COMMUNICATION SCIENCES

AND

ENGINEERING

XIII. STATISTICAL COMMUNICATION THEORY*

Prof. Y. W. Lee
Prof. A. G. Bose
Prof. M. Schetzen
Prof. H. L. Van Trees
Prof. V. R. Algazi
Prof. J. D. Bruce
Prof. A. V. Oppenheim
R. Alter
M. Austin
A. B. Baggeroer
R. F. Bauer
L. D. Collins
T. A. Froeschle
J. W. Giffin
L. M. Goodman
K. Grace, Jr.
T. Huang
D. E. Nelsen
T. H. Nyman
M. O. Pace
R. B. Parente
L. R. Poulo
J. E. Schindall
D. L. Snyder
J. J. Wawzonek
D. H. Wolaver

A. WORK COMPLETED

[Titles followed by a dagger (†) are theses that were supervised by members of this group, although the work was not sponsored by the Research Laboratory of Electronics. Summaries are included because they might be of interest to workers in this field.]

1. BIOELECTRIC CONTROL OF PROSTHESES

This study has been completed by R. Alter. In August 1965, he submitted the results to the Department of Electrical Engineering, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Doctor of Science. This study will appear as Technical Report 446 of the Research Laboratory of Electronics.

A. G. Bose

2. FUNCTIONAL ANALYSIS OF SYSTEMS CHARACTERIZED BY NONLINEAR DIFFERENTIAL EQUATIONS

This study has been completed by R. B. Parente. In August 1965, he submitted the results to the Department of Electrical Engineering, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Doctor of Philosophy. This study will appear as Technical Report 444 of the Research Laboratory of Electronics.

Y. W. Lee

3. OPTIMUM LAGUERRE EXPANSION OF SYMMETRIC N^{th}-ORDER FUNCTIONS

This study has been completed by J. W. Giffin. In August 1965, he submitted the results to the Department of Electrical Engineering, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Master of Science.

M. Schetzen

*This work was supported in part by the Joint Services Electronics Program (Contract DA36-039-AMC-03200(E)), the National Science Foundation (Grant GP-2495), and the National Aeronautics and Space Administration (Grants NsG-334 and NsG-496).

4. SOME PROBLEMS IN THE STUDY OF NONLINEAR SYSTEMS WITH FEEDBACK LOOPS†

This study has been completed by Cynthia L. K. Whitney. In August 1965, she submitted the results to the Department of Electrical Engineering, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Master of Science.

M. Schetzen

5. MEASUREMENT OF VOLTERRA KERNELS OF A NONLINEAR SYSTEM OF FINITE ORDER

This study has been completed by T. Huang. In August 1965, he submitted the results to the Department of Electrical Engineering, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Master of Science.

M. Schetzen

6. SEVERAL ADAPTIVE BINARY DETECTION PROBLEMS†

This study has been completed by D. W. Boyd. In August 1965, he submitted the results to the Department of Electrical Engineering, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Master of Science.

H. L. Van Trees

7. DIGITAL SIMULATION OF ANALOG MODULATION TECHNIQUES OVER THE RAYLEIGH CHANNEL†

This study has been completed by T. J. Cruise. In June 1965, he submitted the results to the Department of Electrical Engineering, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Master of Science.

H. L. Van Trees

8. SPACE-TIME SIGNAL PROCESSING

This study has been completed by K. Grace, Jr. In August 1965, he submitted the results to the Department of Electrical Engineering, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Master of Science.

H. L. Van Trees

9. ANALOG COMMUNICATION THROUGH SEPARABLE MULTIPATH CHANNELS CHARACTERIZED BY TIME-VARYING PATH DELAYS†

This study has been completed by R. R. Kurth. In August 1965, he submitted the results to the Department of Electrical Engineering, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Master of Science.

H. L. Van Trees

10. PREDISTORTION IN NO-MEMORY FILTERING AND IN QUANTIZATION

This study has been completed by M. O. Pace. In June 1965, he submitted the results to the Department of Electrical Engineering, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Master of Science and the degree of Electrical Engineer.

V. R. Algazi

11. A STATISTICAL STUDY OF VLF ATMOSPHERIC NOISE†

This study has been completed by R. A. Grant, Jr. In August 1965, he submitted the results to the Department of Electrical Engineering, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Master of Science.

V. R. Algazi

B. DETERMINATION OF OPTIMUM NONLINEAR SYSTEMS FOR GAUSSIAN INPUTS BY CROSSCORRELATION

1. Optimum Systems with White Gaussian Inputs

In the Wiener theory of nonlinear systems, the input, $x(t)$, of a system A, as shown in Fig. XIII-1, is a white Gaussian process. The output, $y_a(t)$, of the system is represented by the orthogonal expansion

$$y_a(t) = \sum_{n=0}^{\infty} G_n[h_n, x(t)] \tag{1}$$

in which $\{h_n\}$ is the set of Wiener kernels of the nonlinear system A, and $\{G_n\}$ is a complete set of orthogonal functionals. The orthogonal property of the functionals is expressed by the fact that the time average $\overline{G_m[h_n, x(t)]\, G_n[h_n, x(t)]} = 0$ for $m \neq n$. The power density spectrum of the Gaussian input, $x(t)$, is $\Phi_{xx}(\omega) = \frac{K}{2\pi}$ watts per radian per second so that the autocorrelator function of the input is $\phi_{xx}(\tau) = K\mu(\tau)$ where $\mu(\tau)$ is the unit impulse function.

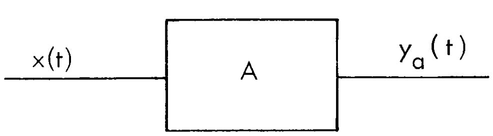

Fig. XIII-1. Nonlinear system with white Gaussian input.

If the desired output of the nonlinear system A is $z(t)$, the error, $\epsilon_a(t)$, is

$$\epsilon_a(t) = z(t) - y_a(t) \tag{2}$$

We shall show in this report that the Wiener kernels of the optimum nonlinear system A for which the mean-square error, $\overline{\epsilon_a^2(t)}$, is a minimum are given by

$$h_n(\sigma_1, \dots, \sigma_n) = \begin{cases} \dfrac{1}{n!K^n} \overline{z(t)\, x(t-\sigma_1) \dots x(t-\sigma_n)} & \sigma_i \geq 0 \quad i = 1, 2, \dots, n \quad n = 0, 1, 2, \dots \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

except when two or more σ's are equal.

To show this result, let us write the n^{th}-degree functional with $x(t-\sigma_1) \dots x(t-\sigma_n)$ as the leading term in an orthogonal set $\{H_n[k_n, x(t)]\}$ as

$$\left.\begin{aligned} &H_o[k_o, x(t)] = 1 \\ &H_n[k_n, x(t)] = \int \dots \int k_n(\tau_1, \dots, \tau_n)\, x(t-\tau_1) \dots x(t-\tau_n)\, d\tau_1 \dots d\tau_n + F \\ &\qquad n = 1, 2, 3, \dots \end{aligned}\right\} \tag{4}$$

in which F is a sum of homogeneous functionals of degrees lower than n and

$$k_n(\tau_1, \dots, \tau_n) = u(\tau_1-\sigma_1) \dots u(\tau_n-\sigma_n); \tag{5}$$

in which u(t) is the unit impulse function. It has been shown[1] that

$$\overline{y_a(t)\, H_n[k_n, x(t)]} = n!K^n h_n(\sigma_1, \dots, \sigma_n) \qquad n = 0, 1, 2, \dots \tag{6}$$

in which there are no restrictions on the σ's.

Now construct the system A with the Wiener kernels given by

$$h_n(\sigma_1, \dots, \sigma_n) = \begin{cases} \dfrac{1}{n!K^n} \overline{z(t)\, H_n[k_n, x(t)]}; & \sigma_i \geq 0 \quad n = 0, 1, 2, \dots \\ 0 & \text{any } \sigma_i < 0 \end{cases} \tag{7}$$

This system is the optimum nonlinear system. To show this, we first observe from (6) and (7) that by our construction A,

$$\overline{y_a(t)\, H_n[k_n, x(t)]} = \overline{z(t)\, H_n[k_n, x(t)]} \qquad \sigma_i \geq 0 \qquad n = 0, 1, 2, \dots \tag{8}$$

so that from (2)

$$\overline{\epsilon_a(t)\, H_n[k_n, x(t)]} = 0 \qquad \text{for } \sigma_i \geq 0 \qquad n = 0, 1, 2, \dots . \tag{9}$$

Equation 9 implies that

$$\overline{\epsilon_a(t)\, x(t-\sigma_1) \dots x(t-\sigma_n)} = 0 \qquad \text{for } \sigma_i \geq 0 \qquad n = 0, 1, 2, \dots \tag{10}$$

This is easily seen by induction from Eqs. 4 and 9, since

$$\overline{\epsilon_a(t)\, H_o[k_o, x(t)]} = \overline{\epsilon_a(t)} = 0 \tag{11}$$

$$\overline{\epsilon_a(t)\, H_1[k_1, x(t)]} = \overline{\epsilon_a(t)\, x(t-\sigma_1)} = 0; \qquad \sigma_1 \geq 0 \tag{12}$$

For $n = 2$ in Eq. 9

$$\begin{aligned} \epsilon_a(t)\, H_2[k_2, x(t)] &= \overline{\epsilon_a(t)\, x(t-\sigma_1)\, x(t-\sigma_2)} + \overline{\epsilon_a(t) F} \\ &= 0 \qquad \text{for } \sigma_1, \sigma_2 \geq 0 \end{aligned} \tag{13}$$

Since F in Eq. 13 is the sum of homogeneous functionals of degrees less than 2, we have from (11) and (12) that in Eq. 13, $\overline{\epsilon_a(t)\, F} = 0$ and thus

$$\overline{\epsilon_a(t)\, x(t-\sigma_1)\, x(t-\sigma_2)} = 0 \qquad \text{for } \sigma_1, \sigma_2 \geq 0. \tag{14}$$

By continuing in this manner, the validity of Eq. 10 for any value of n can be established. We note that Eq. 10 implies that the average of the product of $\epsilon_a(t)$ with any realizable functional of x(t) is zero, since

$$\begin{aligned} &\overline{\epsilon_a(t) \int_0^\infty \cdots \int_0^\infty g_n(\tau_1, \ldots, \tau_n)\, x(t-\tau_1) \ldots x(t-\tau_n)\, d\tau_1 \ldots d\tau_n} \\ &\quad = \int_0^\infty \cdots \int_0^\infty g_n(\tau_1, \ldots, \tau_n)\, \overline{\epsilon_a(t)\, x(t-\tau_1) \ldots x(t-\tau_n)}\, d\tau_1 \ldots d\tau_n = 0. \end{aligned} \tag{15}$$

By use of this result, we can now show that no other nonlinear system of the Wiener class can have a mean-square error smaller than $\overline{\epsilon_a^2(t)}$ so that the System A is the optimum system. To show this, we consider another system B with the output $y_b(t)$ for the input x(t). Let $\{g_n\}$ be the set of Wiener kernels of System B so that

$$y_b(t) = \sum_{n=0}^{\infty} G_n[g_n, x(t)]. \tag{16}$$

The error, $\epsilon_b(t)$, obtained when using System B is

$$\begin{aligned} \epsilon_b(t) &= z(t) - y_b(t) \\ &= z(t) - y_a(t) + y_a(t) - y_b(t) \\ &= \epsilon_a(t) + y_a(t) - y_b(t). \end{aligned} \tag{17}$$

The mean-square error thus can be written

$$\overline{\epsilon_b^2(t)} = \overline{\epsilon_a^2(t)} + \overline{[y_a(t)-y_b(t)]^2} + \overline{2\epsilon_a(t)\,[y_a(t)-y_b(t)]}. \tag{18}$$

Now

$$y_a(t) - y_b(t) = \sum_{n=0}^{\infty} G_n[h_n-g_n, x(t)]. \tag{19}$$

Thus by use of Eq. 15, the last term of Eq. 18 is zero and

$$\overline{\epsilon_b^2(t)} = \overline{\epsilon_a^2(t)} + \overline{[y_a(t)-y_b(t)]^2}. \tag{20}$$

From Eq. 20, $\overline{\epsilon_b^2(t)}$ is a minimum if $y_b(t) = y_a(t)$ which implies that System B is identical with System A. Thus no other system can have a mean-square error smaller than $\overline{\epsilon_a^2(t)}$ and System A with the Wiener kernels given by Eq. 7 is the optimum system. If no two σ's are equal, it can be shown that F in Eq. 4 is zero so that Eq. 3 follows from Eq. 7 and our result is proved.

To develop a procedure of measurement that is valid for all values of the σ's, we need not construct the functional F in Eq. 4. The restriction in (3) on the equality of the σ's arises from the presence of G-functionals of order lower than n which produce an impulse when two or more σ's are equal.[1] For example, the restriction in the determination of h_2 is due to G_o which produces an impulse when $\sigma_1 = \sigma_2$. From (3), $G_o[h_o, x(t)] = \overline{z(t)}$. Hence if we subtract G_o from the desired output, we have

$$h_2(\sigma_1, \sigma_2) = \frac{1}{2!K^2} \overline{\{z(t)-G_o[k_o, x(t)]\}\, x(t-\sigma_1)\, x(t-\sigma_2)}$$

$$\sigma_1, \sigma_2 \geq 0 \quad \text{including } \sigma_1 = \sigma_2. \tag{21}$$

In general, when we determine the n^{th}-order kernel, all of the lower order kernels have been determined so that all of the G-functionals of order less than n could be formed. For the determination of h_n, instead of (3) we would have the unrestricted expression

$$h_n(\sigma_1, \dots, \sigma_n) = \frac{1}{n!K^n} \overline{\left\{z(t) - \sum_{m=0}^{n-1} G_m[h_m, x(t)]\right\} x(t-\sigma_1) \dots x(t-\sigma_n)}$$

$$\sigma_i \geq 0 \qquad i = 1, 2, \dots, n \tag{22}$$

in which there is no restriction on the equality of the σ's

2. Optimum Systems with Non-White Gaussian Inputs

The theory that has been presented can be generalized to the case for a non-white Gaussian process. Consider that the optimum nonlinear system N shown in Fig. XIII-2 is to be determined for a desired output z(t) and an input, v(t), which is a non-white Gaussian process for which the power density spectrum is factorable.[2] It can be written

$$\Phi_{vv}(\omega) = \Phi_{vv}^{+}(\omega)\ \Phi_{vv}^{-}(\omega) \tag{23}$$

in which $\Phi_{vv}^{+}(\omega)$ is the complex conjugate of $\Phi_{vv}^{-}(\omega)$; also all of the poles and zeros of $\Phi_{vv}^{+}(\omega)$ are in the left-half of the complex s-plane in which $s = \sigma + j\omega$. Thus $\Phi_{vv}^{+}(\omega)$ and $\frac{1}{\Phi_{vv}^{+}(\omega)}$ are each realizable as the transfer function of a linear system. We then can

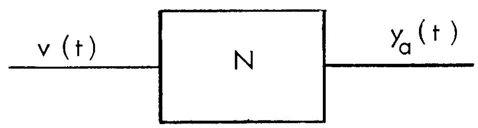

Fig. XIII-2. Nonlinear system with non-white Gaussian input.

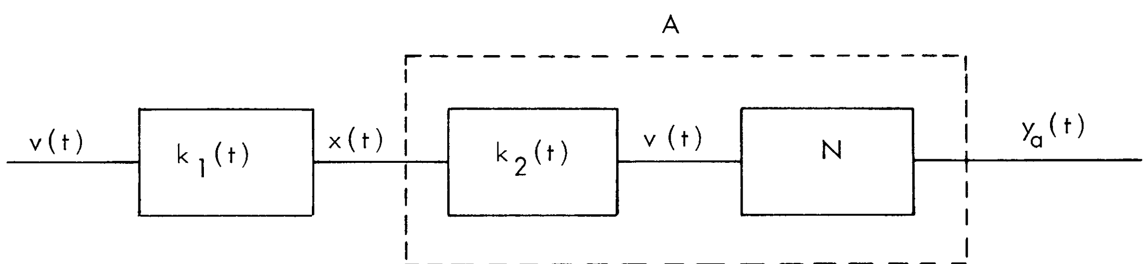

Fig. XIII-3. Equivalent form of the nonlinear system N.

consider the system of Fig. XIII-2 in the equivalent form shown in Fig. XIII-3 in which the transfer functions of the two linear systems $k_1(t)$ and $k_2(t)$ are

$$K_1(\omega) = \frac{1}{\Phi_{vv}^{+}(\omega)} \tag{24}$$

and

$$K_2(\omega) = \Phi_{vv}^{+}(\omega) \tag{25}$$

Also as shown, the system A is the system formed by the tandem connection of the linear system $k_2(t)$ and the system N. We observe that x(t), the input to the system A, is a white Gaussian process whose power density spectrum is 1 watt per radian per second;

the output of the system A is $y_a(t)$. The Wiener kernels of the optimum nonlinear system A for which the mean-square error, $\overline{\epsilon_a^2(t)} = \overline{[z(t)-y_a(t)]^2}$ is a minimum, according to Eq. 3, are

$$h_n(\sigma_1, \ldots, \sigma_n) = \frac{(2\pi)^n}{n} \overline{z(t)\, x(t-\sigma_1) \ldots x(t-\sigma_n)} \qquad \text{for } \sigma_i \geq 0 \qquad \begin{array}{l} i = 1, 2, \ldots, n \\ n = 0, 1, 2, \ldots \end{array} \tag{26}$$

except when two or more σ's are equal. The resulting system N is the optimum nonlinear system. To show this, we first observe that

$$\overline{\epsilon_a(t)\, v(t-\sigma_1) \ldots v(t-\sigma_n)} = 0 \qquad \text{for } \sigma_i \geq 0 \qquad \begin{array}{l} i = 1, 2, \ldots, n \\ n = 0, 1, 2, \ldots \end{array} \tag{27}$$

This result is obtained by substituting the relation

$$v(t) = \int_0^\infty k_2(\sigma)\, x(t-\sigma)\, d\sigma \tag{28}$$

in Eq. 27 and making use of Eq. 10. By the use of Eq. 27 and an argument identical with that given for a white Gaussian process, it is easy to show that no other system can have a mean-square error smaller than $\overline{\epsilon_a^2(t)}$ so that the system N is the optimum system.

The desired crosscorrelation function of Eq. 26 can be expressed in terms of only v(t) by substituting the relation

$$x(t) = \int_0^\infty k_1(\sigma)\, v(t-\sigma)\, d\sigma \tag{29}$$

in Eq. 26. The result is

$$\overline{z(t)\, x(t-\sigma_1) \ldots x(t-\sigma_n)} = \int_0^\infty k_1(\tau_1)\, d\tau_1 \ldots \int_0^\infty k_1(\tau_n)\, d\tau_n\, \overline{z(t)\, v(t-\sigma_1-\tau_1) \ldots v(t-\sigma_n-\tau_n)}. \tag{30}$$

other forms and interpretations for this crosscorrelation function have been given elsewhere.[3]

M. Schetzen

References

1. Y. W. Lee and M. Schetzen, "Measurement of the Kernels of a Nonlinear System by Crosscorrelation," Quarterly Progress Report No. 60, Research Laboratory of Electronics, M.I.T., January 15, 1961, pp. 126-129.

2. Y. W. Lee, Statistical Theory of Communication (John Wiley and Sons, Inc., New York, 1960), Chap. 14, Sec. 8.

3. M. Schetzen, "Measurement of the Kernels of a Nonlinear System by Crosscorrelation with Gaussian Non-White Inputs," Quarterly Progress Report No. 63, Research Laboratory of Electronics, M.I.T., October 15, 1961, pp. 113-117; also see Errata, Quarterly Progress Report No. 64, Research Laboratory of Electronics, M.I.T., January 15, 1962, p. 163.

C. USEFUL APPROXIMATIONS TO OPTIMUM QUANTIZATION

The quantization of random signals has been considered by Max[1] and, more recently, by Bruce.[2] Max considers the selection of an optimum step size in a uniform quantizer and also the determination of the optimum nonuniform quantizer and carries out computations for a Gaussian input and a mean-square distortion measure. Bruce gives a computer algorithm for the determination of the optimum nonuniform quantizer for an arbitrary distortion measure. These exact approaches to quantization give little insight and confront someone who has a new quantization problem with a considerable amount of digital computation. Schteyn[3] and Roe[4] have proposed approximations to the optimum nonuniform quantizer which are of practical interest. In this report we present useful approximations to the optimum quantizer and to the resulting distortion in uniform and nonuniform quantization for arbitrary distortion measures.

1. Uniform Quantization

The equation for the quantization step that minimizes the distortion has been given by Max[1] and can be solved on a computer by an iterative procedure. Here we obtain a simple solution by using the following facts.

1. The first-order effect will be due to the truncation of the tail of the distribution.
2. Except for the tails, the probability density of the signal can be simply approximated between successive quantization steps.

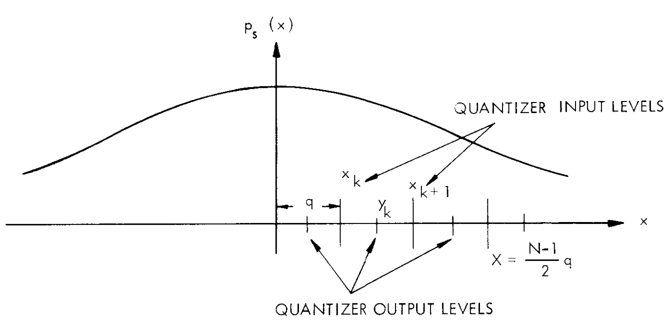

Fig. XIII-4. Uniform quantizer.

Consider Fig. XIII-4 which illustrates the problem. It is clear that a difficulty in selecting the step size will occur whenever the signal probability density has a long tail. Let D_k be the conditional distortion whenever the input signal, x, is between the input quantizer levels of x_k and x_{k+1}, and let the probability of this event be p_k. Let W(e) be an even error-weighting function and $W_1(e)$, its antiderivative. Then D_k is given approximately by

$$D_k = \frac{2W_1\left[\frac{q}{2}\right]}{q} \tag{1}$$

in which q is the size of the uniform step. This result is easily obtained by assuming that the $p_s(x)$ is well approximated by a straight line between x_k and x_{k+1}. Since D_k is independent of k, we have

$$D = \sum_k p_k D_k + p_T D_T \tag{2}$$

in which D_T is the distortion in the tails. Let p_T be the probability of the tails. We have then

$$\sum_k p_k = 1 - p_T$$

from which we obtain

$$D = \frac{2W_1\left[\frac{q}{2}\right]}{q}\left[1 - \int_{-\infty}^{\frac{N-1}{2}q} p_s(x)\,dx - \int_{\frac{N-1}{2}q}^{\infty} p_s(x)\,dx\right] + \int_{-\infty}^{-\frac{N-1}{2}q} W\left[x + \frac{N-1}{2}q\right] p_s(x)\,dx + \int_{\frac{N-1}{2}q}^{\infty} W\left[x - \frac{N-1}{2}q\right] p_s(x)\,dx \tag{3}$$

when $p_s(x)$ is even we have the simpler expression

$$D = 2W_1\left[\frac{q}{2}\right]\left[1 - 2\int_{\frac{N-1}{2}q}^{\infty} p_s(x)\,dx\right] + 2\int_{\frac{N-1}{2}q}^{\infty} W\left[x - \frac{N-1}{2}q\right] p_s(x)\,dx. \tag{4}$$

Equation 4 is an approximate expression for the distortion as a function of the step size and has to be minimized by proper choice of q. We could formally set the derivative equal to zero, but it is generally quite simple to obtain D as a function q and get an idea of the sensitivity of the distortion to the step size. In Fig. XIII-5 we give as an illustration the distortion versus the quantization range $x = \frac{N-1}{2}q$ for a Gaussian probability density, a mean-square distortion measure, and 8 quantization steps. We observe that the curve has a well-marked minimum and that the proper choice of step size is definitely worthwhile. Equation 4 gives

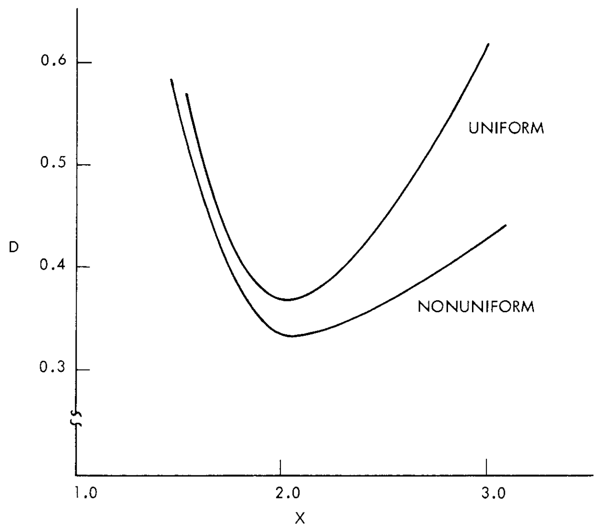

Fig. XIII-5. Distortion vs quantization range.

minimum distortions in good agreement (within 0.2 db) with the exact values given by Max for a Gaussian probability density.

2. Nonuniform Quantization

We consider the nonuniform quantizer to be the cascade of two nonlinear devices and of a uniform quantizer as shown in Fig. XIII-6. For a given uniform quantizer and a given signal probability density $p_s(x)$ the two nonlinear devices f( ) and g( ) are chosen

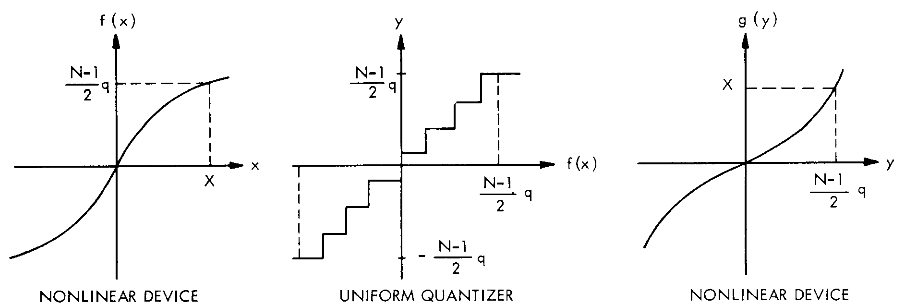

Fig. XIII-6. Nonuniform quantizer.

so as to minimize the distortion $D = E\{W[s-g(y)]\}$. The distortion D is made up of two parts: (i) distortion D_n obtained within the range of the uniform quantizer, and (ii) distortion in the tails D_T for which the nonlinear devices f( ) and g( ) are completely ineffective.

The distortion D_n can be conveniently discussed in terms of an analog model in which the uniform quantizer is replaced by additive noise. We obtain an expression for the minimum D_n as a function of X such that $f(X) = \frac{N-1}{2} q$. The optimum selection of X is then done as for uniform quantization by trading off between D_n and D_T.

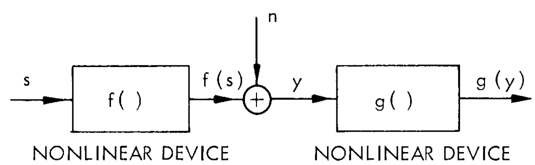

Fig. XIII-7. Analog model.

Consider the analog system shown in Fig. XIII-7. The additive, statistically independent noise, n, models the effect of the uniform quantizer. Note that Widrow[5] has shown that the quantization noise was independent of the signal for a large number of quantization levels. Here again the nonlinear devices f( ) and g( ) are chosen so as to minimize the distortion.

$$D_n = E\{W[s-g(y)]\}.$$

To keep the problem relevant to quantization we shall assume that the noise is small, and it can then be shown that f( ) and g( ) are inverses.

We can now solve easily the analog filtering problem. We write

$$D_n = \iint W\{g[f(x)+\beta]-x\}\, p_s(x)\, p_n(\beta)\, dx d\beta. \tag{5}$$

For small noise we have

$$g[f(x)+\beta] \approx g[f(x)] + \beta g'[f(x)].$$

But since f( ) and g( ) are inverses

$$g[f(x)] = x \qquad g'[f(x)] = \frac{1}{f'(x)}$$

and we have

$$D_n = \iint W\left[\frac{\beta}{f'(x)}\right] p_s(x)\, p_n(\beta)\, dx d\beta. \tag{6}$$

To proceed with an arbitrary error-weighting function W(e) we have to model the noise as uniform from $-q/2$ to $q/2$; again, this is a reasonable model quantization noise.[5] Then we carry out the integral with respect to β and use the calculus of variations to minimize D_n by proper choice of f( ). More specific results can be obtained whenever

$$W\left[\frac{\beta}{f'(x)}\right] = W[\beta]\, W\left[\frac{1}{f'(x)}\right] \tag{7}$$

which is the class of error-weighting functions $W(e) = |e|^c$. We have then

$$D_n = \int |\beta|^c \, p_n(\beta) \, d\beta \int \frac{p_s(x)}{[f'(x)]^c} \, dx. \tag{8}$$

It can be shown by the calculus of variations that D_n is minimized by taking

$$f(x) = K_1 \int [p_s(x)]^{\frac{1}{c+1}} \, dx + K_2 \tag{9}$$

which corresponds to Roe's expression for the approximate quantizer. The constants K_1 and K_2 have to be chosen to give the best fit for the quantization problem. Roe examines the behavior of the exact quantizer for a large number of levels and large inputs (mean-square error) and determines accordingly the approximate quantizer. Pace[6] selected the constant so as to match the exact two-level quantizer. We proceed here as for uniform quantization and obtain an approximate expression for the resulting distortion. Note that under the approximation that the probability density at the input of the uniform quantizer is piecewise linear, we have again

$$\int |\beta|^c \, p_n(\beta) \, d\beta = \frac{2W_1\left[\frac{q}{2}\right]}{q} = \frac{q^c}{2^c(c+1)}.$$

Now we select K_1 and K_2 in Eq. 9 to give the total range of the uniform quantizer. From Fig. XIII-5 we have

$$f(x_2) - f(x_1) = (N-1)q$$

from which we get

$$f'(x) = \frac{(N-1) \, q[p_s(x)]^{\frac{1}{c+1}}}{\int_{x_1}^{x_2} [p_s(x)]^{\frac{1}{c+1}} \, dx} \tag{10}$$

and by substituting Eq. 10 in Eq. 8 we obtain

$$D_n = \frac{1}{2^c(c+1)(N-1)^c} \left[\int_{x_1}^{x_2} [p_s(x)]^{\frac{1}{c+1}} \, dx \right]^{c+1}. \tag{11}$$

If we take into account the probability of occurrence of D_n and D_T, we have

$$D = D_n\left[1 - \int_{-\infty}^{x_1} p_s(x)\, dx - \int_{x_2}^{\infty} p_s(x)\, dx\right] + \int_{-\infty}^{x_1} |x-x_1|^c\, p_s(x)\, dx$$
$$- \int_{x_2}^{\infty} |x-x_2|^c\, p_s(x)\, dx \tag{12}$$

For an even signal probability density we have $x_2 = -x_1 = X$, and

$$D = \frac{2}{(c+1)(N-1)^c}\left[\int_0^X [p_s(x)]^{\frac{1}{c+1}}\, dx\right]^{c+1}\left[1 - 2\int_X^{\infty} p_s(x)\, dx\right]$$
$$+ 2\int_X^{\infty} |x-X|^c\, p_s(x)\, dx \tag{13}$$

which has to be minimized by proper choice of X. Once X is determined, we have

$$f(x) = \frac{(N-1)q \int_0^x [p_s(a)]^{\frac{1}{c+1}}\, da}{2\int_0^X [p_s(a)]^{\frac{1}{c+1}}\, da}. \tag{14}$$

Note that the step size of the uniform quantizer will not affect the resulting nonuniform quantizer.

As an example we consider a Gaussian signal, mean-square error and 8 quantization steps. The distortion versus X is shown in Fig. XIII-5, and its minimum is in good agreement with the value given by Max.

3. Uniform Versus Nonuniform Quantization

By comparing the distortion for uniform and nonuniform quantization in Fig. XIII-5, it appears that there is little to be gained by nonuniform quantization and the added complexity in equipment. To discuss this point more generally we consider again Eqs. 4 and 13. If we rewrite Eq. 4 for $W(e) = |e|^c$ and take $X = \frac{N-1}{2} q$ we have for uniform quantization

$$D_u = \frac{X^c}{(N-1)(c+1)}\left[1 - 2\int_X^{\infty} p_s(x)\, dx\right] + 2\int_X^{\infty} |x-X|^c\, p_s(x)\, dx. \tag{15}$$

By comparing Eqs. 15 and 13, we see that the two expressions are quite similar except for the two factors

$$F_u \triangleq X^c$$

$$F_{nu} \triangleq 2\left[\int_0^X [p_s(x)]^{\frac{1}{c+1}} dx\right]^{c+1}.$$

The factors F_u and F_{nu} affect the distortion within the quantization range. It can be shown that F_{nu} is maximized by a signal probability density $p_s(x)$ that is uniform between $-X$ and X. We then have $F_{nu} = X^c$ as expected. An indication of the effect of nonuniform quantization can be obtained by forming the ratio $F(X) = F_{nu}/F_u$ of the quantization errors within the quantization range. For a Gaussian $p_s(x)$ and large X we obtain

$$F(X) = \frac{[2\pi\sigma^2(c+1)]^{\frac{c+1}{2}}}{X^c}.$$

Since F(X) goes to zero as X goes to infinity, we obtain a large improvement in distortion by nonuniform quantization as N goes to infinity. Note, however, that $D_{nu}/D_u = 0.7$ for N = 36 and c = 2; therefore, the asymptotic behavior is not too meaningful here.

V. R. Algazi

References

1. J. Max, "Quantizing for Minimum Distortion," IRE Trans., Vol. IT-6, No. 1, pp. 7-12, March 1960.
2. J. D. Bruce, "Optimum Quantization," Technical Report 429, Research Laboratory of Electronics, M.I.T., Cambridge, Massachusetts, March 1, 1965.
3. V. M. Schteyn, "On Group Transmission with Frequency Division of Channels by the Pulse Code Modulation Method," Telecommunications (Russian) 1959, pp. 178-180.
4. G. M. Roe, "Quantizing for Minimum Distortion," IEEE Trans., Vol. IT-10 (Letter), October 1964.
5. B. Widrow, "A Study of Rough Amplitude Quantization by Means of Nyquist Sampling Theory," Sc.D. Thesis, Department of Electrical Engineering, Massachusetts Institute of Technology, 1956.
6. M. O. Pace, "Predistortion in No-Memory Filtering and in Quantization," E.E. Thesis Massachusetts Institute of Technology, June 1965.

D. TIME JITTER IN TUNNEL DIODE THRESHOLD-CROSSING DETECTORS

In Quarterly Progress Report No. 78 a model was presented which describes the time jitter arising in a tunnel diode threshold-crossing detector.[1] Analysis of the model by dimensional methods indicated close agreement with experimental observations. Since then a more complete and exact analysis of this model has been performed with the help of the IBM 7094 computer. It has been predicted that the jitter should have a Gaussian distribution. The mean and standard deviation of this distribution is related to the shot noise that is present near the peak of the tunnel diode i-v relation, to the slope of the input signal, and to other circuit parameters. This predicted behavior agrees closely with experimental measurements made thus far.

In this report we shall outline the method used in analyzing the model, discuss the derived results, and present comparisons of these results with experimental observations.

1. Model for Predicting Switching Jitter

The model[1] used for predicting the switching behavior is shown in Fig. XIII-8. The circuitry to the right of the input current source is a commonly accepted model for the

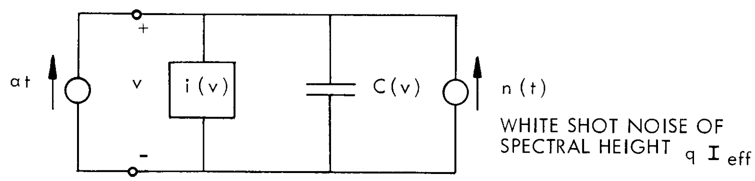

Fig. XIII-8. Equivalent-circuit model for tunnel-diode threshold detector.

tunnel diode.[2,3] Lead inductance and resistance are neglected. We consider i(v), the familiar static tunnel diode i-v curve, to be an instantaneous relation between i and v. The total capacitance across the junction, C(v), is considered to be constant and equal to C in the vicinity of the current peak. I_{eff} is the effective shot noise current, and in the region near the peak is approximately equal to the actual tunnel diode current.

Using Kirchhoff's current law, we may write for the network of Fig. XIII-8

$$C \frac{dv}{dt} + i(v) = \alpha t + n(t). \tag{1}$$

If we translate our coordinate system so that its origin is at the peak of i(v) and fit a parabola to the curve at that point, we obtain a new i-v relation

$$i(v) = -kv^2, \tag{2}$$

where k is a measure of the curvature at the tunnel diode peak, and v and t are understood to be new variables. Substituting this new relation in (1), we obtain

$$C \frac{dv}{dt} - kv^2 = \alpha t, \tag{3}$$

where n(t) is white noise of spectral height, N_o. It is from this "switching" equation that the statistics of the jitter are derived.

2. Analysis of the Switching Equation (3)

By suitably grouping the parameters a, k, and C, we obtain the new dimensionless variables

$$v' = \left(\frac{k^2}{Ca}\right)^{1/3} v$$

$$t' = \left(\frac{ka}{C^2}\right)^{1/3} t \tag{4}$$

$$N_o' = \frac{k}{C^2 a} N_o.$$

By substituting these variables in Eq. 3, we obtain the dimensionless equation

$$\frac{dv'}{dt'} - v'^2 = t' + n'(t'), \tag{5}$$

where $n'(t')$ is white noise of spectral height N_o'.

A property of this equation is that if the right side is negative, then $v'(t')$ will tend to some stable finite value. If the right side becomes positive, then $v'(t')$ will grow until it reaches infinity at some finite time. We shall consider this to be the time at which the tunnel diode switches and denote it by the variable T_s'.

For a given set of initial conditions and $n'(t') = 0$ the system will always "switch" at the same time, T_s. When noise is added, however, T_s becomes a random variable, taking on values distributed around some mean. We shall call this distribution of T_s',

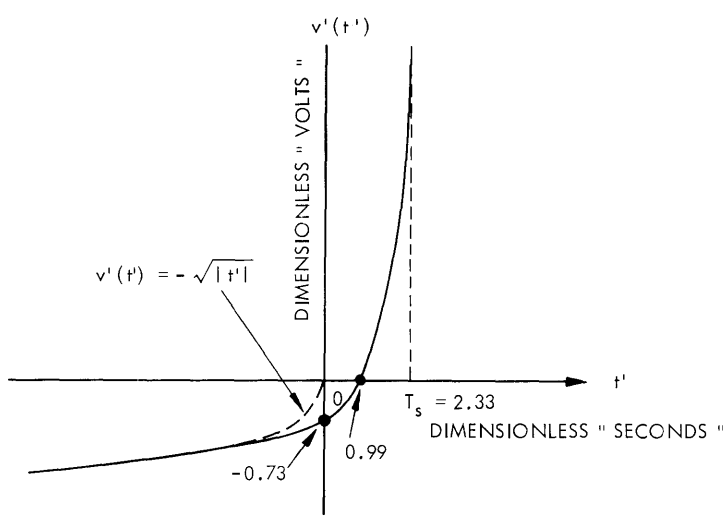

Fig. XIII-9. Solution of the dimensionless switching equation with $n'(t') = 0$.

$P_{T_s'}(\tau'; N_o')$, where τ' is a range variable for T_s'. The distribution will depend on the noise spectral height N_o' and also on the initial conditions.

We shall henceforth take the initial conditions to be those for which the system would be in "equilibrium" at a large negative time t_o:

$$v_o' = -\sqrt{|t_o'|} . \tag{6}$$

These initial conditions ensure that the operating point essentially follows the curve $i'(v')$ until the region of switching is reached. The solution for these conditions and no noise is shown in Fig. XIII-9.

With noise present the standard deviation and mean of T_s' will be functions of N_o'. We shall denote these functions by $\sigma_T'(N_o')$ and $\overline{T_s'}(N_o')$. The computation of these functions will be described.

3. Computer Solution of the Switching Equation

The computation was performed using the Fortran language on the IBM 7094 computer. The dimensionless switching Eq. 5 was solved by using standard one-step difference methods. Noise was introduced by adding a random number at each iteration of the difference equation. The random number sequence was obtained by using the "RANNOF" routine. This routine generates a pseudo-random sequence of numbers that are uniformly distributed from 0 to 1 and which, for our purposes, can be considered to be mutually independent. The sequence was then adjusted to have zero mean and to have a variance corresponding to a given spectral height N_o'.

The computation was started far enough back in time, subject to the initial conditions of (6), to ensure that the process would appear to have been going indefinitely. When $v'(t')$ became large enough to ensure that the noise would have negligible effect on the future course of the signal, the computation was stopped and the final values of v' and t' were substituted in an asymptotic form of the solution that is valid for large v'. From this asymptotic form the switching time T_s could be obtained.

N_o' was set to some specified value and this solution procedure was carried out 1000 times, resulting in that many values of the random variable T_s'. By means of standard computing techniques, the mean, standard deviation, and distribution of T_s' were calculated. These statistics were obtained for N_o' varying over a range that was slightly wider than that covered in the actual measurements.

4. Results of the Computation

Graphs of $\sigma_T'(N_o')$ and $\overline{T_s'}(N_o')$ are presented in Figs. XIII-10 and XIII-11. The fact that $\sigma_T'(N_o')$ varies linearly with $\sqrt{N_o'}$ indicates that the result obtained previously[1] by assuming this linearity is valid at least over the range of the present computation. It

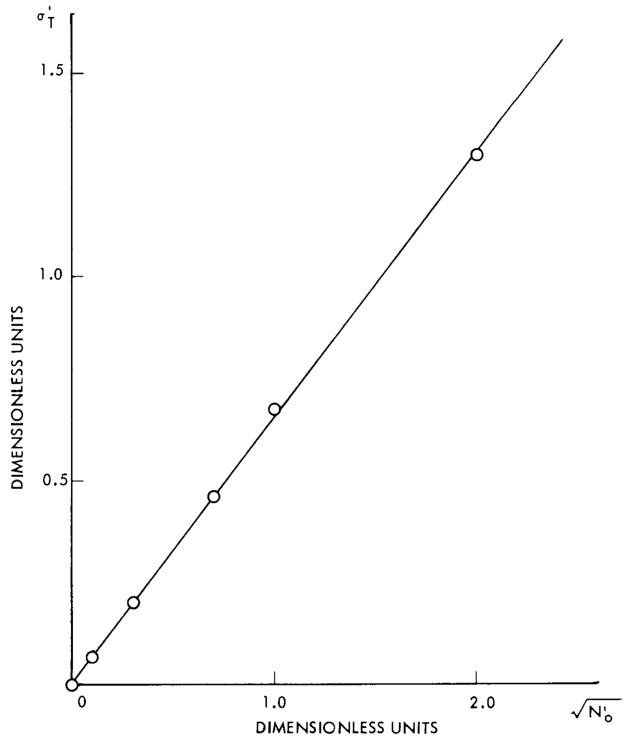

Fig. XIII-10. Jitter standard deviation σ'_T vs $\sqrt{N'_o}$ (in dimensionless units).

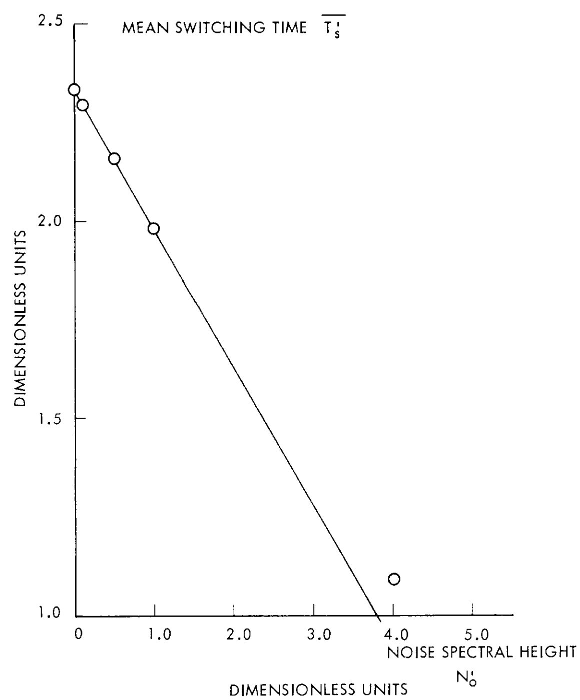

Fig. XIII-11. Mean switching time $\overline{T'_s}$ vs N'_o (in dimensionless units).

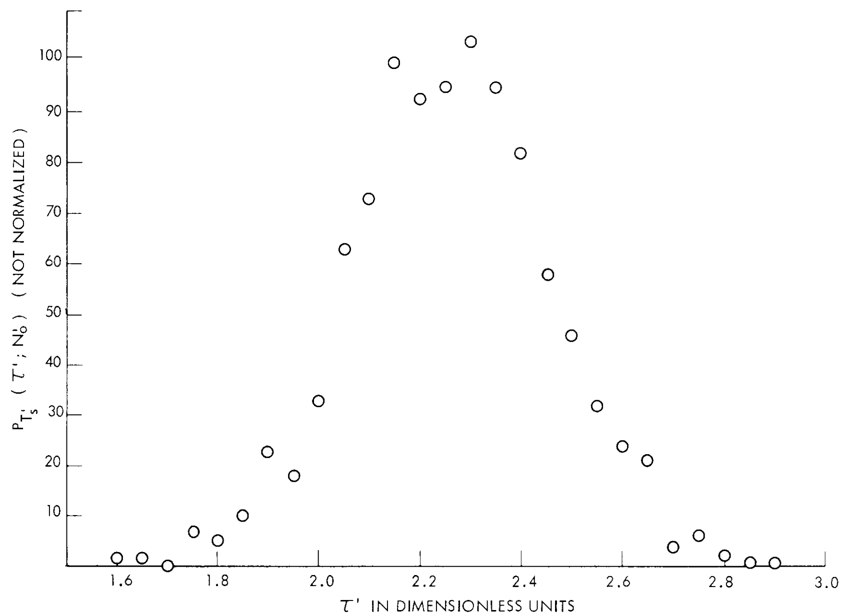

Fig. XIII-12. Un-normalized form of $P_{T'_s}$ vs $(\tau';N'_o)$ for $N'_o = 0.1$.

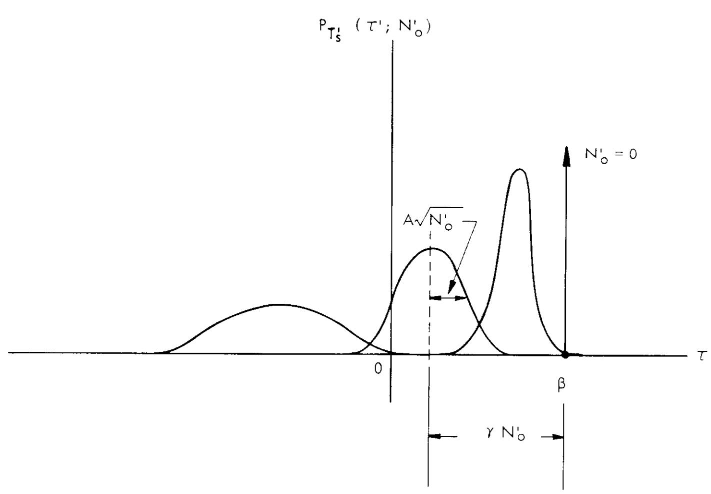

Fig. XIII-13. Gaussian-shaped distribution function moves toward the left and grows wider as noise N_o is increased.

is interesting that, on the average, switching occurs earlier as noise is increased. The advance in the mean switching time is directly proportional to N_o'.

It appears that $T_s'(N_o')$ is beginning to deviate from its linear behavior at large values of N_o'. Since the required computation time increases with $\sqrt{N_o'}$, we have hesitated to push the range further until we are certain that an analytic solution could not be obtained by some other method. Knowledge of how this linear behavior breaks down (if it does) would be useful for a complete description of the statistical behavior of the jitter.

Within the accuracy of the statistics, the distributions obtained were Gaussian. A typical distribution is shown in Fig. XIII-12.

The functions $\sigma_T'(N_o')$ and $\overline{T_s'}(N_o')$ can be approximated analytically from Figs. XIII-10 and XIII-11 by

$$\sigma_T'(N_o') = A\sqrt{N_o'}$$

and (7)

$$\overline{T_s'}(N_o') = \beta - \gamma N_o,$$

where A, β, and γ are constants.

Using these forms, we can express the jitter distribution analytically as

$$P_{T_s'}(\tau'; N_o') = \frac{1}{\sqrt{2\pi N_o'}\, A} \exp - \frac{1}{2}\left[\frac{(\tau'-\beta+\gamma N_o')^2}{A^2 N_o'}\right]. \quad (8)$$

The behavior of $P_{T_s'}(\tau'; N_o')$ with increasing noise is shown in Fig. XIII-13. It is interesting to note that $P_{T_s'}(\tau'; N_o')$ satisfies the equation for diffusion in a moving medium. Thus far we have not been able to relate this diffusion equation to the switching process in any fundamental manner.

5. Transformation of $\sigma_T'(N_o')$ and $\overline{T_s'}(N_o')$ back into the Dimensional Domain

Using the relations of Eq. 4, we can transform $\sigma_T'(N_o')$ and $\overline{T_s'}(N_o')$ back into the dimensional domain and thus relate the jitter statistics to the circuit parameters a, C, k, and N_o. Performing this operation, we obtain

$$\sigma_T = \frac{A N_o^{1/2} k^{1/6}}{a^{5/6} C^{1/3}} \quad (9)$$

and

$$\overline{T_s} = \beta\left(\frac{C^2}{ka}\right)^{1/3} - \gamma\left(\frac{k^{2/3} N_o}{C^{4/3} a^{4/3}}\right). \quad (10)$$

The distribution of the jitter can be obtained as a function of the circuit parameters by substituting σ_T and $\overline{T_s}$ in the Gaussian distribution

$$p(\tau) = \frac{1}{\sqrt{2\pi}\,\sigma_T} \exp\left[-\frac{1}{2}\frac{(\tau-\overline{T_s})^2}{\sigma_T^2}\right]. \tag{11}$$

6. Comparison of Results with Experimental Observations

Using the value of A determined from Fig. XIII-11 and values of N_o, C, and k corresponding to those existing in a circuit studied experimentally, we plot the relation between σ_T and a expressed by Eq. 9. Experimentally measured points superimposed on this curve are shown in Fig. XIII-14.

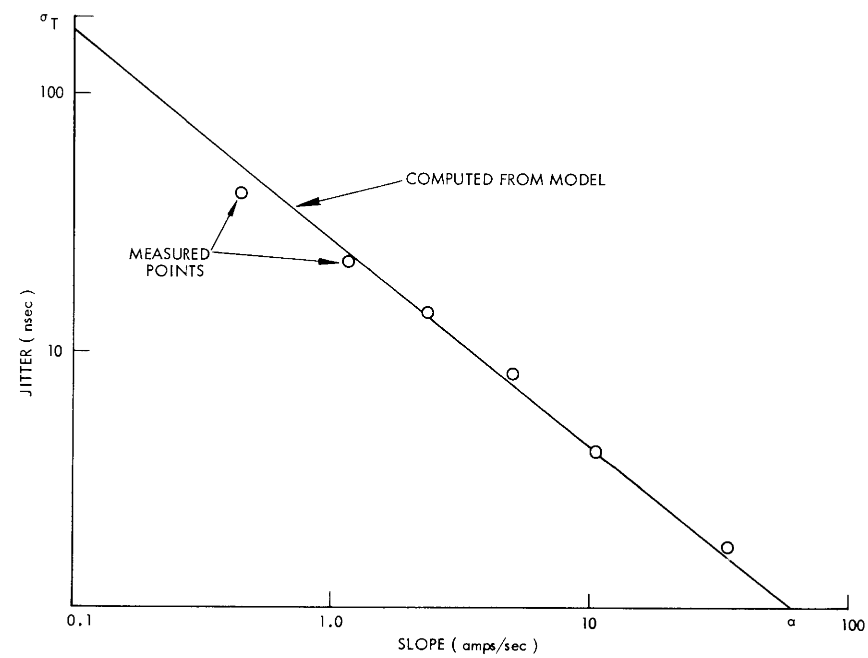

Fig. XIII-14. Jitter standard deviation σ_T vs slope a: a comparison between behavior predicted by model and experimentally observed behavior.

The experimentally measured distributions were Gaussian.[1] This observation is in agreement with the results obtained from the model.

We have not yet checked the validity of (10) experimentally. Since a differential method was used for measuring the switching jitter, all information concerning the mean T_s was cancelled out.

This work was done partly at the Computation Center of the Massachusetts Institute of Technology, Cambridge, Massachusetts.

D. E. Nelsen

References

1. D. E. Nelsen, "Time Jitter in Tunnel Diode Threshold-Crossing Detectors," Quarterly Progress Report No. 78, Research Laboratory of Electronics, M. I. T., July 15, 1965, pp. 178-188.
2. R. A. Pucel, "The Equivalent Noise Current of Esaki Diodes," Proc. IRE 49, 49 (1961).
3. B. E. Turner and R. E. Burgess, "Direct Tunnel-Current Noise in Tunnel Diodes for Small Biases," Can. J. Phys. 42, 1046 (June 1964).

E. NONLINEAR MINIMUM-MEAN SQUARE FILTERING WITH APPLICATION TO ANALOG COMMUNICATION

The purpose of this report is to briefly describe an approach to nonlinear, minimum-mean-square filtering and estimation and to mention some applications of the approach which have been made to analog communications through randomly time-varying channels. The approach differs significantly from the linear, minimum-mean-square approach of Wiener[1] and of Kalman and Bucy[2] because the estimate is not restricted to being a linear transformation of the observed process. On the other hand, the approach bears a resemblance to the technique of Kalman and Bucy, rather than to the more classic technique of Wiener, because of the use of the state variable representation of random processes. This representation is used because it is the most convenient way to represent continuous Markov processes, upon which the approach is theoretically based, and also because it allows the consideration of multilevel estimation problems without any added theoretical or manipulative difficulties.

The results presented here are an extension of those of Snyder[3] for the scalar case. The procedure closely parallels that used for the simpler case.

1. Notation

Underscored, lower-case letters denote column vectors and capital letters denote matrices. Supperscript "τ" and "-1" denote transpose and inverse. The exact and approximate minimum-mean-square estimates of $\underline{x}(t)$ are denoted by $\hat{\underline{x}}_{mv}(t)$ and $\underline{x}^*_{mv}(t)$, respectively.

$D[\underline{f}(t:\underline{x})]$ denotes the Jacobian associated with any vector $\underline{f}[t:\underline{x}(t)]$ whose components are memoryless transformations of $\underline{x}(t)$. The i-row, j-column element of $D[\underline{f}(t:\underline{x})]$ is $\frac{\partial}{\partial x_i} f_j(t:\underline{x})$.

2. Estimation Model and Equation for $\underline{x}^*_{mv}(t)$

Define the two vector processes, $\underline{x}(t)$ and $\underline{y}(t)$, by

$$d\underline{x}(t) = F\,\underline{x}(t)\,dt + d\underline{\chi}(t) \tag{1}$$

$$d\underline{y}(t) = \underline{g}[t:\underline{x}(t)]\,dt + d\underline{\eta}(t) \tag{2}$$

where the components of $\underline{\chi}(t)$ and $\underline{\eta}(t)$ are Wiener processes and

$$E[\underline{\chi}(t)\,\underline{\chi}^T(t)] = X \min(t,u) \tag{3}$$

$$E[\underline{\eta}(t)\,\underline{\eta}^T(t)] = N \min(t,u). \tag{4}$$

Here, $\underline{\chi}(t)$ and $\underline{\eta}(t)$ are assumed to be independent, $\underline{g}[t:\underline{x}(t)]$ represents a memoryless transformation of $\underline{x}(t)$. As defined by (1) and (2), $\underline{x}(t)$ and $\underline{y}(t)$ jointly form a continuous vector Markov process.

It is assumed that the observed process, $\underline{r}(t) = \frac{d}{dt}\underline{y}(t)$, is available from an initial time, t_0, until the present time, t. The observed waveform is denoted by $\underline{r}_{0,t}$.

Given $\underline{r}_{0,t}$ we seek to determine $\hat{\underline{x}}_{mv}(t)$. An equation for $\hat{\underline{x}}_{mv}(t)$ can be obtained in a straightforward way from the equation for the conditional probability density functional, $p(\underline{x};t|\underline{r}_{0,t})$, correctly derived by Kushner.[4] Using the fact that $\hat{\underline{x}}_{mv}(t)$ is the conditional mean, the result is

$$d\hat{\underline{x}}_{mv}(t) = F\,\hat{\underline{x}}_{mv}(t) + E[\{\underline{x}-\hat{\underline{x}}_{mv}(t)\}\underline{g}^T(t:\underline{x})]\,N^{-1}[d\underline{y}(t)-E\,\underline{g}(t:\underline{x})\,dt], \tag{5}$$

where the expectations are with respect to $p(\underline{x};t|\underline{r}_{0,t})$. An alternative expression for $\hat{\underline{x}}_{mv}(t)$ can be obtained by substituting the multidimensional Taylor expansion for $\underline{g}(t:\underline{x})$, which is assumed to exist, in (5). The resulting expression cannot be solved nor readily implemented. By assuming, however, that the error, $\underline{x} - \hat{\underline{x}}_{mv}(t)$, is small, an approximate estimate, $\underline{x}^*_{mv}(t)$, can be obtained. Keeping terms leading to the second moment of the error, we obtain

$$\frac{d}{dt}\underline{x}^*_{mv}(t) = F\,\underline{x}^*_{mv}(t) + VD\left[\underline{g}\left(t:\underline{x}^*_{mv}\right)\right]N^{-1}\{\underline{r}(t)-\underline{g}[t:\underline{x}^*_{mv}(t)]\}, \tag{6}$$

where V = V(t) is an error-covariance matrix satisfying

$$\frac{d}{dt}V(t) = FV + VF^T + X + VD\left[D\left[g\left(t:\underline{x}^*_{mv}\right)\right]N^{-1}\left\{\underline{r}(t)-\underline{g}\left(t:\underline{x}^*_{mv}\right)\right\}\right]V. \tag{7}$$

Under steady-state conditions (7) reduces to

$$0 = FV + VF^T + X - \overline{VD\left[\underline{g}\left(t:\underline{x}^*_{mv}\right)\right]N^{-1}D^T\left[\underline{g}\left(t:\underline{x}^*_{mv}\right)\right]V} \tag{8}$$

in which the bar indicates time or ensemble averaging which are assumed to be equivalent.

3. Communication Model

The communication model is shown in Fig. XIII-15. $\underline{a}(t)$ and $\underline{b}(t)$ are defined by

$$d\underline{a}(t) = F_a\, \underline{a}(t)\, dt + d\underline{\alpha}(t) \tag{9}$$

$$d\underline{b}(t) = F_b\, \underline{b}(t)\, dt + d\underline{\beta}(t), \tag{10}$$

where the components of $\underline{\alpha}(t)$ and $\underline{\beta}(t)$ are Wiener processes with the associated covariance matrices A min (t, u) and B min (t, u). $\underline{a}(t)$ and $\underline{b}(t)$ can represent one or more Gaussian processes occurring as messages and channel disturbances.

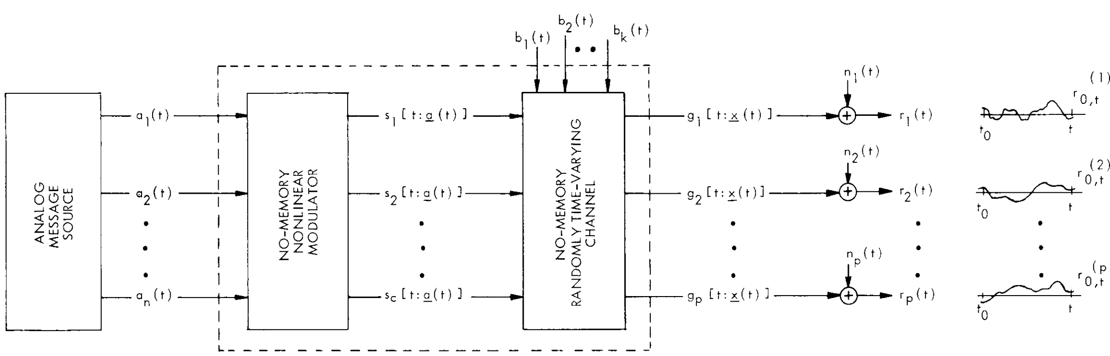

Fig. XIII-15. Communication model.

$\underline{a}(t)$ is transformed by a memoryless, nonlinear modulator into c signals represented by $\underline{s}[t{:}\underline{a}(t)]$. By a suitable interpretation of $\underline{a}(t)$, modulation schemes with memory, such as FM, fall within the scope of the model.

$\underline{s}[t{:}\underline{a}(t)]$ is transformed into p signals by the "randomly time-varying" portion of the channel. The resulting signals are represented by $\underline{g}[t{:}\underline{x}(t)]$ and are observed in additive white Gaussian noise.

The relationship between the estimation model and the communication model is evident when it is noted that $\underline{x}(t)$ represents the vector obtained by adjoining $\underline{a}(t)$ and $\underline{b}(t)$. $\underline{x}^*_{mv}(t)$ is then a vector whose elements are the approximate minimum-mean-square estimates of the message vector and channel-disturbance vector.

4. Examples

When $\underline{g}[t{:}\underline{x}(t)]$ is a linear transformation of x(t), the exact and approximate estimates are equal, and (6) and (7) reduce to the equations of Kalman and Bucy.[2] Communication

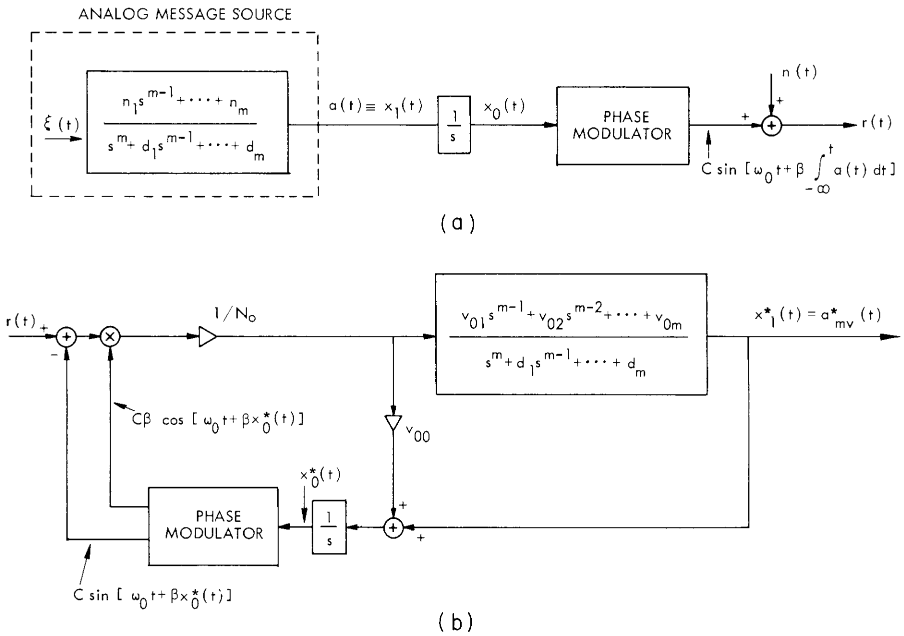

Fig. XIII-16. (a) FM communication model.
(b) Approximate minimum-mean-square FM demodulator.

models with linear modulation schemes, such as suppressed-carrier AM and single-sideband AM, fall within this case and can be easily treated.

For the case when $\underline{g}[t:\underline{x}(t)]$ is a nonlinear transformation of $\underline{x}(t)$, several examples have been studied and reported by Snyder.[5] For brevity, we give only the result for frequency modulation and cite other examples of interest which have been discussed elsewhere.[5] These are listed here.

1. Single message, general modulation, additive channel.
2. Single message, phase modulation, additive channel.
3. Single message, frequency modulation, additive channel.
   (The communication model and resulting demodulator for this case are shown in Fig. III-16. The v_{ij} occurring in the demodulator are the components of V in the steady-state and $\xi(t)$ is a white Gaussian process.)
4. Single message, general modulation, c diversity or multipath channels.
5. Single message, phase modulation, c diversity or multipath channels.
   (For this case, the demodulator structure is in the form of a maximal ratio combiner followed by a phase-locked loop.)
6. Single message, phase-modulation, simple multiplicative channel.
   (In this instance, the demodulator is in the form of a joint message and channel estimator.)
7. Single message, phase modulation, Rayleigh channel.

8. Single message, phase modulation, random phase channel (oscillator instability).
9. m messages, PM_m/PM, additive channel.

5. Conclusion

An approach has been outlined for nonlinear, minimum-mean-square filtering. The resulting fibers (or demodulators) bear a close relation to the demodulators obtained by the maximum a posteriori estimation procedure described, for example, by Van Trees.[6] The minimum-mean-square demodulators are identical to the realizable portion of the cascade realization of the maximum a posteriori demodulators. The communication model discussed in this report was used by Van Trees[7] who studied it with the alternative approach.

D. L. Snyder

References

1. N. Wiener, "The Extrapolation, Interpolation, and Smoothing of Stationary Time Series," John Wiley and Sons, Inc., New York, 1949.
2. R. E. Kalman and R. Bucy, "New Results in Linear Filtering and Prediction Theory," ASME J. Basic Eng. 83, 95-108 (March 1961).
3. D. L. Snyder, "An Application of an Equation for the Conditional Probability Density Functional of Markov Processes to Nonlinear Minimum Variance Filtering and Estimation," Quarterly Progress Report No. 78, Research Laboratory of Electronics, M. I. T., July 15, 1965, pp. 192-201.
4. H. J. Kushner, "On the Differential Equations Satisfied by Conditional Probability Densities of Markov Processes, with Applications," J. SIAM Control, Ser. A, Vol. 2, No. 1, pp. 106-119, 1964.
5. D. L. Snyder, "An approach to nonlinear minimum-variance filtering with application to analog communication via randomly-time-varying channels," Internal Memorandum No. IM-DS-5, M. I. T., August 1965.
6. H. L. Van Trees, "Analog Modulation and Continuous Estimation Theory," Class notes for Course 6. 576, Massachusetts Institute of Technology, Cambridge, Mass., 1965 (printed).
7. H. L. Van Trees, "Analog Communication over Randomly-Time-Varying Channels," 1964 WESCON Convention Record, Part 4, August 1964.