Reprinted from INFORMATION AND CONTROL, Volume 1, No. 2, May 1958
Academic Press Inc. *Printed in U.S.A.*

INFORMATION AND CONTROL **1,** 153–158 (1958)

On The Quantization of Finite Dimensional Messages[1]

MARCEL P. SCHÜTZENBERGER[2]

Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts

Let L be the average value of a measure of quantization noise, and let H be the negentropy of the quantized signal. Some reciprocal relationship exists between these quantities, since, for example, increasing the number of possible quantized values reduces L but increases H. We give a lower bound to L as a function of H and show that it may be realized up to a constant factor. Roughly speaking, this shows that every bit added to H multiplies L by a factor depending on the dimensionality of the message and the measure of quantization noise used.

I. INTRODUCTION

Let the message ξ be an n-dimensional continuous variate with *a priori* probability density $f(\xi)$. Before it can be transmitted through a discrete channel it has to be replaced by a quantized signal $[\xi]$, that is, by some approximate quantity taking only a finite number of distinct values.

We assume here that the channel is perfectly noiseless so that the only source of error lies in the quantization $\xi \to [\xi]$. The accuracy is usually measured by the average L of some given nondecreasing function $\ell(|\xi - [\xi]|)$ over the *a priori* distribution of ξ. We shall consider only those functions ℓ which are of the form $c|\xi - [\xi]|^{\alpha}(\alpha > 0)$ and our results will consequently cover the case of the so-called rms criterion $(\alpha = 2)$.

Shannon's theory of noiseless communication indicates that the natural measure of the cost of transmission is the negentropy H of the quantized signal $[\xi]$. With these conventions, the optimum is obtained when, for a given value of H (or of L), the other quantity is as small as possible. Intuitively, some general relationship must presumably exist between H and L, since any action which tends to decrease one of them (for in-

[1] This work was supported in part by the U. S. Army (Signal Corps), the U. S. Air Force (Office of Scientific Research, Air Research and Development Command), and the U. S. Navy (Office of Naval Research).

[2] Present address: Faculté des Sciences de Poitiers, France.

stance the multiplication of the number of different values of $[\xi]$) has exactly the opposite effect on the other.

Under some broad conditions we give here a lower bound to the value of L as a function of H and we show that this bound may be reached up to a constant proportionality factor. Loosely speaking, these two results mean that every bit of information allows on the average a reduction of L by a factor no more but not less than $2^{-\alpha/n}$, whatever be the density function $f(\xi)$. In particular, $L \geqq KN^{-\alpha/n}$ with K, a constant, for every quantization with N different quantized values $[\xi]$.

II. HYPOTHESES

We state first our hypotheses:

A. The message is an n-dimensional variate ($n < \infty$) admitting a continuous, bounded density $f(\xi)$ in its domain of variation $E = \{\xi : f(\xi) > 0\}$ Further,

$$\left| \int_E f(\xi) \log f(\xi)\, d\xi \right| < \infty.$$

B. There exists a finite θ for which

$$\int_E | \xi |^{\alpha+\theta} f(\xi)\, d\xi < \infty.$$

A quantization $\xi \to [\xi]$ will be identified with a partition $W = \{E_i\}$ of E; each E_i is the set of the ξ's admitting the same quantized value $[\xi] = a_i$. For any W we define:

$$H(W) = -\sum P_i \log P_i$$

where

$$P_i = \int_{E_i} f(\xi)\, d\xi$$

and

$$L(W) = \int_E c \,| \xi - [\xi] |^\alpha f(\xi)\, d\xi = \sum P_i L_i'$$

where

$$L_i' = P_i^{-1} \int_{E_i} c\, | \xi - [\xi] |^\alpha f(\xi)\, d\xi.$$

Finally, we say that a sequence of quantizations $W_1 = \{E_{1_i}\}$, $W_2 =$

$\{E_{2i}\}\ \cdots$, $W_j = \{E_{ji}\}\ \cdots$, is *systematically convergent* if, for all j, every $E_{j+1,i}$ is entirely contained in some $E_{ji'}$ and if $\overline{\lim}_{j\to\infty} L(W_j) = 0$.

First inequality. If $f(\xi)$ satisfies A, there exists a constant K with the property that $L(W) \geqq K\,(\exp\, -\alpha/n)\, H(W)$ for all possible quantizations of ξ.

Second inequality. If $f(\xi)$ satisfies A and B, there exists a constant K' and a systematically convergent sequence $\{W_j\}$ with the property that $L(W_j) \leqq K'\,(\exp\, -\alpha/n)\, H(W_j)$ for all j.

III. PROOF OF THE INEQUALITIES

In what follows $g_1, g_2, \cdots$ denote geometric constants which are functions of α and n only; $k_1, k_2, \cdots$ denote nonzero finite constants whose values depend upon $f(\xi)$ but not upon the quantization considered.

We shall use twice the fact that for any partition $W = \{E_i\}$ the sum $|-\sum P_i \log f_i|$, where f_i is the value of $f(\xi)$ at some inner point of E_i, is uniformly bounded. This results immediately from the hypotheses by the following inequalities

$$\left|\sum P_i \log 1/f_i\right| = \left|\sum P_i \log f^*/f_i - \sum P_i \log f^*\right|$$

$$\leqq \sum P_i \left|\log f^*/f_i\right| + \left|\log f^*\right| < \int_E f(\xi)\left|\log f^*/f(\xi)\right| d\xi + \left|\log f^*\right|$$

$$\leqq \int_E f(\xi) \log 1/f(\xi)\, d\xi + 2\left|\log f^*\right|$$

where

$$f^* = \sup_{\xi\epsilon E} f(\xi).$$

First Inequality

We take a fixed arbitrary number p $(0 < p < 1)$ and, for each E_i of W we define a value f_i and a subset E_i' of E_i by the relations:

$$E_i' = \{\xi\epsilon E_i\,;\, f(\xi) \geqq f_i\}; \qquad \int_{E_i'} f(\xi)\, d\xi = p\int_{E_i} f(\xi)\, d\xi = pP_i\,.$$

We have

$$P_i L_i' \geqq \inf_x \int_{E_i} c\,|\,\xi - x\,|^\alpha f(\xi)\, d\xi = c\int_{E_i} |\,\xi - x_i\,|^\alpha f(\xi)\, d\xi$$

$$\geqq c\int_{E_i'} |\,\xi - x_i\,|^\alpha f(\xi)\, d\xi \geqq cf_i \int_{E_i} |\,\xi - x_i\,|^\alpha\, d\xi = cf_i L_i''.$$

It is a classical result that for a fixed value of meas(E_i'), the sum L_i'' is a minimum when E_i' is an n-dimensional sphere with radius ρ_i centered at x_i. Consequently, $P_iL_i' \geqq cg_1f_i\rho_i^{n+\alpha}$ where ρ_i is defined by $\text{meas}(E_i') = g_2\rho_i^n$ and where g_1 and g_2 are geometric constants. If we now define $\bar{f}_i$ by the equality $\bar{f}_i \text{ meas}(E_i') = pP_i = \int_{E_i'} f(\xi)\, d\xi$, we can eliminate ρ_i and meas(E_i'). Thus we obtain

$$P_iL_i' \geqq P_i^{1+\alpha/n} cg_3 p^{1+\alpha/n} f_i \bar{f}_i^{-1-\alpha/n}.$$

Taking into account the remark made at the beginning of this section, we find that

$$L_i' \geqq cg_3(p/f^*)^{1+\alpha/n} p_i^{\alpha/n} f_i$$

and

$$- \sum P_i \log L_i' \leqq \left(\frac{\alpha}{n}\right) H(W) + \log k_1 - \sum P_i \log f_i$$

$$\leqq \left(\frac{\alpha}{n}\right) H(W) - \log K.$$

This concludes the proof, since we have

$$L(W) = \sum P_iL_i' \geqq \exp - \sum P_i \log L_i' \quad [= K \exp - \alpha/nH(W)]$$

because of the convexity of the function $\log 1/x$.

Second Inequality

The construction of a systematically convergent sequence can be carried out in many ways. We indicate here one method which is probably among the simplest ones. In the first place we observe that the classical inequality on the absolute moments

$$\left[\int_{E'} |\xi|^{\alpha} f(\xi)\, d\xi\right]^{1/\alpha} \leqq \left[\int_{E'} |\xi|^{\alpha+\theta} f(\xi)\, d\xi\right]^{(\alpha+\theta)^{-1}} \left[\int_{E'} f(\xi)\, d\xi\right]^{\theta(\alpha+\theta)^{-1}}$$

gives under the hypothesis B

$$\int_{E'} |\xi^{\alpha}| f(\xi)\, d\xi \leqq \left[\int_{E'} |\xi|^{\alpha+\theta} f(\xi)\, d\xi\right]^{\alpha(\alpha+\theta)^{-1}} \left[\int_{E'} f(\xi)\, d\xi\right]^{\theta(\alpha+\theta)^{-1}}$$

$$= k_2 \left[\int_{E'} f(\xi)\, d\xi\right]^{\theta(\alpha+\theta)^{-1}}$$

for any subset E' of E.

Let us take now an arbitrary length d and construct a connected domain F around the origin made up of the juxtaposition of n-dimensional cubes C_i, with d the length of the side of each cube. We can make F big enough so that

$$e = \int_{E-F} f(\xi)\, d\xi$$

satisfies the relation $e^{\theta/\alpha+\theta} \leqq d^\alpha$. We consider the quantization W in which, $[\xi] = x_i$, the center of C_i, when $\xi \epsilon C_i$ and $[\xi] = 0$ when $\xi \epsilon E - F$. We have

$$H(W) = -\sum P_i \log P_i - e \log e - P_i \log f_i + n \log 1/d - e \log e$$

(where, again, f_i is the value of $f(\xi)$ at some inner point of C_i), that is,

$$n \log 1/d \geqq H(W) - k_3 + e \log e.$$

Had we considered instead of F some domain F' for which

$$e' = \int_{E-F'} f(\xi)\, d\xi \leqq \int_{E-F} f(\xi)\, d\xi = e$$

the last inequality would still have been valid, for $x \log 1/x$ is a decreasing function of x. Consequently $d^n \leqq K'' \exp - H(W)$ for some K''. We compute now $L(W)$.

$$\sum P_i L_i' = c \int_{E-F} |\, \xi \,|^\alpha f(\xi)\, d\xi + \sum c \int_{C_i} |\, \xi - x_i \,|^\alpha f(\xi)\, d\xi$$

but, for any C_i:

$$\int_{C_i} |\, \xi - x_i \,|^\alpha f(\xi)\, d\xi \leqq \int_{C_i} \left| \sup_{\xi \epsilon C_i} |\, \xi - x_i \,| \right|^\alpha f(\xi)\, d\xi$$

$$\leqq g_4\, d^\alpha \int_{C_i} f(\xi)\, d\xi = g_4\, d^\alpha\, P_i$$

and

$$\int_{E-F} |\, \xi \,|^\alpha f(\xi)\, d\xi \leqq k_2 e^{\theta(\alpha+\theta)^{-1}} \leqq k_2\, d^\alpha.$$

Thus

$$L(W) = \sum P_i L_i' \leqq K''' d^\alpha \leqq K' \exp - \alpha/n H(W).$$

By construction the constant K' can be chosen such that it does not depend upon W. We consider now the partition $W = W_1$ as the first term of the sequence $\{W_j\}$ and we take a second value d' such that d is equal to some multiple of d'.

We subdivide every C_i into smaller cubes C_i' with length of the side d' and we add new cubes of the same size around F so as to obtain a domain F' for which, as above,

$$\int_{E-F'} f(\xi)\, d\xi \leqq d'^{(\alpha+\theta)\alpha\theta^{-1}}.$$

Obviously the partition $W_2 = W'$ satisfies $L(W') \leqq L(W)$; $L(W') \leqq K' \exp - \alpha/nH(W')$, and this concludes the proof since we can choose, by iterating the same method, a sequence d, d', $\cdots$ converging to zero.

IV. REMARKS

i. The hypotheses A and B are sufficient but obviously not necessary for the validity of the results. In the same manner, the assumption that the "loss function" $\ell(r)$, $(r = |\xi - [\xi]|)$, has the form cr^α could be weakened and the results would hold substantially, in an asymptotic fashion, for any $\ell(r)$ with $\lim_{r\to 0} rd/dr \log \ell(r) = \alpha > 0$. But this would definitely not be true for arbitrary $\ell(r)$ (as, for example, $\exp -1/r$ or $r \log 1/r$) and the normalization function $H(W)$ does not seem then to play the same natural role.

ii. A more detailed computation allows one to get closer estimates of the constants K and K'. However, they remain different and their ratio tends to infinity with n. For $n = 2$, a better "second inequality" can be obtained by use of a covering of the plane with hexagons instead of squares. Our present ignorance concerning the most elementary properties of the coverings of the space for $n \geqq 3$ seems to lie at the root of the discrepancy between K and K'.

RECEIVED: September 6, 1957.