XIII. PROCESSING AND TRANSMISSION OF INFORMATION*

Prof. R. M. Fano
Prof. R. G. Gallager
Prof. F. C. Hennie III
Prof. I. M. Jacobs
Prof. A. M. Manders
Prof. W. F. Schreiber
Prof. C. E. Shannon
Prof. J. M. Wozencraft
Dr. C. L. Liu
M. H. Bender
E. R. Berlekamp
D. G. Botha
J. E. Cunningham
H. Dym
P. M. Ebert
D. D. Falconer
E. F. Ferretti
G. D. Forney, Jr.
U. F. Gronemann
A. R. Hassan
J. L. Holsinger
T. S. Huang
N. Imai
J. E. Savage
J. R. Sklar
K. D. Snow
I. G. Stiglitz
W. R. Sutherland
O. J. Tretiak
W. J. Wilson
H. L. Yudkin

A. VECTOR REPRESENTATION OF TIME-CONTINUOUS CHANNELS WITH MEMORY

Analysis of time-continuous channels is usually accomplished by representing both signals and noise as vectors (or n-tuples) in an n-dimensional vector space[1,2]; the components of the vectors are the coordinates of signal or noise with respect to the set of basis functions used. For memoryless channels with additive white Gaussian noise, this results in the simple representation

$$\underline{y} = \underline{x} + \underline{n}, \tag{1}$$

where $\underline{x}$, $\underline{n}$, and $\underline{y}$ are the vector representations with respect to a common basis of the channel input, additive noise, and channel output, respectively, and the components of $\underline{n}$

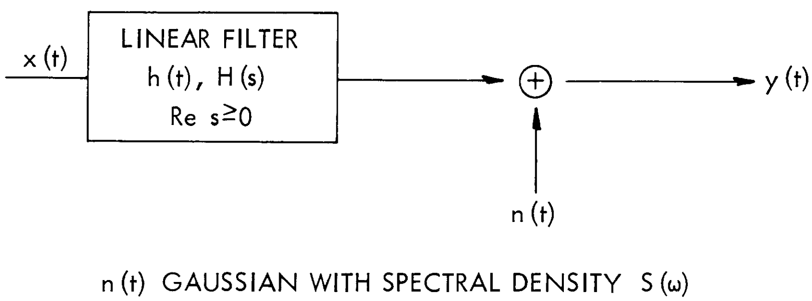

Fig. XIII-1. Gaussian channel with memory.

are statistically independent and identically distributed Gaussian random variables. This report demonstrates that for the class of channels with memory which is defined in Fig. XIII-1 there exist basis functions such that

*This research was supported in part by Purchase Order DDL BB-107 with Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology, with the joint support of the U.S. Army, Navy, and Air Force under Air Force Contract AF19(604)-7400; and in part by the National Institutes of Health (Grant MH-04737-03), and the National Science Foundation (Grant G-16526); additional support was received under NASA Grant NsG-496.

$$\underline{y} = [\sqrt{\lambda}]\underline{x} + \underline{n}, \tag{2}$$

where $\underline{x}$, $\underline{n}$, and $\underline{y}$ are again the (column) vector representations[3] of the channel input, additive noise, and channel output, respectively, the components of $\underline{n}$ are statistically independent and identically distributed Gaussian random variables, and $[\sqrt{\lambda}]$ is a diagonal matrix. The important result here is that $[\sqrt{\lambda}]$ is diagonal and the noise components are identically distributed, since a representation similar to Eq. 2 is always possible when these conditions are not satisfied. This diagonalization is accomplished by using different but closely related basis functions for the representation of x(t) and y(t) in contrast to the common bases of Eq. 1. In fact, it will be seen later that Eq. 1 is a special case of Eq. 2 with $h(t) = \delta(t)$ and $S(\omega)$ = constant.

The main results of this report are presented in the form of several theorems. First, however, some assumptions and simplifying notation will be introduced.

Assumptions

(i) The time functions x(t), n(t), and y(t) are in all cases real.

(ii) x(t) is nonzero only on 0, T and has finite energy, that is, $x(t) \in \mathscr{L}_2$ and thus

$$\int_0^T x^2(t)\, dt < \infty.$$

(iii) The time scale for y(t) is shifted to remove any pure delay in h(t).

Notation

(i) The standard inner product on the interval 0, T' is written $(f, g)_{T'}$, that is,

$$(f, g)_{T'} = \int_0^{T'} f(t)\, g(t)\, dt.$$

(ii) The linear integral operation on $f(t, t'')$ by $k(t, t')$ is written $kf(t, t'')$, that is,

$$kf(t, t'') = \int_0^T k(t, t')\, f(t', t'')\, dt'.$$

(iii) The generalized inner product on 0, T' is written $(f, kg)_{T'}$, that is,

$$(f, kg)_{T'} = \int_0^{T'} f(t)\, kg(t)\, dt = \int_0^{T'} \int_0^T f(t)\, k(t, t')\, g(t')\, dt'\, dt.$$

With these preliminaries the pertinent theorems can now be stated. It is worth noting that the following results are closely related to the spectral decomposition of linear self-adjoint operators on a Hilbert space.

THEOREM 1: Let $S(\omega) = N_o$, define a symmetric function $R(\tau, \tau') = R(\tau', \tau)$ in terms of the filter impulse response by

$$R(\tau', \tau) = \int_0^{T_1} h(t-\tau)\, h(t-\tau')\, dt, \qquad T_1 > T,$$

and define a set of eigenfunctions and eigenvalues by

$$\lambda_i \phi_i(t) = R\phi_i(t) \qquad \begin{array}{l} 0 \leq t \leq T \\ i = 0, 1, \dots . \end{array} \tag{3}$$

(Here and throughout this report it is assumed that $\phi_i(t) = 0$ for $t < 0$ and $t > T$.) Then the vector representation of $x(t)$ on $0, T$ is $\underline{x}$, where $x_i = (x, \phi_i)_T$, and the vector representation of $y(t)$ on the interval $0, T_1$ is given by Eq. 2 in which the components of $\underline{n}$ are statistically independent Gaussian random variables with mean zero and variance N_o and

$$[\sqrt{\lambda}] = \begin{bmatrix} \sqrt{\lambda_o} & & 0 & \\ & \ddots & & \\ 0 & & \sqrt{\lambda_n} & \\ & & & \ddots \end{bmatrix}. \tag{4}$$

The basis functions for $\underline{y}$ are $\{\theta_i(t)\}$, where

$$\theta_i(t) = \begin{cases} \dfrac{1}{\sqrt{\lambda_i}} h\phi_i(t) & 0 \leq t \leq T_1 \\ 0 & t < 0, \quad t > T_1, \end{cases} \tag{5}$$

and the i^{th} component of $\underline{y}$ is

$$y_i = (y, \theta_i)_{T_1}.$$

PROOF: First, it must be shown that the $\phi_i(t)$ defined by Eq. 3 exist and are a complete basis for representing $x(t)$. Since the kernel of Eq. 3 is symmetric, it is well known[5] that at least one nonzero solution exists. Furthermore, proof that

$$(f, Rf)_T > 0 \tag{6}$$

for any nonzero $f(t) \in \mathscr{L}_2$ is necessary and sufficient to prove completeness of the $\{\phi_i(t)\}$.[6] But

$$(f, Rf)_T = \int_0^{T_1} \left[\int_0^T f(\tau)\, h(t-\tau)\, d\tau\right]^2 dt.$$

Thus it suffices to prove that if

$$\int_0^T f(\tau)\, h(t-\tau)\, d\tau \equiv 0 \qquad 0 \leq t \leq T_1,$$

then $f(t) \equiv 0$ almost everywhere on $0, T$. Let

$$r(t) = \int_0^T f(\tau)\, h(t-\tau)\, dt,$$

and assume that

$$r(t) = \begin{cases} 0 & t \leq T_1 \\ z(t-T_1) & t > T_1, \end{cases}$$

where $z(t)$ is zero for $t < 0$ and is arbitrary except that it must be the result of passing some $\mathscr{L}_2$ signal through $h(t)$. Then

$$R(s) = \int_0^\infty r(t)\, e^{-st}\, dt = e^{-sT_1} Z(s) = F(s)\, H(s) \qquad \text{Re } s \geq 0,$$

and it follows that $Z(s)$ must contain any zeros of $H(s)$. (Note that $H(s)$ contains no exponential term by Assumption 3.) Thus

$$f(t) = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} e^{-sT_1} \frac{Z(s)}{H(s)} e^{st}\, ds,$$

which, when evaluated by residue calculus, yields $f(t) \equiv 0$ for $t < T_1$, since the integrand contains no right-hand plane singularities. Recall that by assumption, $T_1 > T$. Thus it follows that any $x(t) \in \mathscr{L}_2$ can be represented as

$$x(t) = \sum_{i=0}^\infty x_i \phi_i(t) \qquad 0 \leq t \leq T, \tag{7}$$

where $x_i = (x, \phi_i)$ and mean-square convergence is assumed.

Second, define $\theta_i(t)$ as in Eq. 5. Then the signal at the filter output on the interval $0, T_1$ is given by

$$r(t) = \sum_{i=0}^\infty \sqrt{\lambda_i}\, x_i \theta_i(t), \tag{8}$$

and thus for $0 \leq t \leq T_1$

$$y(t) = r(t) + n(t) = \sum_{i=0}^\infty \sqrt{\lambda_i}\, x_i \theta_i(t) + n(t), \tag{9}$$

however,

$$(\theta_i, \theta_j)_{T_1} = \frac{1}{\sqrt{\lambda_i \lambda_j}} (h\phi_i, h\phi_j)_{T_1} = \frac{1}{\sqrt{\lambda_i \lambda_j}} (\phi_i, R\phi_j)_T = \sqrt{\frac{\lambda_j}{\lambda_i}} (\phi_i, \phi_j)_T = \delta_{ij}.$$

Thus the final result is obtained:

$$y(t) = \sum_{i=0}^{\infty} y_i\theta_i(t) \qquad 0 \leq t \leq T_1, \tag{10a}$$

where $y_i = (y, \theta_i) = \sqrt{\lambda_i}\, x_i + n_i$, $n_i = (n, \theta_i)$, and $\overline{n_i n_j} = N_o \delta_{ij}$.

Q.E.D.

Note that Eq. 10a should actually be written as

$$y(t) = \sum_{i=0}^{\infty} y_i\theta_i(t) + r_n(t), \tag{10b}$$

where $(\theta_i, r_n)_{T_1} = 0$ $(i = 0, 1, \ldots)$, since, in general, the $\{\theta_i(t)\}$ do not form a complete basis for n(t). However, $r_n(t)$ is, in effect, "outside" the space of signals that can be obtained from the filter output and thus is of no interest in subsequent calculations.

THEOREM 2: Let H(s) be restricted to be a ratio of polynomials in s (that is, the transfer function of a lumped parameter network) and define two polynomials in s^2, $N(s^2)$ and $D(s^2)$ by

$$H(s)\, H(-s) = \frac{N(s^2)}{D(s^2)},$$

then the $\{\phi_i(t)\}$ defined by Eq. 3 satisfy the differential equation

$$N\left(\frac{d^2}{dt^2}\right)\phi_i(t) - \lambda_i D\left(\frac{d^2}{dt^2}\right)\phi_i(t) = 0 \qquad 0 \leq t \leq T.$$

PROOF: Let

$$H(s) = \frac{N^+(s)}{D^+(s)}.$$

Then

$$h(t) = \int_{-\infty}^{\infty} \frac{N^+(s)}{D^+(s)}\, e^{st}\, df, \quad s = j\, 2\pi f,$$

and from Eq. 3 it follows that

$$\lambda_i\phi_i(t) = \int_0^T \phi_i(t') \int_0^{T_1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{N^+(s)}{D^+(s)} \frac{N^+(s')}{D^+(s')}\, e^{s(t''-t')}\, e^{s'(t''-t)}\, df\,df'\,dt''\,dt'$$

$$0 \leq t \leq T < T_1.$$

Differentiating both sides gives

$$\lambda_i D^+\left(-\frac{d}{dt}\right)\phi_i(t) = \int_0^T \phi_i(t') \int_0^{T_1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{N^+(s)}{D^+(s)} N^+(s')\, e^{s(t''-t')}\, e^{s'(t''-t)}\, df df' dt'' dt',$$

and thus

$$\lambda_i D^+\left(-\frac{d}{dt}\right)\phi_i(t) = N^+\left(-\frac{d}{dt}\right) \int_0^T \phi_i(t') \int_0^{T_1} \int_{-\infty}^{\infty} \frac{N^+(s)}{D^+(s)} e^{s(t''-t')} \int_{-\infty}^{\infty} e^{s'(t''-t)}\, df df' dt'' dt'$$

$$= N^+\left(-\frac{d}{dt}\right) \int_0^T \phi_i(t') \int_{-\infty}^{\infty} \frac{N^+(s)}{D^+(s)} \int_0^{T_1} e^{s(t''-t')}\, \delta(t''-t)\, dt'' df dt'$$

$$= N^+\left(-\frac{d}{dt}\right) \int_0^T \phi_i(t') \int_{-\infty}^{\infty} \frac{N^+(s)}{D^+(s)} e^{s(t-t')}\, df dt',$$

which, after further, and similar, manipulations, becomes

$$\lambda_i D\left(\frac{d^2}{dt^2}\right)\phi_i(t) = N\left(\frac{d^2}{dt^2}\right) \int_0^T \phi_i(t') \int_{-\infty}^{\infty} e^{s(t-t')}\, df dt'$$

$$= N\left(\frac{d^2}{dt^2}\right)\phi_i(t) \qquad 0 \leq t \leq T.$$

Q. E. D.

The functions $\{\phi_i(t)\}$ and the representation of Eq. 2 have some other properties of interest. The optimum (matched filter) detector for detecting the presence of x(t) by observing y(t) only on the interval $0, T_1$ makes a decision based on the quantity

$$\int_0^{T_1} y(t)\, r(t)\, dt = (y, hx)_{T_1} = \underline{y} \cdot \underline{r},$$

where $\underline{r} = [\sqrt{\lambda}]\ \underline{x}$ and $\cdot$ denotes the usual vector inner product. The probability of error in this case is determined by E/N_o, where

$$\frac{E}{N_o} = \frac{(hx, hx)_{T_1}}{N_o} = \frac{\underline{r} \cdot \underline{r}}{N_o} = \frac{\sum_{i=0}^{\infty} \lambda_i x_i^2}{N_o}.$$

Also, the input energy is given by

$$(x, x)_T = \underline{x} \cdot \underline{x} = \sum_{i=0}^{\infty} x_i^2.$$

By convention, $\lambda_o \geq \lambda_1 \geq \lambda_2 \geq \ldots$. Thus it follows that for fixed input-signal energy the probability of error is minimum (that is, the output-signal energy is maximum) when $x(t) = \phi_o(t)$.[7] More generally, it is known[8] that $x(t) = \phi_j(t)$ is the function giving maximum output energy under the constraints

$$(x, \phi_i)_T = 0 \qquad i = 0, 1, \ldots, j-1$$

and

$$(x, x)_T = 1.$$

Finally, it should be noted that (a) when $T_1 = \infty$, Eq. 3 reduces to the well-known[9] and studied[10] integral equation involved in the expansion of a random process with autocorrelation function $\mathcal{R}(\tau-\tau') = R(\tau, \tau') = R(\tau-\tau')$, and (b) when $R(\tau, \tau')$ is defined as

$$R(\tau, \tau') = R(\tau-\tau') = \int_{-\infty}^{\infty} h(t-\tau)\, h(t-\tau')\, dt$$

and h(t) is specialized to $h(t) = (\sin \pi t)/\pi t$, then the $\{\phi_i(t)\}$ are the prolate spheroidal wave functions studied by Slepian and Pollak.[11]

The following theorem specifies basis functions for nonwhite noise and $T_1 = \infty$. For physical, as well as mathematical, reasons the following conditions are assumed.

$$\int_{-\infty}^{\infty} |H(\omega)|^2\, df < \infty$$

$$\int_{-\infty}^{\infty} S(\omega)\, df < \infty$$

$$\int_{-\infty}^{\infty} \frac{|H(j\omega)|^2}{S(\omega)}\, df < \infty.$$

THEOREM 3: Let the noise of Fig. XIII-1 have a spectral density $S(\omega)$, define two symmetric functions $K(\tau-\tau')$ and $K_1(\tau-\tau')$ by

$$K(\tau-\tau') = \int_{-\infty}^{\infty} \frac{|H(j\omega)|^2}{S(\omega)} \exp[j\omega(\tau-\tau')]\, df$$

and

$$K_1(\tau-\tau') = \int_{-\infty}^{\infty} \frac{1}{S(\omega)} \exp[j\omega(\tau-\tau')]\, df,$$

and define a set of eigenfunctions and eigenvalues by

$$\lambda_i\phi_i(\tau) = K\phi_i(\tau) \qquad 0 \leq \tau \leq T. \tag{11}$$

Then the vector representation of x(t) on the interval 0, T is $\underline{x}$, where $x_i = (x, \phi_i)$, and the vector representation of y(t) on the interval $0, \infty$ is given by Eq. 2, in which the components of $\underline{n}$ are statistically independent, identically distributed, Gaussian random variables with mean zero and variance unity, and $[\sqrt{\lambda}]$ is given by Eq. 2. The basis functions for $\underline{y}$ are $\{\theta_i(t)\}$, where

$$\theta_i(t) = \begin{cases} \dfrac{1}{\sqrt{\lambda_i}} h\phi_i(t) & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{12}$$

and the i^{th} component of $\underline{y}$ is

$$y_i = \int_{-\infty}^{\infty} y(t) \int_0^{\infty} K_1(t-t')\, \theta_i(t')\, dt'dt \equiv (y, K_1\theta_i)_\infty.$$

PROOF: Under the conditions assumed for this theorem, the functions $\{\phi_i(t)\}$ of Eq. 11 are simply a special case of Eq. 3 with $T_1 = \infty$ and $R(\tau, \tau') = K(\tau-\tau')$. Thus they are complete and the representation for x(t) follows directly. By defining $\theta_i(t)$ as in Eq. 12, the filter output r(t) for $t \geq 0$ is given by Eq. 8, and thus y(t) is given by Eq. 9, but

$$(\theta_i, K_1\theta_j)_\infty = \frac{1}{\sqrt{\lambda_i\lambda_j}} (h\phi_i, K_1 h\phi_j)_\infty = \frac{1}{\sqrt{\lambda_i\lambda_j}} (\phi_i, K\phi_j)_T = \sqrt{\frac{\lambda_j}{\lambda_i}} (\phi_i, \phi_j)_T = \delta_{ij}.$$

Thus y(t) is given by Eq. 10a for which

$$y_i = (y, K_1\theta_i)_\infty = \sqrt{\lambda_i}\, x_i + n_i,$$

and

$$n_i = (n, K_1\theta_i)_\infty,$$

with

$$\overline{n_i n_j} = \overline{(n, K_1\theta_i)_\infty (n, K_1\theta_j)_\infty}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{R}_n(t-t') \int_0^{\infty} K_1(t-t'')\, \theta_i(t'')\, dt'' \int_0^{\infty} K_1(t'-t''')\, \theta_j(t''')\, dt'''dtdt'$$

$$= \int_0^{\infty} \int_0^{\infty} dt''\, dt'''\theta_i(t'')\, \theta_j(t''') \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{R}_n(t-t')\, K_1(t-t'')\, K_1(t'-t''')\, dtdt'$$

$$= \int_0^{\infty} \int_0^{\infty} \theta_i(t)\, \theta_j(t')\, K_1(t-t')\, dtdt'$$

$$= (\theta_i, K_1\theta_j)_\infty = \delta_{ij}.$$

Note that since $K_1\theta_i(t)$ is in general nonzero for $t < 0$ the inner product for y_i and n_i must be over the doubly infinite interval. Q. E. D.

THEOREM 4: Let $H(s)$, $N(s^2)$, and $D(s^2)$ be defined as in Theorem 2, let $S(\omega)$ be the spectrum obtained by passing white noise through a lumped parameter filter, and define

$$S(\omega) = \left. \frac{N_n(s^2)}{D_n(s^2)} \right|_{s^2=-\omega^2}.$$

Then the solutions of Eq. 11 satisfy the differential equation

$$N\left[\frac{d^2}{dt^2}\right] D_n\left[\frac{d^2}{dt^2}\right] \phi_i(t) - \lambda_i D\left[\frac{d^2}{dt^2}\right] N_n\left[\frac{d^2}{dt^2}\right] \phi_i(t) = 0.$$

PROOF: This follows directly from Theorem 2 with $T_1 = \infty$. Q. E. D.

A final property of the expansion of Theorem 3 is as follows. It is known[12] that the optimum detector for detecting the presence of $x(t)$ by observing $y(t)$ over the doubly infinite time interval makes a decision based on the quantity

$$\int_{-\infty}^{\infty} y(t)\, q(t)\, dt,$$

where

$$q(t) = \int_{-\infty}^{\infty} \frac{X(\omega)\, H(\omega)}{S(\omega)} e^{j\omega t}\, df.$$

In vector notation this quantity becomes

$$\int_{-\infty}^{\infty} y(t)\, q(t)\, dt = (y, K_1 r)_\infty = \underline{y} \cdot \underline{r} = \underline{y} \cdot [\sqrt{\lambda}]\, \underline{x}.$$

The "generalized E/N_o" for this case is

$$\int_{-\infty}^{\infty} \frac{|X(\omega)|^2\, |H(\omega)|^2}{S(\omega)} df = (r, K_1 r)_\infty = \underline{r} \cdot \underline{r} = \sum_{i=0}^{\infty} \lambda_i x_i^2,$$

and, since the input energy is $\underline{x} \cdot \underline{x} = \sum_{i=0}^{\infty} x_i^2$, it follows that again $x(t) = \phi_o(t)$ is the optimum signal to be used to minimize the probability of error for fixed input energy.

Other results have been obtained on expansions for colored noise with a finite-time observation interval and for a class of $\{\phi_i(t)\}$ that are again time-limited after passing through the filter.

The author would like to thank Dr. Robert Price, Lincoln Laboratory, M.I.T., for his careful reading of this report.

J. L. Holsinger

References

1. C. E. Shannon, Communication in the presence of noise, Proc. IRE 37, 10 (1949).

2. C. E. Shannon, Probability of error for optimal codes in a Gaussian channel, Bell System Tech. J. 38, 611 (1959).

3. It should not be construed from this that the basis functions used to represent y(t) are complete in the noise space; in general, they are not. See Eq. 10b.

4. All of the properties of the $\{\phi_i(t)\}$ and $\{\theta_i(t)\}$ still follow when $T_1 < T$, except that, in general, the $\{\phi_i(t)\}$ will not be complete.

5. R. Courant and D. Hilbert, Methods of Mathematical Physics (Interscience Publishers, Inc., New York, 1953), p. 122.

6. F. G. Tricomi, Integral Equations (Interscience Publishers, Inc., New York, 1957), p. 109.

7. For the special case $T_1 = \infty$, this property appears to have been first recognized by J. H. H. Calk, The optimum pulse-shape for pulse communication, Proc. IRE 97, Part III, 88 (1950).

8. R. Courant and D. Hilbert, op. cit., p. 128.

9. W. B. Davenport, Jr. and W. L. Root, Random Signals and Noise (McGraw-Hill Book Co., New York, 1958).

10. D. C. Youla, Solution of an homogeneous Wiener-Hopf integral equation occurring in the expansion of second-order stationary random functions, Trans. IRE, Vol. IT-3, p. 187, September 1957.

11. D. Slepian and H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty – I, Bell System Tech. J. 40, 43 (1961).

12. C. W. Helstrom, Statistical Theory of Signal Detection (Pergamon Press, New York, 1960), p. 108.