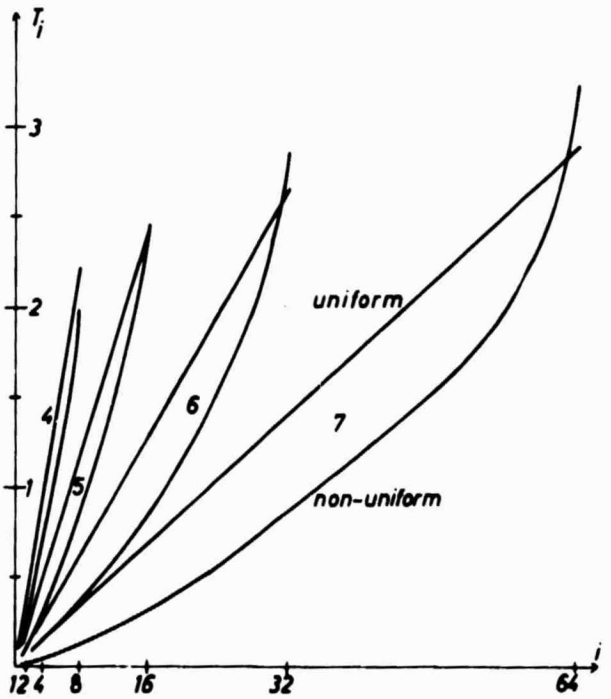

Figure 9. Thresholds versus rank.

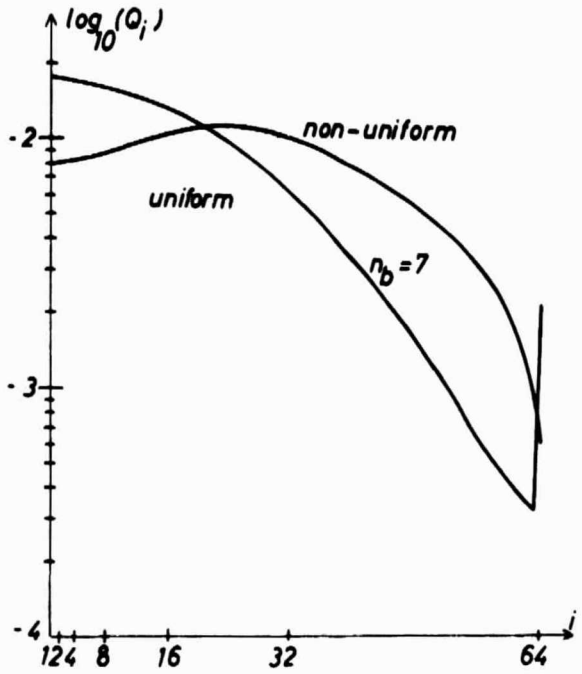

Figure 10. Output probabilities versus rank.

Although the non-uniform quantisers have slightly better performances (lower distortion, higher output entropy) than the uniform ones, the uniform quantisers are the only quantisers practically available: varying ΔT is essentially the same as varying the gain of a signal amplifier before quantisation.

REFERENCES

1. MAX, J.: 'Quantizing for minimum distortion'. *Trans. IRE.* IT-6, pp. 7–12, 1960.
2. LLOYD, S.P.: personal communication to T. Fine.
3. BERGMANS, P.: personal communication.
4. HASTINGS, C.: *Approximation for digital computers.* Princeton University Press, Princeton, New Jersey, 1955.

APPENDIX

APPENDIX

a) $P(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{\xi^2}{2}\right) . d\xi$

b) $x = P^{-1}(y)$. That is, P^{-1} computes the value x when $y = \text{Prob}(\xi \leqslant x) = P(x)$ is given.

These two functions are not available in the standard P.D.P. 11/40 software. We use modified numerical approximations[4].

a) $$P(x) = \begin{cases} \frac{1}{2}\left(1 - \text{erf}\left(-\frac{x}{\sqrt{2}}\right)\right), & x < 0 \\ \frac{1}{2}\left(1 + \text{erf}\left(\frac{x}{\sqrt{2}}\right)\right), & x \geqslant 0 \end{cases}$$

The error function itself is approximated numerically by:

$$\text{erf}^*(x) = 1 - \left(\sum_{i=1}^{5} a_i \eta^i\right) . \frac{2}{\sqrt{\pi}} \exp(-x^2)$$

where $\eta = \frac{1}{1 + cx}$ with

c	= 0.3275911	a_3	= 1.259695130
a_1	= 0.225836846	a_4	= −1.287822453
a_2	= −0.252128668	a_5	= 0.940646070

b) For $y \geqslant 0.5$ we have the following approximations:

$$[P^{-1}(y)]^* = \eta - \frac{\sum_{i=0}^{2} a_i \eta^i}{\sum_{i=0}^{3} b_i \eta^i} \quad \text{where } \eta = \sqrt{\log_e (1 - y)^{-2}} \text{ with}$$

a_0	= 2.515517	b_0	= 1.0
a_1	= 0.802853	b_1	= 1.432788
a_2	= 0.010328	b_2	= 0.189269
		b_3	= 0.001308