COMMUNICATION SCIENCES AND ENGINEERING

XVI. STATISTICAL COMMUNICATION THEORY*

Prof. Y. W. Lee	R. F. Bauer	P. L. Konop
Prof. A. G. Bose	E. M. Bregstone	A. J. Kramer
Prof. D. J. Sakrison	J. D. Bruce	D. E. Nelsen
Prof. M. Schetzen	A. M. Bush	J. K. Omura
Prof. H. L. Van Trees, Jr.	J. K. Clemens	A. V. Oppenheim
V. R. Algazi	A. G. Gann	R. B. Parente
R. Alter	C. E. Gray	W. S. Smith
D. S. Arnstein	W. F. Kelly	D. W. Steele
M. E. Austin	T. G. Kincaid	W. S. Widnall

RESEARCH OBJECTIVES

This group is interested in a variety of problems in statistical communication theory. Our current research is concerned primarily with the following problems:

1. A simple two-state modulation system has been designed and is under study. We are making a theoretical investigation of the linearity of the system and of its noise performance. The system will be applied to the problems of high-efficiency power amplification and regulation, multiplication, and low-frequency tape recordings.

2. The study of factors that influence the recording and reproduction of sound continues. The relative roles of the normal modes of the reproducing room and of the transducer are under investigation.

3. In the Wiener theory of nonlinear systems, a nonlinear system is characterized by a set of kernels. A method for the determination of these kernels was reported in Quarterly Progress Report No. 60 (pages 118-130). Work on the method, both theoretical and experimental, is being continued.

4. The central idea in the Wiener theory of nonlinear systems is to represent the output of a system by a series of orthogonal functionals with the input of the system being a white Gaussian process. An attempt is being made to extend the orthogonal representation to other types of inputs that may have advantages in the practical application of the theory.

5. A study is being made of the various aspects of error in filtering when the noise is additive and statistically independent of the signal. Emphasis is placed on nonlinear no-memory filters.

6. Noise sources in space can be located by means of higher order correlation functions. A study is being made of the errors, caused by finite observation time, incurred in locating sources by this method.

7. Many physical processes can be phenomenologically described in terms of a large number of interacting oscillators. An experimental investigation is being made of some theoretical results that have been obtained.

8. A nonlinear system can be characterized by a set of kernels. The synthesis of a nonlinear system involves the synthesis of these kernels. A study is being made of efficient methods for synthesizing these kernels.

9. The advantages of pseudo-noise carrier systems are well known. One disadvantage is the large bandwidth requirement. This disadvantage can be compensated for in multiplex systems by using a large number of carriers in the same frequency band. If one uses a simple demodulation scheme for a particular channel, the other carriers act as additive noise. By using a more sophisticated demodulation scheme, this interference can be reduced. An experimental system is being constructed.

---

*This work was supported in part by the National Institutes of Health (Grant MH-04737-02).

10. A reasonably simple adaptive, coherent, binary communication system has been developed and analyzed. Theoretical results indicate that even in channels with reasonably fast fading there is an appreciable improvement over a conventional incoherent system.

11. Experimental work on threshold effects in phase-locked loop discriminators continues.

12. Interesting results pertaining to the analysis of nonlinear, randomly time-variant systems have been obtained. Work in this area continues and some experimental verification on typical systems is planned.

Y. W. Lee

A. THE SYNTHESIS OF A CLASS OF NONLINEAR SYSTEMS

In the Wiener theory of nonlinear systems,[1] a nonlinear system is characterized by a set of kernels, h_n. For a Gaussian input, the kernels can be determined by crosscorrelation.[2,3] If only N kernels are used to characterize the nonlinear system, then the resulting representation can be expressed in the form of a Volterra series

$$y(t) = \sum_{n=1}^{N} \int_0^\infty \cdots \int_0^\infty k_n(\tau_1, \ldots, \tau_n)\, x(t-\tau_1) \ldots x(t-\tau_n)\, d\tau_1 \ldots d\tau_n \tag{1}$$

in which y(t) is the response of the nonlinear system for the input x(t). If the system is realizable, then

$$k_n(\tau_1, \ldots, \tau_n) = 0 \quad \text{for any } \tau_j < 0,\ j = 1, 2, \ldots, n. \tag{2}$$

A problem in the practical application of these results is the synthesis of systems whose kernels are the kernels, k_n, of the Volterra series. If an n^{th}-order kernel of a system is separable, so that it can be written in the form

$$k_n(\tau_1, \ldots, \tau_n) = k_{a_1}(\tau_1)\, k_{a_2}(\tau_2) \ldots k_{a_n}(\tau_n), \tag{3}$$

then the system can be synthesized in the form depicted in Fig. XVI-1. The system is one in which the outputs of the n linear systems with impulse responses $k_{a_j}(t)$ are

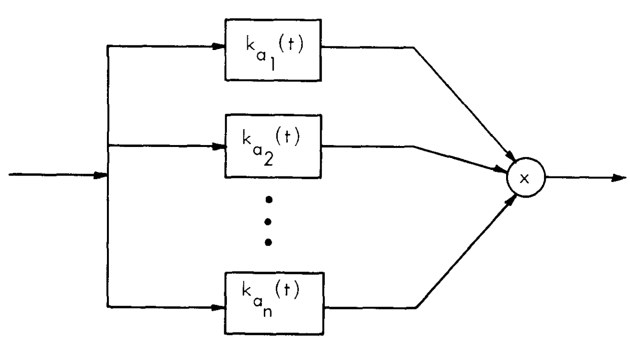


Fig. XVI-1. Synthesis of an n^{th}-order separable kernel.

multiplied together. However, in general, a kernel is not separable. A method suggested by Wiener is to expand the kernels in terms of a set of orthogonal functions such as the Laguerre functions.[1] Thus, if $\{\ell_n(t)\}$ is the set of Laguerre functions, then

$$k_1(\tau_1) = \sum_{n=0}^{\infty} c_n \, \ell_n(\tau_1)$$

$$k_2(\tau_1,\tau_2) = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} c_{n_1,n_2} \, \ell_{n_1}(\tau_1) \, \ell_{n_2}(\tau_2) \tag{4}$$

$$\vdots$$

$$k_N(\tau_1,\dots,\tau_N) = \sum_{n_1=0}^{\infty} \dots \sum_{n_N=0}^{\infty} c_{n_1,\dots,n_N} \, \ell_{n_1}(\tau_1) \dots \ell_{n_N}(\tau_N),$$

in which

$$c_{n_1,\dots,n_p} = \int_0^{\infty} \dots \int_0^{\infty} k_p(\tau_1,\dots,\tau_p) \, \ell_{n_1}(\tau_1) \dots \ell_{n_p}(\tau_p) \, d\tau_1 \dots d\tau_p. \tag{5}$$

The synthesis of a first- and a second-order kernel in terms of Eq. 4 is depicted in Fig. XVI-2. However, the synthesis of a given system by this procedure will, in general, require an infinite number of multipliers. In this report, we shall present a method of determining whether a system can be synthesized by using only a finite number of multipliers. We also shall present a method for the synthesis of such systems. To explain the method, the analysis of a system with only a second-order kernel will be presented. The extension to systems with higher-order kernels will then be given.

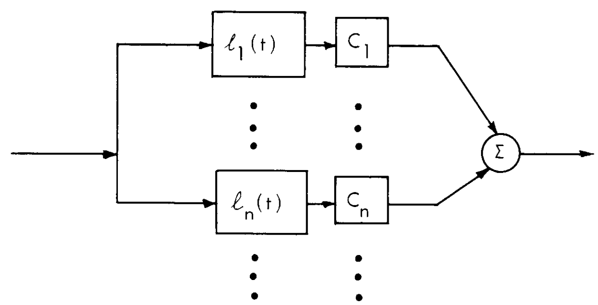
1. Second-order Systems with Only One Multiplier

We shall first discuss the class of second-order kernels that can be synthesized by means of only one multiplier. A system consisting of only one multiplier whose kernel is clearly the most general second-order kernel that can be obtained is depicted in Fig. XVI-3. The second-order kernel of this system is

$$k_2(\tau_1,\tau_2) = \int_0^{\infty} k_a(\tau_1-\sigma) \, k_b(\tau_2-\sigma) \, k_c(\sigma) \, d\sigma. \tag{6}$$

The kernel transform of this system is the two-dimensional Laplace transform of its kernel:

$$K_2(s_1,s_2) = \int_0^{\infty} \int_0^{\infty} k_2(\tau_1,\tau_2) \, e^{-s_1\tau_1-s_2\tau_2} \, d\tau_1 \, d\tau_2 = K_a(s_1) \, K_b(s_2) \, K_c(s_1+s_2), \tag{7}$$



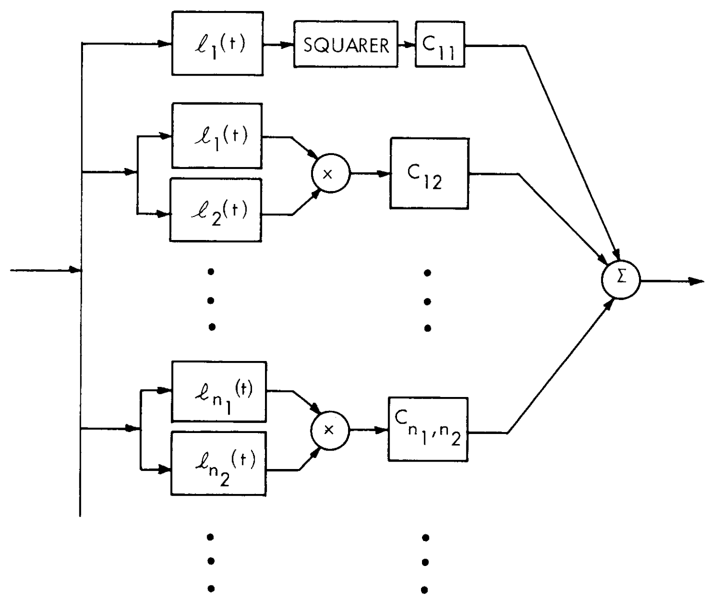


Fig. XVI-2. (a) Synthesis of a first-order kernel by means of Laguerre functions. (b) Synthesis of a second-order kernel by means of Laguerre functions.

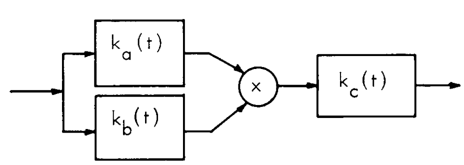


Fig. XVI-3. The most general second-order kernel with only one multiplier.

in which

$$K(s) = \int_0^\infty k(t)\, e^{-st}\, dt, \tag{8}$$

so that $K_a(s)$, $K_b(s)$, and $K_c(s)$ are the transfer functions of the linear systems with impulse responses $k_a(t)$, $k_b(t)$, and $k_c(t)$, respectively. To facilitate our discussion, we shall write the transfer function, $K(s)$, of a linear system in the form

$$K(s) = \frac{P(s)}{Q(s)}. \tag{9}$$

The zeros of $P(s)$ are the zeros of $K(s)$, and the zeros of $Q(s)$ are the poles of $K(s)$. Note, however, that $P(s)$ and $Q(s)$ are not necessarily polynomials in s. In terms of Eq. 9, we can write the kernel transform, $K_2(s_1,s_2)$, as given by Eq. 7, as

$$K_2(s_1,s_2) = \frac{P_a(s_1)\, P_b(s_2)\, P_c(s_1+s_2)}{Q_a(s_1)\, Q_b(s_2)\, Q_c(s_1+s_2)}. \tag{10}$$

2. Second-order Systems with N Multipliers

We now note that a system consisting of N multipliers, whose kernel is the most general second-order kernel that can be obtained, is one whose output

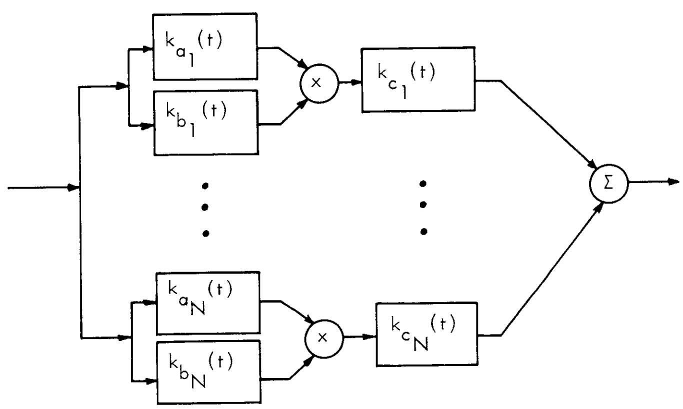


Fig. XVI-4. The most general second-order kernel with N multipliers.

is the sum of the outputs of N systems of the form depicted in Fig. XVI-3. Such a system is shown in Fig. XVI-4. By use of Eq. 6, the second-order kernel of this system is

$$k_2(\tau_1,\tau_2) = \sum_{n=1}^{N} \int_0^\infty k_{a_n}(\tau_1-\sigma)\, k_{b_n}(\tau_2-\sigma)\, k_{c_n}(\sigma)\, d\sigma, \tag{11}$$

and from Eq. 7, its kernel transform is

$$K_2(s_1,s_2) = \sum_{n=1}^{N} K_{a_n}(s_1)\, K_{b_n}(s_2)\, K_{c_n}(s_1+s_2). \tag{12}$$

By use of Eq. 10, the kernel transform can be written in the form

$$K_2(s_1,s_2) = \sum_{n=1}^{N} \frac{P_{a_n}(s_1)\, P_{b_n}(s_2)\, P_{c_n}(s_1+s_2)}{Q_{a_n}(s_1)\, Q_{b_n}(s_2)\, Q_{c_n}(s_1+s_2)}. \tag{13}$$

The sum of the series of Eq. 13 is of the form

$$K_2(s_1,s_2) = \frac{P_2(s_1,s_2)}{Q_2(s_1,s_2)}. \tag{14}$$

The zeros of $P_2(s_1,s_2)$ are the zeros of $K_2(s_1,s_2)$, and the zeros of $Q_2(s_1,s_2)$ are the poles of $K_2(s_1,s_2)$. From Eq. 13, we note that

$$P_2(s_1,s_2) = \sum_{n=1}^{N} R_{a_n}(s_1)\, R_{b_n}(s_2)\, R_{c_n}(s_1+s_2), \tag{15}$$

in which

$$R_{\alpha_n}(s) = P_{\alpha_n}(s) \prod_{\substack{j=1 \\ j\neq n}}^{N} Q_{\alpha_j}(s) \qquad \alpha = a,\ b \text{ or } c \tag{16}$$

and

$$Q_2(s_1,s_2) = F_a(s_1)\, F_b(s_2)\, F_c(s_1+s_2) \tag{17}$$

in which

$$F_\alpha(s) = \prod_{n=1}^{N} Q_{\alpha_n}(s) \qquad \alpha = a,\ b \text{ or } c. \tag{18}$$

Thus $Q_2(s_1,s_2)$ is expressible as the product of three functions: a function of s_1, a function of s_2, and a function of (s_1+s_2); also, $P_2(s_1,s_2)$ is expressible as the sum of N such products.

We note that if N multipliers are used to synthesize a second-order kernel, then the second-order kernel transform must be of the form

$$K_2(s_1,s_2) = \frac{\sum_{n=1}^{N} R_{a_n}(s_1)\, R_{b_n}(s_2)\, R_{c_n}(s_1+s_2)}{F_a(s_1)\, F_b(s_2)\, F_c(s_1+s_2)}. \tag{19}$$

Furthermore, if a second-order kernel transform is expressible in the form of Eq. 19, then it can be synthesized by means of N multipliers.

3. Illustration of the Method of Synthesis

We shall illustrate, by means of two examples, a procedure by which this synthesis can be accomplished.

As our first example, we desire to synthesize the following second-order kernel transform:

$$K_2(s_1,s_2) = \frac{P_2(s_1,s_2)}{Q_2(s_1,s_2)}$$

$$= \frac{s_1^2s_2 + 4s_1^2 + 12s_1s_2 + s_1s_2^2 + 2s_2^2 + 12s_1 + 48}{s_1^2s_2 + 4s_1^2 + 12s_1s_2 + s_1s_2^2 + 2s_2^2 + 32s_1 + 20s_2 + 48}. \tag{20}$$

This kernel transform can be synthesized by means of a finite number of multipliers. Thus $Q_2(s_1,s_2)$ is expressible in the form $F_a(s_1)\, F_b(s_2)\, F_c(s_1+s_2)$. To determine these three functions, we first let $s_1 = 0$ and $s_2 = s$. Then the zeros of $Q_2(0,s)$ are the zeros of $F_a(0)\, F_b(s)\, F_c(s)$. Thus,

$$F_a(0)\, F_b(s)\, F_c(s) = 2s^2 + 20s + 48$$

$$= 2(s+4)(s+6). \tag{21}$$

Second, let $s_1 = s$ and $s_2 = 0$. Then the zeros of $Q_2(s,0)$ are the zeros of $F_a(s)\, F_b(0)\, F_c(s)$. Thus,

$$F_a(s)\, F_b(0)\, F_c(s) = 4s^2 + 32s + 48$$

$$= 4(s+2)(s+6). \tag{22}$$

Third, let $s_1 = s$ and $s_2 = -s$. Then the zeros of $Q_2(s,-s)$ are the zeros of $F_a(s)\, F_b(-s)\, F_c(0)$. Thus,

$$F_a(s)\, F_b(-s)\, F_c(0) = -6s^2 + 12s + 48$$

$$= 6(s+2)(-s+4). \tag{23}$$

By comparing the common zeros of Eqs. 21-23, we note that

$$\begin{aligned} F_a(s) &= (s+2) \\ F_b(s) &= (s+4) \\ F_c(s) &= (s+6) \end{aligned} \tag{24}$$

so that

$$Q_2(s_1,s_2) = (s_1+2)(s_2+4)(s_1+s_2+6). \tag{25}$$

By substituting Eq. 25 in Eq. 20 and expanding by partial fractions, we obtain

$$K_2(s_1,s_2) = 1 - \frac{20(s_1+s_2)}{(s_1+2)(s_2+4)(s_1+s_2+6)}. \tag{26}$$

Figure XVI-5 depicts the system whose second-order kernel transform is given by Eq. 26.

For our second example, we desire to synthesize a second-order system whose second-order kernel is

$$k_2(\tau_1,\tau_2) = \begin{cases} e^{-a\tau_1-b\tau_2} & \text{for } 0 \le \tau_1 \le c\tau_2 \\ 0 & \text{otherwise} \end{cases} \tag{27}$$

in which $a > 0$, $b > 0$, and $c > 0$. To accomplish this, we first must determine the second-order kernel transform, $K_2(s_1,s_2)$.

$$K_2(s_1,s_2) = \int_0^\infty d\tau_2 \int_0^{c\tau_2} d\tau_1\, e^{-a\tau_1-b\tau_2}\, e^{-s_1\tau_1-s_2\tau_2}$$

$$= \frac{-c}{[b+s_2][b+ac+cs_1+s_2]}. \tag{28}$$

We note that for $c \ne 1$, the second-order kernel transform cannot be synthesized with a finite number of multipliers. This can be seen by assuming that the denominator,

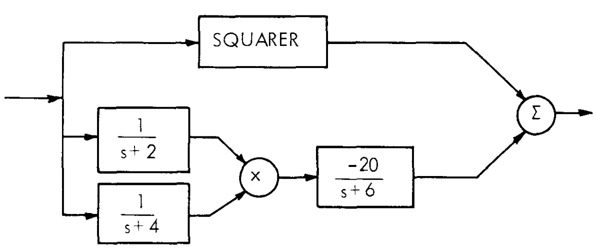


Fig. XVI-5. Second-order system whose kernel transform is given by Eq. 26.

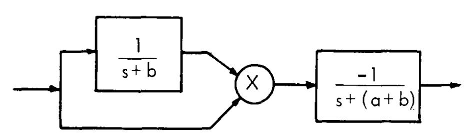


Fig. XVI-6. Second-order system whose kernel is given by Eq. 27 for c = 1.

$Q_2(s_1,s_2)$, can be expressed in the form $F_a(s_1)\, F_b(s_2)\, F_c(s_1+s_2)$. Then, by following the procedure used in our first example, we have

$$\begin{aligned} Q_2(s,0) &= b(b+ac+cs) \\ Q_2(0,s) &= (b+s)(b+ac+s) \\ Q_2(s,-s) &= (b-s)(b+ac+cs-s). \end{aligned} \tag{29}$$

For $c \neq 1$, the Eqs. 29 have no common zeros and thus the denominator, $Q_2(s_1,s_2)$, cannot be expressed in the form as given by Eq. 17.

However, for the special case in which $c = 1$, we have

$$K_2(s_1,s_2) = \frac{-1}{(b+s_2)(b+a+s_1+s_2)}. \tag{30}$$

This second-order kernel transform can be synthesized as shown in Fig. XVI-6.

4. Systems with Higher-Order Kernels

We now note that our results for second-order systems can be extended directly to higher-order systems. For example, a system consisting of only two multipliers whose kernel is clearly the most general third-order kernel that can be obtained is depicted in Fig. XVI-7. The third-order transfer function of this system is

$$K_3(s_1,s_2,s_3) = K_a(s_1)\, K_b(s_2)\, K_c(s_1+s_2)\, K_d(s_3)\, K_e(s_1+s_2+s_3). \tag{31}$$

Thus, if 2N multipliers are used to synthesize a third-order transfer function, then it must be of the form

$$K_3(s_1,s_2,s_3) = \frac{\sum_{n=1}^{N} R_{a_n}(s_1)\, R_{b_n}(s_2)\, R_{c_n}(s_1+s_2)\, R_{d_n}(s_3)\, R_{e_n}(s_1+s_2+s_3)}{F_a(s_1)\, F_b(s_2)\, F_c(s_1+s_2)\, F_d(s_3)\, F_e(s_1+s_2+s_3)}. \tag{32}$$

Furthermore, if a third-order transfer function is expressible in the form of Eq. 32, then it can be synthesized by means of 2N multipliers. A procedure

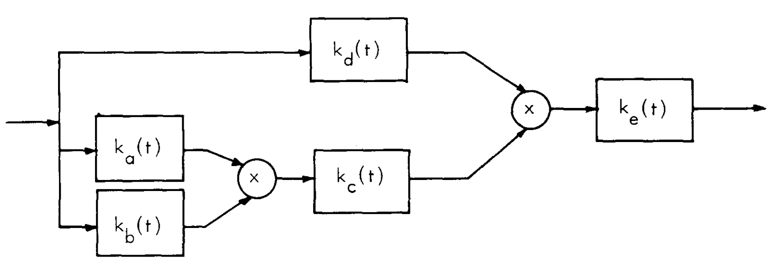


Fig. XVI-7. The most general third-order kernel with only two multipliers.

by which this can be accomplished is similar to the one described for the synthesis of second-order systems.

M. Schetzen

References

1. N. Wiener, Nonlinear Problems in Random Theory (The Technology Press of Massachusetts Institute of Technology, Cambridge, Mass., and John Wiley and Sons, Inc., New York, 1958).

2. 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. 118-130.

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.

B. MAXIMA OF THE MEAN-SQUARE ERROR IN OPTIMUM NONLINEAR NO-MEMORY FILTERS

1. Introduction

In this report we are interested in the problem of mean-square filtering for the class of no-memory filters shown in Fig. XVI-8.

We consider the case

$$x(t) = m(t) + n(t)$$

$$z(t) = m(t),$$

in which m(t), the message, and n(t), the noise, are statistically independent. For this case we have obtained and reported previously[1] a relation between the amplitude probability density of the input and the characteristic of the optimum nonlinear no-memory filter. In the present report we use those results to derive a simple expression for the

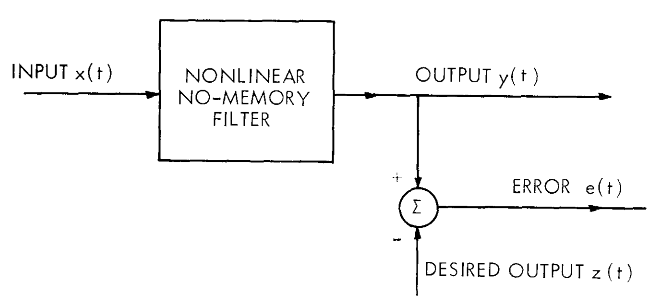


Fig. XVI-8. Filter operation considered in this report.

resulting mean-square error in the cases of Poisson noise and Gaussian noise. In the case of additive Gaussian noise, this expression involves only the amplitude probability density of the input and its first derivative. We take advantage of this characteristic to find the message probability density corresponding to the largest mean-square error in an optimum nonlinear no-memory filter under the constraint that the average message power be constant. The resulting message probability density is Gaussian.

2. Expression for the Mean-Square Error

The relation between the probability density of the input of the filter and the characteristic of the optimum nonlinear no-memory filter which we obtained previously is repeated here:

$$[x-g(x)]\, q(x) = \int q(x_1)\, f(x-x_1)\, dx_1.$$

Here, $q(x) = p_{m+n}(x)$ is the amplitude probability density of the input, $g(x)$ is the optimum nonlinear no-memory filter characteristic and

$$f(x) = \frac{1}{2\pi} \int F(t)\, e^{-jtx}\, dt$$

$$F(t) = \frac{d[\ln P_n(t)]}{d(jt)}$$

in which $P_n(t)$ is the characteristic function for the noise. Note that we use an integral sign without limits to indicate integration from $-\infty$ to $+\infty$. We can express $g(x)$ in terms of $q(x)$ by writing

$$g(x) = \frac{-\int q(x_1)\, f(x-x_1)\, dx_1}{q(x)} + x.$$

Whenever $f(x)$ is a singularity function, this relation for $g(x)$ leads to a simple expression for the mean-square error in terms of $q(x)$, the input probability density. The two well-known, nontrivial noise characteristics that give a singularity function for $f(x)$ are Gaussian noise and Poisson noise. We shall use the following expression for the mean-square error:

$$\overline{\mathscr{E}^2} = \overline{m^2} - \int g^2(x)\, q(x)\, dx,$$

in which $\overline{m^2}$ is the mean-square value of the message. This expression is obtained without difficulty by using the known result that the error resulting from the optimum mean-square filter is uncorrelated with the output of all nonlinear no-memory filters with the same input.

a. Poisson Noise

$$P(x=k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$P(t) = e^{\lambda(e^{jt}-1)}$$

hence

$$f(x) = \lambda u(x-1),$$

and

$$g(x) = -\lambda \frac{q(x-1)}{q(x)} + x.$$

$$\overline{\mathscr{E}^2} = m^2 - \int \left[-\lambda \frac{q(x-1)}{q(x)} + x\right]^2 q(x)\, dx$$

$$= m^2 - \int x^2 q(x)\, dx + 2\lambda \int xq(x-1)\, dx - \lambda^2 \int \frac{q^2(x-1)}{q(x)}\, dx$$

Since

$$\int xq(x-1)\, dx = \int (x+1)q(x)\, dx = \overline{m} + \lambda + 1$$

and

$$\int x^2 q(x)\, dx = m^2 + \lambda + \lambda^2 + 2\overline{m}\lambda,$$

then we have

$$\overline{\mathscr{E}^2} = \lambda + \lambda^2 - \lambda^2 \int \frac{q^2(x-1)}{q(x)}\, dx. \tag{1}$$

Equation 1 holds whether or not the message has zero mean.

b. Gaussian Noise

$$p_n(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

$$P_n(t) = \exp\left(-\frac{\sigma^2 t^2}{2}\right),$$

hence

$$f(x) = -\sigma^2 u_1(x)$$

and

$$g(x) = \sigma^2 \frac{q'(x)}{q(x)} + x.$$

$$\overline{\mathscr{E}^2} = m^2 - \int \left[\sigma^2 \frac{q'(x)}{q(x)} + x\right]^2 q(x)\, dx$$

Since

$$\int x^2 q(x)\, dx = m^2 + \sigma^2$$

and

$$\int xq'(x)\, dx = xq(x)\Big|_{-\infty}^{+\infty} - \int q(x)\, dx = -1,$$

then we have

$$\overline{\mathscr{E}^2} = \sigma^2 - \sigma^4 \int \frac{q'^2(x)}{q(x)}\, dx. \tag{2}$$

3. Maximum of the Error under Constraints for Additive Gaussian Noise

Since we have an expression for the mean-square error solely in terms of the input probability density, we can find extrema of the error under constraint by the method of calculus of variations. We shall consider, for the present, a power constraint on the input.

a. Power Constraint

We consider here a filtering problem characterized by additive Gaussian noise of known average power. We consider all possible messages of fixed average power, and in each case use the optimum nonlinear no-memory filter in the mean-square sense to separate the message from the noise. We now undertake to find the message probability density that gives an extremum of the mean-square error. Since the message and the noise are statistically independent, a constraint on the input average power is equivalent to a constraint on the message average power, and we write

$$\int x^2 q(x)\, dx = m^2 + \sigma^2.$$

Other constraints are

$$\int q(x)\, dx = 1$$

$$q(x) \geq 0 \qquad \text{for all } x$$

We take care of the last constraint by letting

$$y^2(x) \triangleq q(x),$$

and we have

$$\overline{\mathscr{E}^2} = \sigma^2 - 4\sigma^4 \int y'^2(x)\, dx. \tag{3}$$

Because $q(x)$ is the result of convolving a Gaussian probability density with the message probability density $p_m(x)$, we would need another constraint on $q(x)$ to ensure that $p_m(x)$ is positive. However, this constraint cannot be handled analytically, and we shall have to select among the solutions obtained for $q(x)$ those leading to an acceptable probability density. In terms of $y(x)$, using the Lagrange multiplier, we look for extrema of

$$J = \int \left[y'^2 + \lambda_1 x^2 y^2 + \lambda_2 y^2\right] dx,$$

in which λ_1 and λ_2 are Lagrange multipliers. This leads to the following Euler-Lagrange equation:

$$y'' + y\left[\lambda_1 x^2 + \lambda_2\right] = 0. \tag{4}$$

We have obtained here the Weber-Hermite differential equation. Since we are looking for solutions that are square integrable, we have the boundary conditions

$$y \to 0 \quad \text{for} \quad |x| \to \infty.$$

The differential equation has solutions that satisfy these boundary values[2] only if it is in the form

$$\frac{d^2y}{du^2} + y\left[n + \frac{1}{2} - \frac{u^2}{2}\right] = 0 \tag{5}$$

in which n, a non-negative integer, is the eigenvalue. The corresponding solutions or eigenfunctions are the Hermite functions

$$y_n(u) = D_n(u) \triangleq \exp\left(-\frac{u^2}{2}\right) 2^{-n/2} H_n\left(\frac{u}{\sqrt{2}}\right)$$

in which $H_n(v)$ is the Hermite polynomial.

$$H_n(v) \triangleq (-1)^n e^{v^2} \frac{d^n e^{-v^2}}{dv^n}$$

To put Eq. 4 in the form of Eq. 5, we let $x = cu$, in which c is a constant, and thus obtain the solution

$$y_n(x) = A\, D_n\left(\frac{x}{c}\right).$$

Here, A, an arbitrary constant, appears because the linear differential equation to be satisfied is an homogeneous equation. The solution for the amplitude probability density of the input becomes

$$q_n(x) = A^2 D_n^2\left(\frac{x}{c}\right).$$

It can be shown that the minimum of the integral $\int y'^2\, dx$ that appears with a minus sign in the expression for the mean-square error (Eq. 3) corresponds to the eigenvalue $n = 0$. For $n = 0$ we have

$$q(x) = A^2 \exp\left(-\frac{x^2}{2c^2}\right)$$

which is, therefore, the amplitude probability of the input giving the maximum mean-square error.

We satisfy the constraints by letting $A^2 = 1/\sqrt{2\pi}\, c$, and $c^2 = \sigma^2 + m^2$. Therefore,

$$q(x) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2 + m^2}} \exp\left(-\frac{x^2}{2(\sigma^2+m^2)}\right).$$

The probability density of the message now is

$$p_m(x) = \frac{1}{\sqrt{2\pi}\, m} \exp\left(-\frac{x^2}{2m^2}\right).$$

Hence, when the noise is Gaussian and additive, and the message has a fixed average power, the maximum mean-square error is obtained whenever the message is also Gaussian. In such a case, the optimum no-memory filter reduces to an attenuator and

$$\overline{\mathscr{E}^2} = \sigma^2 - \sigma^4 \int \frac{q'^2(x)}{q(x)}\, dx = \frac{\sigma^2 m^2}{\sigma^2 + m^2}.$$

One might wonder if some interpretation can be given in the context to higher-order eigenvalues and eigenfunctions ($n = 1, 2$, etc.) which correspond to stationary values of the expression for mean-square error.

However, although $(q(x)=A^2D_n^2(x/c))$, the probability density of the input, is positive for all x, the corresponding message probability density $p_m(x)$ is not strictly positive for $n > 0$ and does not correspond to a physical situation.

Although we did not obtain it by the present formulation, an interesting result would be to find out whether a minimum of the optimum mean-square error exists under the same constraint. It is possible to show that an arbitrarily small error can be achieved unless additional constraints are used. Work is under way in this area.

V. R. Algazi

References

1. V. R. Algazi, Nonlinear no-memory filters of prescribed form, Quarterly Progress Report No. 66, Research Laboratory of Electronics, M.I.T., July 15, 1962, p. 201.

2. See, for instance, G. Birkhoff and G. C. Rota, Ordinary Differential Equations (Ginn and Company, Boston, 1962), p. 280.