#2

CODING WITH LINEAR SYSTEMS

JOHN P. COSTAS

LOAN COPY only

TECHNICAL REPORT NO. 226

FEBRUARY 20, 1952

RESEARCH LABORATORY OF ELECTRONICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS

The research reported in this document was made possible through support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, jointly by the Army Signal Corps, the Navy Department (Office of Naval Research) and the Air Force (Air Materiel Command), under Signal Corps Contract No. DA36-039 sc-100, Project No. 8-102B-0; Department of the Army Project No. 3-99-10-022.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
RESEARCH LABORATORY OF ELECTRONICS

Technical Report No. 226 February 20, 1952

CODING WITH LINEAR SYSTEMS

John P. Costas

Abstract

Message transmission over a noisy channel is considered. Two linear networks are to be designed: one to treat the message before transmission; the second to filter the treated message plus channel noise at the receiving end. The mean-square error between the actual transmission circuit output and the desired output is minimized for a given allowable average signal power by proper network design. The results of a numerical example are given and discussed.

CODING WITH LINEAR SYSTEMS

I. Introduction

The importance of the statistical approach to filter design as originated by Wiener (1, 2) and developed by Lee (3) has been discussed by this author in a forthcoming report (4). Briefly, for a given message plus noise input as shown in Fig. 1, we desire to find a filter characteristic which will give the best possible performance. By "best" we mean that filter which minimizes the mean-square error between the actual output $f_o(t)$ and the desired output $f_d(t)$. That is $\mathcal{E}$, the mean-square error of filtering, will be given by

$$\mathcal{E} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \left[f_o(t) - f_d(t) \right]^2 dt. \tag{1}$$

Such an error criterion is certainly a reasonable one from a physical point of view, but contrary to a popular notion it is by no means the only error measure amenable to mathematical treatment (5).

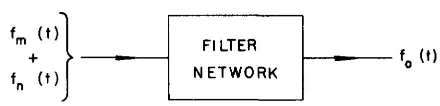

Fig. 1
Conventional filter problem:
$f_m(t)$ = message function;
$f_n(t)$ = noise function;
$f_o(t)$ = actual filter output;
$f_d(t)$ = desired filter output.

The desired filter output is usually the message function $f_m(t)$. However, it may happen that an output other than the message may be required. For example, one might ask for a network design which would filter the message from the noise, predict the message by α seconds, and differentiate the result. Thus, we could require prediction, filtering, and differentiation in one operation (3). In this sense a network may be considered as an operator rather than simply a filter (1).

If linear systems only are to be considered, the statistical parameters needed for design are known as the correlation functions. The crosscorrelation function $\phi_{12}(\tau)$ between random functions $f_1(t)$ and $f_2(t)$ is defined by

$$\phi_{12}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} f_1(t) f_2(t+\tau) dt. \tag{2}$$

The autocorrelation function $\phi_{11}(\tau)$ of the random function $f_1(t)$ is defined by

$$\phi_{11}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} f_1(t) f_1(t+\tau) dt. \tag{3}$$

The Fourier transform pair $g(t)$, $G(\omega)$ are related by

$$G(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(t)\, \epsilon^{-j\omega t}\, dt \tag{4}$$

and

$$g(t) = \int_{-\infty}^{\infty} G(\omega) \epsilon^{+j\omega t}\, d\omega. \tag{5}$$

By the Laplace transform we shall mean relations Eqs. 4 and 5 except that ω is replaced by λ where

$$\lambda = \omega + j\sigma. \tag{6}$$

An important theorem due to Wiener (3, 6) states that the power-density spectrum of a random function $f_1(t)$ is given by the Fourier transform of the autocorrelation function of $f_1(t)$. That is,

$$\Phi_{11}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{11}(\tau) \epsilon^{-j\omega\tau}\, d\tau. \tag{7}$$

In a similar manner, we may define a cross-power spectrum $\Phi_{12}(\omega)$ between random functions $f_1(t)$ and $f_2(t)$ as

$$\Phi_{12}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{12}(\tau) \epsilon^{-j\omega\tau}\, d\tau. \tag{8}$$

Let us define the unit impulse u(t) by

$$u(t) = \lim_{a \to \infty} \frac{a}{\sqrt{\pi}} \epsilon^{-a^2 t^2}.$$

Then we may show using Eq. 4 that the transform $\mathcal{U}(\omega)$ of u(t) is given by

$$\mathcal{U}(\omega) = \frac{1}{2\pi}. \tag{9}$$

Now if h(t) is the response of a linear system to a unit impulse input, it may be shown(3) that the output $f_o(t)$ of the linear system to an arbitrary input $f_i(t)$ will be given by

$$f_o(t) = \int_{-\infty}^{\infty} h(\sigma) f_i(t-\sigma)\, d\sigma. \tag{10}$$

Let $e_o(t)$ be the transient output of the linear system due to a transient input $e_i(t)$. If a system function $H(\omega)$ is defined for the linear system such that

$$H(\omega) = \frac{E_o(\omega)}{E_i(\omega)} \tag{11}$$

it may be shown that $H(\omega)$ and $h(t)$ are related by

$$H(\omega) = \int_{-\infty}^{\infty} h(t)\,\epsilon^{-j\omega t}\,dt \tag{12}$$

and

$$h(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega)\,\epsilon^{+j\omega t}\,d\omega \tag{13}$$

which correspond to Eqs. 4 and 5 except for the location of the 2π term.

II. The Transmission Problem

The filter design problem of Fig. 1 has been treated in great detail by both Wiener (1) and Lee (3), with experimental verification available in the work of C. A. Stutt (8). Therefore, in this report we shall consider the more general situation described by Fig. 2. In most communication systems the opportunity exists to modify or "code" the message to be transmitted before its introduction into the transmission channel. Network $H(\omega)$ must be designed so that the message is "pre-distorted" or "coded" in such a way as to enable the "decoding" or filtering network $G(\omega)$ to produce an output which is a better mean-square approximation to the desired output than would be possible if the untreated message were put into the transmission channel.

The mean-square error between $f_o(t)$ and $f_d(t)$ of Fig. 2 will be

$$\mathcal{E} = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} dt\left[\int_{-\infty}^{\infty} d\sigma\, g(\sigma) f_n(t-\sigma) + \int_{-\infty}^{\infty} d\sigma\, g(\sigma)\int_{-\infty}^{\infty} d\nu\, h(\nu) f_m(t-\sigma-\nu) - f_d(t)\right]^2. \tag{14}$$

When expanded Eq. 14 may be rewritten in terms of correlation functions as

$$\mathcal{E} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\sigma d\nu\, g(\sigma) g(\nu)\phi_{nn}(\sigma-\nu) + 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\xi d\sigma d\nu\, g(\xi) g(\sigma) h(\nu)\phi_{nm}(\xi-\sigma-\nu)$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\psi d\xi d\sigma d\nu\, g(\psi) h(\xi) g(\sigma) h(\nu)\phi_{mm}(\psi+\xi-\sigma-\nu) - 2\int_{-\infty}^{\infty} d\sigma\, g(\sigma)\phi_{nd}(\sigma)$$

$$- 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\sigma d\nu\, g(\sigma) h(\nu)\phi_{md}(\sigma+\nu) + \phi_{dd}(0). \tag{15}$$

Impulse response functions $g(t)$ and $h(t)$ must be found which minimize $\mathcal{E}$ of Eq. 15 and which may be realized by physical networks. That is

$$g(t), h(t) = 0 \quad \text{for } t < 0. \tag{16}$$

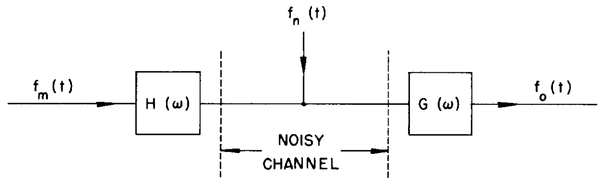

Fig. 2
Transmission circuit.

An additional constraint must be imposed on the coding network H(ω) with regard to average transmitted signal power, but this need not be considered at this point.

As a first approach let us assume that H(ω) is fixed and solve for an optimum G(ω). This may be done by letting g(t) of Eq. 15 take on an admissible variation $\epsilon\eta(t)$ where

$$\eta(t) = 0 \quad \text{for } t < 0 \tag{17}$$

and ϵ is a parameter independent of η and h. That is, we replace g(t) of Eq. 15 by $g(t) + \epsilon\eta(t)$ and replace $\mathcal{E}$ by $\mathcal{E} + \delta\mathcal{E}$. Now if a certain g(t) gives minimum mean-square error, then certainly this optimum g(t) must also satisfy

$$\left.\frac{\partial(\mathcal{E} + \delta\mathcal{E})}{\partial\epsilon}\right|_{\epsilon=0} = 0. \tag{18}$$

An expansion of Eq. 18 by using Eq. 15 results in

$$\int_{-\infty}^{\infty} d\nu\, g(\nu)\phi_{nn}(\sigma-\nu) + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\xi d\nu\, h(\nu)g(\xi)\phi_{nm}(\xi-\sigma-\nu)$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\xi d\nu\, h(\nu)g(\xi)\phi_{nm}(\sigma-\xi-\nu)$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\psi d\xi d\nu\, h(\xi)h(\nu)g(\psi)\phi_{mm}(\psi+\xi-\sigma-\nu)$$

$$- \phi_{nd}(\sigma) - \int_{-\infty}^{\infty} d\nu\, h(\nu)\phi_{md}(\sigma+\nu)$$

$$= q(\sigma) \tag{19}$$

where $q(\sigma)$ is a function defined by

$$q(\sigma) = 0 \quad \text{for } \sigma > 0. \tag{20}$$

Now taking the Laplace transform of Eq. 20 with respect to σ one obtains

$$G(\lambda)F(\lambda) - \Phi_{nd}(\lambda) - H(-\lambda)\Phi_{md}(\lambda) = Q(\lambda) \tag{21}$$

where $F(\lambda)$ is by definition

$$F(\lambda) = H(\lambda)H(-\lambda)\Phi_{mm}(\lambda) + H(\lambda)\Phi_{nm}(\lambda) + H(-\lambda)\Phi_{mn}(\lambda) + \Phi_{nn}(\lambda). \tag{22}$$

We shall assume that $F(\lambda)$ is factorable into

$$F(\lambda) = F^{+}(\lambda) \cdot F^{-}(\lambda) \tag{23}$$

where $F^{+}(\lambda)$ has all its poles and zeros in the upper half of the λ plane, and $F^{-}(\lambda)$ has all its poles and zeros in the lower half of the λ plane. Equation 21 may then be written with the aid of Eqs. 4, 5, and 6 as

$$G(\lambda)F^{+}(\lambda) - \frac{1}{2\pi}\int_{0}^{\infty} \epsilon^{-j\lambda t}\,dt \int_{-\infty}^{\infty} \frac{\Phi_{nd}(w) + H(-w)\Phi_{md}(w)}{F^{-}(w)}\,\epsilon^{+jwt}\,dw$$

$$- \frac{1}{2\pi}\int_{-\infty}^{0} \epsilon^{-j\lambda t}\,dt \int_{-\infty}^{\infty} \frac{\Phi_{nd}(w) + H(-w)\Phi_{md}(w)}{F^{-}(w)}\,\epsilon^{+jwt}\,dw$$

$$= \frac{Q(\lambda)}{F^{-}(\lambda)}. \tag{24}$$

The first two terms on the left-hand side of Eq. 24 have all their poles, if any, in the upper half of the λ plane, while the third term on the left has all its poles in the lower half of the λ plane. Since the right-hand side of Eq. 24 has poles only in the lower-half plane, the first two terms on the left, when taken together, must equal a constant. It can be shown that this constant is zero, thus yielding

$$G(\lambda) = \frac{1}{2\pi F^{+}(\lambda)}\int_{0}^{\infty} \epsilon^{-j\lambda t}\,dt \int_{-\infty}^{\infty} \frac{\Phi_{nd}(w) + H(-w)\Phi_{md}(w)}{F^{-}(w)}\,\epsilon^{+jwt}\,dw. \tag{25}$$

For a fixed $H(\lambda)$ network of Fig. 2, Eq. 25 gives the optimum transfer function for the decoding network. The system function $G(\lambda)$ given by Eq. 25 is always realizable.

Two special cases of Eq. 25 are of interest. First choosing $H(\lambda) = 1$, we have

$$\begin{aligned} F(\lambda) &= \Phi_{mm}(\lambda) + \Phi_{nm}(\lambda) + \Phi_{mn}(\lambda) + \Phi_{nn}(\lambda) \\ &= \Phi_{ii}(\lambda). \end{aligned} \tag{26}$$

Thus, $F(\lambda)$ becomes the Fourier transform of the autocorrelation function of the G network input. We then have for $G(\lambda)$

$$G(\lambda) = \frac{1}{2\pi\Phi_{ii}^{+}(\lambda)}\int_{0}^{\infty} \epsilon^{-j\lambda t}\,dt \int_{-\infty}^{\infty} \frac{\Phi_{id}(w)}{\Phi_{ii}^{-}(w)}\,\epsilon^{+jwt}\,dw \tag{27}$$

where $\Phi_{id}(\lambda)$ represents the cross-power spectrum between the filter input and the desired output. Equation 27 is the optimum filter formula of Wiener and Lee and is the solution to the problem of Fig. 1.

The second special case of interest occurs when the noise function of Fig. 2 is made zero. Then Eq. 25 becomes

$$G(\lambda) = \frac{1}{2\pi H^{+}(\lambda)\Phi^{+}_{mm}(\lambda)} \int_{0}^{\infty} \epsilon^{-j\lambda t} dt \int_{-\infty}^{\infty} \frac{H(-w)\Phi_{md}(w)}{H^{-}(w)\Phi^{-}_{mm}(w)} \epsilon^{+jwt} dw \qquad (28)$$

where

$$H^{+}(\lambda)H^{-}(\lambda) = H(\lambda)H(-\lambda) \qquad (28a)$$

which is the so-called "optimum compensator formula." Equation 28, though previously unpublished, was first derived by Y. W. Lee.

If the network G(w) of Fig. 2 is considered fixed, then by exactly the same methods used above, we will find that the optimum $H(\lambda)$ will be given by

$$H(\lambda) = \frac{1}{2\pi G^{+}(\lambda)\Phi^{+}_{mm}(\lambda)} \int_{0}^{\infty} \epsilon^{-j\lambda t} dt \int_{-\infty}^{\infty} \left[\frac{G(-w)\Phi_{md}(w)}{G^{-}(w)\Phi^{-}_{mm}(w)} - \frac{G^{+}(w)\Phi_{mn}(w)}{\Phi^{-}_{mm}(w)} \right] \epsilon^{+jwt} dw \qquad (29)$$

where

$$G^{+}(\lambda)G^{-}(\lambda) = G(\lambda)G(-\lambda). \qquad (29a)$$

If the channel noise is zero or if the crosscorrelation between the message and noise is zero, Eq. 29 reduces to an optimum compensator formula.

Equation 25 gives the optimum $G(\lambda)$ for a fixed $H(\lambda)$ while Eq. 29 gives the optimum $H(\lambda)$ for fixed $G(\lambda)$. If a simultaneous solution of Eqs. 25 and 29 is performed, one would obtain the optimum coding-decoding pair of networks for the transmission circuit of Fig. 2. However, before such a solution is attempted, it will be convenient to solve for the mean-square error resulting from fixing $H(\lambda)$ and optimizing $G(\lambda)$. Substitution of Eq. 19 into Eq. 15 yields

$$\mathcal{E}_{min}^{(H\ Fixed)} = \phi_{dd}(0) - \int_{-\infty}^{\infty} d\sigma g(\sigma)\phi_{nd}(\sigma) - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\sigma d\nu g(\sigma)h(\nu)\phi_{md}(\sigma+\nu). \qquad (30)$$

This equation may be rewritten in terms of Fourier transforms as

$$\mathcal{E}_{min}^{(H\ Fixed)} = \int_{-\infty}^{\infty} d\omega \left\{ \Phi_{dd}(\omega) - \Phi_{nd}(\omega)G(-\omega) - G(-\omega)H(-\omega)\Phi_{md}(\omega) \right\}. \qquad (31)$$

It must be remembered that $g(\sigma)$ and $G(\omega)$ of Eqs. 30 and 31 are not arbitrary but are solutions of Eq. 19.

III. Long-Delay Solution

In solving for the optimum pair of networks of Fig. 2, we shall assume that the crosscorrelation between the message and the channel noise is zero. We shall further accept a long-delay solution; that is, the desired output shall be the message delayed. Thus

$$f_d(t) = f_m(t-a), \quad a \to \infty \tag{32}$$

and

$$\Phi_{md}(\omega) = \Phi_{mm}(\omega)\,\epsilon^{-j\omega a}, \quad a \to \infty. \tag{33}$$

Under these conditions, Eq. 25 may be rewritten as

$$G(\omega) \to \frac{H(-\omega)\Phi_{mm}(\omega)\,\epsilon^{-ja\omega}}{|H(\omega)|^2\Phi_{mm}(\omega) + \Phi_{nn}(\omega)}, \quad a \to \infty. \tag{34}$$

Since best possible filtering results when long delays are permitted (3), substitution of Eq. 34 into Eq. 31 will result in the lowest possible error, the so-called irremovable error. Thus, we have finally

$$\mathcal{E}_{irr}^{(H\ Fixed)} = \int_{-\infty}^{\infty} \frac{\Phi_{mm}(\omega)\Phi_{nn}(\omega)}{|H(\omega)|^2\Phi_{mm}(\omega) + \Phi_{nn}(\omega)}\,d\omega. \tag{35}$$

Note that the irremovable error is dependent only upon the magnitude of the transfer function of $H(\omega)$, not upon the phase. Equation 34 shows that any phase contribution due to a fixed $H(\omega)$ will be removed by the optimum decoding network $G(\omega)$.

For a given $H(\omega)$, Eq. 35 will give the resulting transmission error provided network $G(\omega)$ is designed according to Eq. 34. Thus, we must find an $H(\omega)$ which makes the error of Eq. 35 a minimum and at the same time keeps the average transmitted signal power constant. That is, we must require that

$$\int_{-\infty}^{\infty} |H(\omega)|^2\Phi_{mm}(\omega)\,d\omega = c_1 \tag{36}$$

since $|H(\omega)|^2\Phi_{mm}(\omega)$ represents the power-density spectrum of the output of the coding network. If we let

$$[y(\omega)]^2 = |H(\omega)|^2\Phi_{mm}(\omega) \tag{37}$$

Eqs. 35 and 36 may be rewritten as

$$\frac{\mathcal{E}_{irr}}{2} = \int_0^{\infty} \frac{\Phi_{mm}(\omega)\Phi_{nn}(\omega)}{[y(\omega)]^2 + \Phi_{nn}(\omega)}\,d\omega \tag{38}$$

and

$$\int_0^{\infty} [y(\omega)]^2\,d\omega = c_{1/2}. \tag{39}$$

We now seek that real-valued function $y(\omega)$ which minimizes Eq. 38 and in addition satisfies Eq. 39. This is the so-called isoperimetric condition of the calculus of variations. By applying the usual techniques (9), one obtains

$$|H(\omega)|^2 \Phi_{mm}(\omega) = -\Phi_{nn}(\omega) + \frac{1}{\sqrt{\gamma}} \sqrt{\Phi_{mm}(\omega) \Phi_{nn}(\omega)} \tag{40a}$$

and

$$|H(\omega)|^2 \Phi_{mm}(\omega) = 0 \tag{40b}$$

where γ is a constant which must be adjusted to satisfy Eq. 39. Equation 40a is used where the right-hand side is positive, otherwise Eq. 40b must be used. Physically this means that $H(\omega)$ may contain stop bands; however, the existence of such stopbands is in no sense a violation of the Paley-Wiener theorem (1, 10, 11) as an infinite delay time through $H(\omega)$ and $G(\omega)$ is assumed.

IV. Discussion of Results

As a check on the results of section III, a noise spectrum

$$\Phi_{nn}(\omega) = a^2 \tag{41}$$

and a message spectrum

$$\Phi_{mm}(\omega) = \frac{\frac{\beta}{\pi}}{\omega^2 + \beta^2} \tag{42}$$

were assumed. When $2a^2\beta$ was chosen equal to $1/5\ \pi$ and c_1 was made unity, Eq. 38 gave a mean-square error of 0.285. Without coding and using the same average signal power, the mean-square error was found to be 0.302. (This was computed from Eq. 35 by setting $|H(\omega)|^2$ equal to unity.) Thus, the optimum coding network gave some transmission improvement but not a considerable amount. For the particular case cited, no transmission is allowed by $H(\omega)$ beyond $\omega = 8.45\beta$, and when the noise level was raised by a factor of five, this upper cut-off frequency moved down to $\omega = 3.25\beta$.

It can be shown that a communication circuit using amplitude modulation can be represented by Fig. 2 (ref. 4). The use of frequency modulation complicates matters but it has been suggested that noise spectra of the form

$$\Phi_{nn}(\omega) = a^2\omega^2 \tag{43}$$

might be meaningful. Using the message spectrum of Eq. 42 and the noise spectrum of Eq. 43, it was found that only a moderate improvement resulted using optimum coding or "pre-emphasis" networks.

The moderate improvement in mean-square error shown above is due in part to the particular spectra assumed. Therefore, in certain instances a considerable improvement might be realized by using proper coding networks.

References

1. N. Wiener: The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications, John Wiley, New York, 1949
2. N. Wiener: Cybernetics, John Wiley, New York, 1948
3. Y. W. Lee: Course 6.563, Statistical Theory of Communication, Class Notes, M.I.T. unpublished
4. J. P. Costas: Interference Filtering, Technical Report No. 185, Research Laboratory of Electronics, M.I.T. March 1951
5. Y. W. Lee: Memorandum on Error Criteria in Filter Design, unpublished
6. N. Wiener: Generalized Harmonic Analysis, Acta Math. 55, 117-258, 1930
7. Y. W. Lee, C. A. Stutt: Statistical Prediction of Noise, Technical Report No. 129, Research Laboratory of Electronics, M.I.T. July 1949
8. C. A. Stutt: An Experimental Study of Optimum Linear Systems, Doctoral Thesis, Dept. of Electrical Engineering, M.I.T. May 15, 1951
9. H. Margenau, G. M. Murphy: The Mathematics of Physics and Chemistry, Van Nostrand, New York, 1943
10. G. E. Valley, Jr., H. Wallman: Vacuum Tube Amplifiers, McGraw-Hill, New York, 1948
11. R. Paley, N. Wiener: Fourier Transforms in the Complex Domain, Am. Math. Soc. Colloq. Pub. 19, 1934

Acknowledgment

The author wishes to thank Professor Y. W. Lee, Dr. C. A. Stutt, and Mr. C. A. Desoer of the Research Laboratory of Electronics, M.I.T. for their many helpful comments and suggestions.