COMMUNICATION UTILIZING FEEDBACK CHANNELS

by

THEODORE JOSEPH CRUISE

S.B., Massachusetts Institute of Technology
(1965)

S.M., Massachusetts Institute of Technology
(1966)

E.E., Massachusetts Institute of Technology
(1966)

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

January, 1969

Signature of Author ______________________________
Department of Electrical Engineering, January 24, 1969

Certified by ______________________________
Thesis Supervisor

Accepted by ______________________________
Chairman, Departmental Committee on Graduate Students

COMMUNICATION UTILIZING FEEDBACK CHANNELS

by

THEODORE JOSEPH CRUISE

Submitted to the Department of Electrical Engineering on January 24, 1969 in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

Communication at essentially error-free rates approaching channel capacity has always involved complex signalling systems. Recently it has been noted that this complexity can be removed at the expense of a noiseless feedback channel from the receiver back to the transmitted. Even simple linear modulation schemes with feedback can signal at error-free information rates approaching channel capacity for a white noise forward channel. Such feedback systems and their characteristics have been analyzed for both digital and analog communications problems. The optimum linear feedback system is given for both situations.

The addition of feedback channel noise makes the communications model more realistic and has also been studied. The optimum linear system remains undetermined for noisy feedback; a class of suboptimal feedback systems yield asymptotically optimal noisy feedback systems and have been studied. The results indicate that linear feedback systems in the presence of feedback noise do provide some performance improvement, but not nearly as much improvement as noiseless feedback.

Also included is the derivation of the channel capacity of a white noise channel with a mean square bandwidth constraint on the transmitted signal. This result is then used to compare angle modulation performance to the rate-distortion bound.

THESIS SUPERVISOR: Harry L. Van Trees
TITLE: Associate Professor of Electrical Engineering

ACKNOWLEDGMENT

I wish to thank Professor Harry L. Van Trees for his continued assistance during my thesis research and throughout my years at M.I.T.

I am also indebted to Professors Robert G. Gallager and Donald L. Snyder who served on my thesis committee.

Elaine Imbornone did an excellent job of typing this thesis.

This work was supported by a National Science Foundation Fellowship.

TABLE OF CONTENTS

To Sherry

My favorite feedback communicator

CHAPTER 1

Introduction to Feedback Communication

In the design of communications systems much effort is devoted to designing systems which perform as well as possible or as well as needed in a particular application. For example, a system operating over a white noise channel has an untimate error-free information rate attainable given by channel capacity; however, systems which signal at rates approaching channel capacity tend to be very complex and involve coding for useful system performance (see Wozencraft and Jacobs [20]). In all practical applications some errors are allowed and a system must be designed to attain the specified performance. If the performance desired is not too severe, a simple system will achieve the desired performance. More often, a simple system is not adequate and the complexities of coding (or other complexities) are necessary to achieve the desired performance. For the most part this thesis is concerned with this latter problem, achieving some specified performance when a simple signalling scheme is not adequate.

Introducing coding complexities will always improve the system performance, but the cost of the coding-decoding apparatus may be great. Recently several authors[4,7,9] have studied the utilization of a feedback link as a means of improving communication over the forward channel. The advantage of such a feedback system is that performance comparable to coding (without feedback) is attainable without the complexities of coding; the main disadvantage, of course, is the addition of the feedback channel (an extra transmitter, receiver, etc.). A feedback channel, then, offers an alternate approach (to coding) to the system designer as a means of improving the performance over the forward channel. Whether or not a feedback system is less expensive (than coding, say) depends on the application. Strictly speaking, feedback systems without coding (the topic of this thesis) should be compared with feedback systems with coding as far as performance and complexity is concerned. Unfortunately such results for coded feedback systems have not yet appeared in the literature.

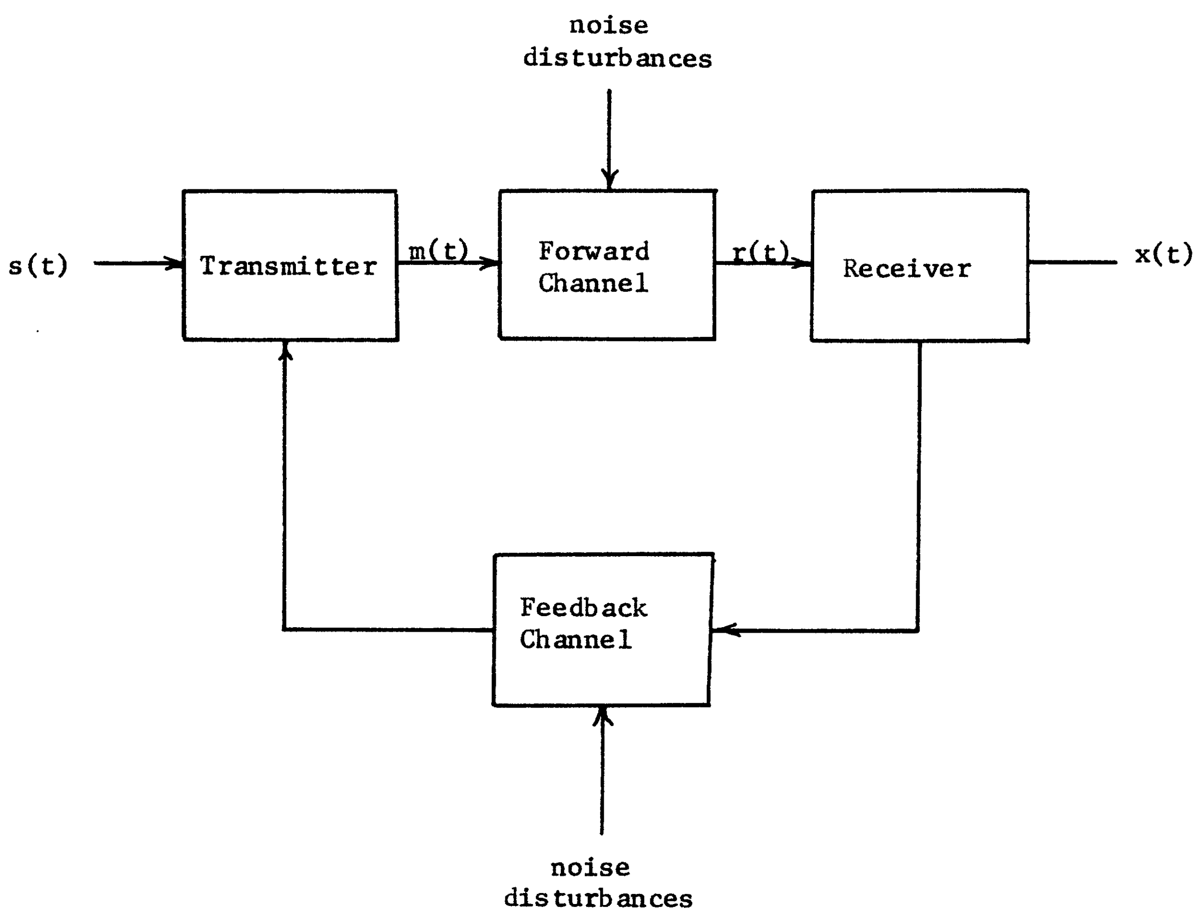
Another advantage of feedback systems is that frequently a simple system with feedback will perform better than a complex coded system without feedback. In other applications where space and/or power may be at a minimum (e.g., in satellite communication) feedback may offer the only solution to system improvement. Feedback can also be added to a completed nofeedback system to improve its performance; even a coded system could be improved (with slight modification) by adding feedback.

The application of feedback can take many forms and consequently give differing levels of system improvement. In a coded system, for example, feedback might only be used to inform the transmitter of each bit received; the transmitter would then alter the transmitted signal according to the incorrect bits received. A more complicated feedback system would continuously inform the transmitter of the "state" of the receiver throughout the baud interval. This second system clearly uses more feedback information and would be expected to offer more improvement than the first. Green[1] distinguishes between these two applications of feedback; the first is called post-decision feedback and the second pre-decision feedback. Obviously post-decision feedback will not give more improvement than pre-decision feedback; but then, it will also be less expensive in terms of complexity.

Thus far, discussion has been limited to digital or coded systems which transmit a single bit or more generally one of a discrete set of messages. Another application of feedback is to analog communications systems. The distinction between analog and digital communication is mainly the distinction between systems with a fidelity criterion (e.g., mean square error) and those with a probability of error (P_e) criterion. Such applications to analog systems are also of interest and are treated in this thesis. Analog communication involves no decisions, but uses a continuous-time or pre-decision feedback for lack of a better word. This thesis is primarily concerned with all types of pre-decision or continuous-time feedback.

1.1 General Feedback Communication System

In Figure 1-1 a block diagram of a general feedback communication system is shown. The information to be conveyed over the forward channel can take any form (digital or analog) depending on the application. For example, the channel might be used to transmit a single bit in T seconds, 20 bits in T seconds, a single bit sequentially, the value of a random variable, or a segment of a



m(t) = transmitted signal

r(t) = forward channel output

x(t) = receiver output (state)

z(t) = feedback signal

y(t) = feedback channel output

Figure 1-1. General feedback communication system random process. At the transmitter the signal s(t) which contains the message is combined with the feedback signal y(t) to generate the transmitted signal m(t). The forward channel could be an additive white noise channel or could contain more complicated disturbances; the results of this thesis are concerned primarily with a white noise forward channel.

The output of the forward channel r(t) is the input to the receiver. The receiver attempts to recover the message from the observed r(t) and also generates the return signal z(t), the input to the feedback channel. The feedback signal z(t) is corrupted by the feedback channel disturbances which could be additive noise, delay, or other types of interference. In many cases this feedback channel will be assumed to be noiseless and without delay. In other words y(t) = z(t).

Typically there are realistic constraints imposed on the system structure or system signals. The transmitted signal m(t) must have a power (peak and/or average) constraint and similarly for z(t) if the feedback channel is noisy. Transmitted signal bandwidth constraints might also be imposed. Given the necessary system constraints, the system must be designed to maximize the overall performance whether the criterion be probability of error for digital systems or mean square error for analog systems.

Although the signals shown in Figure 1-1 are functions of the continuous variable time, most authors who have studied feedback systems previously have studied discrete-time forms of the continuous-time system of Figure 1-1. For the discrete-time version the transmitted signal m(t) becomes m_k (a function of the integer k) where k represents the k-th sample in time or the k-th coordinate of some other expansion. Depending on the expansion employed, such discrete systems may or may not be easily implemented in practice. Even from an analytical point of view such discrete formulations are not always tractable. The analytical comparison of the two analysis procedures is the difference between sums and integrals, difference equations and differential equations. This thesis will treat continuous-time systems except for the following discussion of previous investigations. Most authors subsequently apply their discrete results to continuous-time systems; hence, by always dealing with continuous-time signals such limiting procedures are avoided.

1.2 Summary of Previous Study of Feedback Communications

One of the earliest summaries of feedback communications systems is given by Green [1]. Besides discussing some practical applications of the use of feedback, Green includes a paper by Elias [2] which describes a pre-decision feedback system. Elias describes a system which is able to transmit at the channel capacity of a white noise channel of bandwidth $W_c = k\ W_s$ (W_s = source bandwidth, k = integer) by utilizing a noiseless feedback channel of the same bandwidth. Shannon [5] has shown that even the availability of noiseless feedback does not alter the ultimate error-free transmission rate of the forward channel; hence, throughout this thesis feedback will never improve the ultimate rate of channel capacity, but perhaps make operation at rates approaching capacity easier to achieve.

Elias achieves channel capacity by breaking up the wide-band channel into k separate channels interconnected with k-1 noiseless feedback channels. For Elias operating at channel capacity implies that the suitably defined output signal-to-noise ratio is at the maximum value prescribed by channel capacity. Such a system is said to achieve the rate-distortion bound on mean square error for analog systems although Elias omits reference to the rate-distortion bound. Elias [3] has extended his work to networks of Gaussian channels.

Schalkwijk and Kailath [4] have adapted a stochastic approximation procedure to form a noiseless feedback scheme which can operate at error-free rates up to the ultimate rate given by the forward channel capacity. In their system a message space is defined and a probability of error (P_e) calculated. For message rates less than channel capacity P_e tends to 0 in the limit as the number of messages and the length of the signalling interval increase. Such behavior is usually what is meant when a digital system is said to achieve or approach channel capacity.

Schalkwijk and Kailath consider a discrete time system operating over a T second interval with T/N seconds between samples. The message alphabet of M signals consists of M equally spaced numbers θ_i in the interval [-.5,.5]. The receiver decodes the received signal after T seconds to the θ_I which is closest to the final value of the receiver output x_N. The receiver output x_k is fed-back to the transmitter at each time instant over the noiseless feedback channel. The transmitter attempts to drive the receiver output state x_k to the desired message point θ (a particular member of the M θ_i's) by transmitting at the k-th instant

$$m_k = \alpha \, (x_k - \theta) \tag{1.1}$$

The assumed receiver structure is linear and satisfies the difference equation

$$x_{k+1} = x_k - \frac{1}{\alpha k} (m_k + n_k) \qquad x_0 = 0 \tag{1.2}$$

where

α = constant

n_k = additive noise at k-th time instant

$E[n_k n_j] = N_o/2 \quad \delta_{kj}$

The constants α and N are adjusted so that the average power constraint

$$P_{ave} = \frac{1}{T} E[\sum_{i=0}^{N} \alpha^2 \, (x_i - \theta)^2] \tag{1.3}$$

holds and P_e is minimized.

The performance of this system is shown to have a P_e which tends to 0 at information rates less than

$$C = P_{ave}/N_o \text{ nats/sec} \tag{1.4}$$

as T, M, and N all go to infinity in a prescribed manner. C is the channel capacity for the infinite bandwidth forward channel with or without feedback. The information rate R in nats/sec is defined as

$$R = \frac{1}{T} \ln (M) \tag{1.5}$$

For finite values of T, M, and N Schalkwijk and Kailath's system gives a lower P_e than that obtained for block coding (without feedback). In other words even though both systems have a P_e which approaches 0, the feedback scheme approaches 0 much more rapidly. The feedback system is also structurally simpler and does not involve complex coding-decoding algorithms for the messages.

Schalkwijk [6] in a companion paper shows how to modify the wideband scheme for use over bandlimited channels. A bandlimited channel for bandwidth W_c implies that (for the above scheme)

$$N \simeq 2W_c T \tag{1.6}$$

and that α (in Equation 1.2) becomes a function of k (time). The modified system then achieves error-free transmission (in the limit) at rates up to the bandlimited channel capacity

$$C_{W_c} = W_c \ln(1 + \frac{P_{ave}}{N_o W_c}) \tag{1.7}$$

An important assumption of these two papers is noiseless feedback. In a practical situation there always exists some noise in any system. Both of the above papers calculate the performance of the feedback systems if noise is inserted. The performance exhibits a sharp threshold at the point where the feedback noise dominates the overall system performance. No matter how small the feedback noise is (relative to the forward channel noise), eventually P_e tends to 1 as M, N, and T tend to infinity. The conclusion is that the feedback systems described by Schalkwijk and Kailath cannot achieve channel capacity if the slightest amount of feedback noise is present. In a practical situation where P_e need not be 0 the feedback noise might or might not be small enough for satisfactory operation of the feedback system. No attempt was made by Schalkwijk and Kailath to take into account in system design possible feedback noise.

Omura [7,8] considers the identical discrete-time problem from a different viewpoint. Assuming an arbitrary one-state recursive filter at the receiver, Omura proceeds to determine the best transmitted signal for that receiver (given the receiver state is fedback) and then to optimize over the arbitrary one-state filter. His arbitrary filter is described by

$$\begin{aligned} x_{k+1} &= x_k + \phi_k x_k + g_k\,(m_k + n_k) \\ x_0 &= 0 \end{aligned} \qquad (1.8)$$

where $\{\phi_k\}$ and $\{g_k\}$ are free parameters to be determined. m_k is the transmitted signal which depends on the noiseless feedback signal x_{k-1} and the message point θ; the exact dependence of the transmitted signal on these two inputs is optimally determined. The optimization for $\{m_k\}$, $\{\phi_k\}$, and $\{g_k\}$ can be formulated using dynamic programming and then solved.

The optimal transmitter structure is linear (of the same form as Equation 1.1). For any arbitrary set $\{\phi_k\}$ the optimization yields a particular set $\{g_k\}$ such that all of these systems have identical performance. Omura's system differs slightly from Schalkwijk and Kailath's in that Omura's has a constant average power

$$E[m_k^2] = P_{ave} \qquad \text{(Omura)} \tag{1.9}$$

whereas Schalkwijk and Kailath have a time-varying instantaneous average power

$$E[m_k^2] \sim \frac{1}{k} \qquad \text{(Schalkwijk)} \tag{1.10}$$

Both, of course, satisfy the average power constraint Equation 1.3, but in different ways. Both systems have similar (but not identical) performance; Omura's performs better for finite T, M, and N.

Turin [9,10] and Horstein [11] consider a different system utilizing feedback. They are concerned with transmitting a single bit (or equivalently one of two hypotheses) sequentially or non-sequentially. Thus far only nonsequential systems have been mentioned. The receiver of the sequential system computes the likelihood ratio of the two hypotheses, H_+ and H_-; the ratio is also fedback to the transmitter over a noiseless feedback link. For sequential operation the system is allowed to continue until the likelihood ratio at the output of the receiver reaches one of the two thresholds, Y_+ and Y_-. The time required for each bit to be determined at the receiver will fluctuate, necessitating some data storage capabilities. If the system is operated nonsequentially, the receiver chooses the most likely hypothesis at the end of the fixed transmission interval.

For Turin and Horstein the receiver (likelihood ratio computer) is fixed and the optimal transmitted signal to be determined. In particular they require the transmitted signal to be of the form

$$m_\pm(x,t) = \pm U\ (x)\sigma(t) \tag{1.11}$$

where x = likelihood ratio = receiver output.

The signal transmitted under either hypothesis is the product of a time function $\sigma(t)$ and a weighting $U_\pm(x)$ due to the current state of the receiver. A peak-to-average power ratio is defined

$$\alpha = P_{peak}/P_{ave} \tag{1.12}$$

and a peak power constraint is applied by varying α. Turin considers $\alpha=1$ and $\alpha \geq -\log_2(P_e) = \alpha'$. Horstein considers the remaining values of α.

For α tending to infinity (i.e., no peak power constraint) the sequential system can operate up to an average error-free rate given by channel capacity. For a given (nonzero) P_e and average time/decision $\bar{T}$ the sequential system has an average power advantage of

$$\alpha' = -\log_2(P_e) \tag{1.13}$$

over the same system without feedback. Such a system without feedback would be equivalent to a nonsequential matched-filter likelihood ratio computer at the receiver.

For a finite peak power constraint it is impossible to operate the system at any nonzero rate with $P_e=0$. Without allowing an infinite peak power neither Turin and Horstein nor Schalkwijk and Kailath can achieve channel capacity; Omura's scheme, however, does not require an infinite peak power.

Kashyap [21] has considered a system similar to Schalkwijk and Kailath's, but with noise in the feedback channel. Kashyap's result is that nonzero error-free information rates are possible for rates less than some $R_c<C$. Unfortunately his technique requires an increasing average power in the feedback channel as T, M, and N increase. Basically the transmitted power in the feedback signal is allowed to become infinite so that the feedback link is really noiseless in the limit and nonzero rates can be achieved. That he could only achieve a rate $R_c<C$ must be attributed to his not letting the feedback channel power get large enough fast enough.

Kramer [22] has adapted feedback to an orthogonal signalling system. Orthogonal signalling systems (unlike the linear signalling systems treated thus far) will operate at rates up to channel capacity without errors without feedback. The addition of feedback cannot improve on this error-free rate, but it does improve on P_e for finite T, M, and N. In fact it is not surprising that orthogonal signalling with feedback is much superior to linear signalling with feedback. Of course, the orthogonal system would be somewhat more complex in terms of transmitter-receiver implementation. For the most part this thesis is not concerned with the addition of feedback to already complex systems; the main advantage of feedback appears to be a saving in complexity at the expense of a feedback channel. However, for some channels (e.g., fading) orthogonal signalling is almost a necessity for satisfactory performance.

Kramer also considers noisy feedback, but like Kashyap, lets the feedback channel power approach infinity so that the noise in the feedback link "disappears" allowing capacity to be achieved in the same manner as his noiseless system.

Butman [23,24] has formulated the general linear feedback problem similar to Omura's. Butman assumes a linear transmitter as well as receiver and optimizes over these two linear filters; here a linear receiver is assumed (as with Omura) and the optimal transmitter is shown to be linear. For noiseless feedback Butman's discrete-time system performs better than Omura's, but Omura's system can be made equivalent to Butman's in performance by removing some of Omura's approximations. For noisy feedback Butman has some results for suboptimal systems; analytic solution for the optimal system seems impossible. Unfortunately his partial results cannot be extended to the continuous-time systems treated in this thesis. Butman, however, did impose a finite average power constraint on the feedback transmitter and thereby formulated a realistic problem of interest which others have failed to do.

1.3 Outline of Thesis

In the remainder of this thesis noiseless and noisy feedback systems are studied employing a continuous-time formulation of the problem. The primary concern is not so much with achieving capacity of the forward channel, but with minimizing either the probability of error (P_e) or the mean square error for finite time problems.

Chapter 2 treats the continuous version of Omura's problem with noiseless feedback. Also investigated are the physical characteristics (peak power, bandwidth, etc.) of such noiseless feedback systems.

Chapter 3 treats the topic of analog signalling over noiseless feedback systems. This is a new application of feedback and has not been studied before.

Chapter 4 treats the digital problem in Chapter 2 when there is noise in the feedback link. The results are primarily approximate since analytic solution seems impossible. Nevertheless, such partial results are most useful in systems engineering since the optimal systems (if they could be determined) appear to be not much better than some of the sub-optimum systems studied. Noiseless feedback systems turn out to be very sensitive to the noiseless assumption; noiseless feedback can be viewed almost as a singular system achieving dramatic performance improvement. The addition of noise in the feedback link which makes the system more realistic also cuts down the performance improvement.

Chapter 5 deals with extensions of this work and suggestions for future study.

Chapter 6 contains some results unrelated to feedback systems. They are included for completeness since the results were obtained during my graduate research.

CHAPTER 2

Noiseless Feedback -- Digital

In this chapter a noiseless feedback system will be developed and its performance calculated. The system development is similar to that of Omura [7,8], but the analysis is in terms of a continuous-time variable t instead of a discrete variable k. The mathematics necessary for this formulation involves stochastic differential equations, dynamic programming, and stochastic optimal control theory. No attempt will be made to prove the necessary results from these areas; the reader is directed to the references for further information.

2.1 Definition of Noiseless Feedback System

The receiver structure is assumed to be a simple linear system described by a first order differential equation. The motivation for such a receiver structure is primarily the simplicity and practicality; the prospect of actually having to build the system if it will work satisfactorily is not an unpleasant one. Also, under the Gaussian noise assumption the linear system will turn out to be optimal.

The receiver (by assumption) is an arbitrary one-state linear filter operating on the received signal r(t) in the interval $0 \le t \le T$. The state equation is

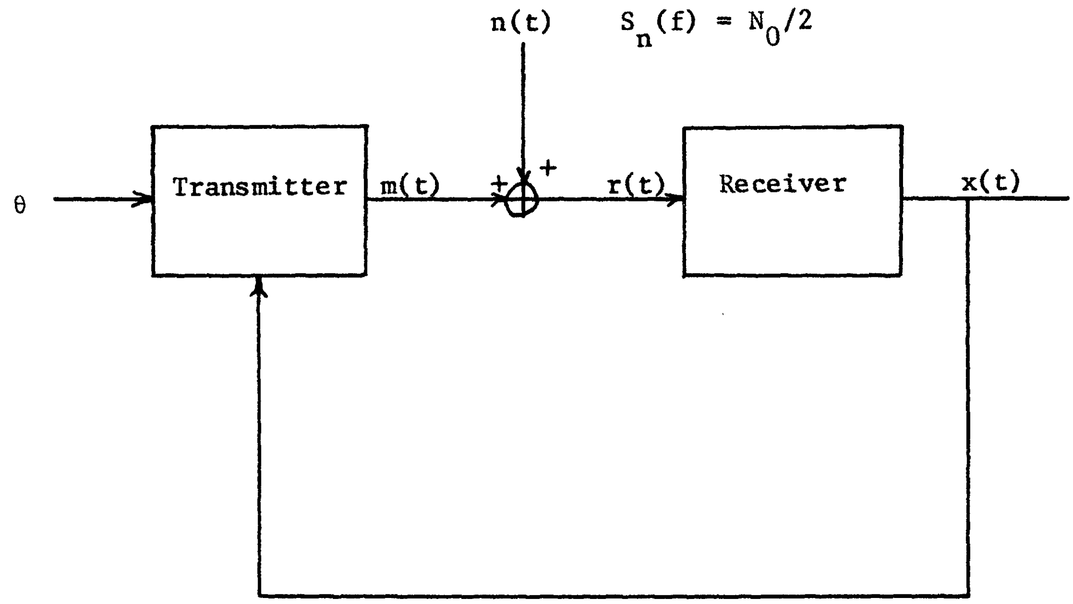
$$\frac{d}{dt} x(t) = \phi(t)\, x(t) + g(t)\, r(t) \qquad (2.1)$$

where $\phi(t)$ and $g(t)$ are to be selected in an optimal manner. A more general linear receiver would be one of higher dimensional state, but the analysis for the one-state system indicates that extra states in the receiver will not improve the system performance; hence, the assumption of a one state receiver does not reduce the ultimate system performance.

The forward channel is an additive Gaussian white noise channel as indicated in Figure 2-1. The feedback channel is noiseless and allows the transmitter to know the state, $x(t)$, of the receiver.

The digital signalling problem consists of transmitting one of M equiprobable messages from the transmitter to the receiver with a minimum probability of error. Assume that the M messages are mapped to M equally spaced points in the unit interval $[-.5,.5]$. The random variable θ takes on the value $-.5 + \frac{i-1}{M-1}$ $(i=1,M)$ depending on which message is transmitted. For this system of coding the transmitter conveys the value of a random variable θ which can be mapped back to the actual message if desired.

The performance criterion for the system is the probability of error (P_e), and ideally this criterion is to be minimized. Unfortunately this criterion is not tractable for selecting the best transmitter structure minimizing P_e. Instead a quadratic criterion is used to optimally select the transmitter structure; the system is designed to minimize the mean square error in estimating θ at time $t=T$.



$$\frac{d}{dt}x(t) = \phi(t)x(t) + g(t)r(t)$$

$$x(0) = 0$$

Figure 2-1. Continuous-time digital feedback system

Several comments about this criterion are appropriate. If the transmitter is assumed linear, forming the minimum variance estimate of at the receiver (and decoding to the nearest message point) is equivalent to minimizing P_e; hence, the solution obtained shortly is the minimum P_e system when the transmitter is constrained to be linear. As will be shown, the best transmitter structure for minimizing the quadratic criterion is linear anyway. In Chapter 3 analog estimation problems are treated; for these problems the criterion is truly a mean square error one so that the results of this chapter are directly applicable.

Besides the message point θ, the transmitter has available x(t), the current state of the receiver. This information is transmitted continuously back to the transmitter over the noiseless (and delayless) feedback link. Knowledge of the state x(t) is sufficient to specify completely all characteristics of the operation of the receiver; hence, any other information supplied over the feedback channel would be redundant. Actually the transmitter also has available the past values of x(t) (i.e., $x(\tau)$ for $0 \leq \tau \leq t$), but these values turn out to be unnecessary. The general structure of the transmitter is arbitrary with $m(t) = f(\theta, x(t), t)$; the optimization implies that f(, ,) is actually linear in the first two arguments.

In the formulation the functions $\phi(t)$ and g(t) which determine the receiver are completely free. In a practical system one or both of these functions might already be specified as part of the system or by cost considerations. Here these two functions will be assumed unconstrained.

If the receiver state at t=T is to be the minimum mean square error estimate of θ, the quadratic criterion to be minimized is

$$\sigma^2 = E[\ (x(T)-\theta)^2\ ] = \text{minimum} \tag{2.2}$$

with the expectation over the forward channel noise and over θ (the message space).

One further constraint remains, that of transmitted power or energy. The transmitted signal m(t) is unspecified, but it must satisfy

$$\int_0^T dt\ E[m^2(t)] \leq E_0 \ = \ P_{ave}T \tag{2.3}$$

as an appropriate transmitter energy constraint. The expectation is over the channel noise and the message space. A constraint on the feedback channel energy has no meaning when the channel is noiseless.

The constraint in Equation 2.3 is only on the average energy used. During any particular T second interval the actual energy used can be more or less than the average E_0. Thus, the transmitter must be able to exceed a transmitted energy of E_0 in T seconds frequently. The average over many intervals (messages), however, is E_0.

Summarizing the problem just formulated, the performance σ^2 in Equation 2.2 is to be minimized subject to the energy constraint in Equation 2.3. The minimization is over the transmitter structure m(t) and the free receiver functions $\phi(t)$ and g(t).

2.2 Stochastic Optimal Control and Dynamic Programming: Formulation

Having specified the problem in the previous section, the solution technique follows by relating the problem to the work of Kushner [12,13]. First of all, some interpretation must be given to systems specified by differential equations with a white noise driving term as in Equation 2.1. Such stochastic differential equations are subject to interpretation according to how one evaluates the limiting forms of difference equations. The two principle interpretations are those of Ito [18] and Stratonovich [19]; the only difference between the two is the meaning of white noise. For this problem, though, the two interpretations are equivalent since the differential equations are linear.

Kushner [12,13] using the Ito interpretation has formulated the stochastic optimal control problem in dynamic programming. Technically Ito differential equations need to be expressed in terms of differentials rather than derivations; throughout this thesis derivative notation will be used for simplicity. Continuing with Kushner's formulation, let the stochastic system to be controlled by specified by a nonlinear vector state equation

$$\frac{d}{dt}\underline{x} = \underline{f}(\underline{x},\underline{u},t) + \underline{\xi}(t) \qquad (2.4)$$

where $\underline{\xi}(t)$ is vector white noise with covariance matrix

$$E[\underline{\xi}(t)\underline{\xi}(u)] = \underline{\underline{S}}(t)\ \delta(t-u) \qquad (2.5)$$

For the communications problem here the state equation is Equation 2.1 or

$$\frac{d}{dt} x(t) = \phi(t)\ x(t) + g(t)[m(t) + n(t)] \tag{2.6}$$

The white noise $\underline{\xi}(t)$ in Equation 2.4 corresponds to g(t)n(t) in Equation 2.6, the covariance function of the latter being

$$E[g(t)n(t)g(u)n(u)] = \frac{N_0}{2} g^2(t)\ \delta(t-u) \tag{2.7}$$

The optimal control problem for Kushner is to determine the control $\underline{u}$ within some control set Ω in the interval of operation [0,T] which minimizes the cost functional

$$J = E_*[\int_0^T dt\ L(\underline{x},\underline{u}) + K(\underline{x}(T))\ ] \tag{2.8}$$

which contains an integral cost plus a terminal cost. The notation $E_*[\ ]$ is the expectation conditioned on all the information which is available to the controller $\underline{u}$. The control variables in the feedback communication problem are m(t), $\phi(t)$, and g(t). $\phi(t)$ and g(t) are simply functions of time, but m(t) is more complicated because it can depend on the feedback signal x(t). The solution proceeds first by determining m(t) (the transmitter structure); then the problem is no longer stochastic control and $\phi(t)$ and g(t) can be found by ordinary means. Therefore, identifying Kushner's control $\underline{u}$ with m(t), the communications problem is to determine the control m(t) within the control set $-\infty<m(t)<\infty$ in the interval [0,T] which minimizes

$$J = E_*[\ \lambda \int_0^T dt\ m^2(t) + (x(T)-\theta)^2\ ] \tag{2.9}$$

The constant λ is a Lagrange multiplier necessary to impose the average energy constraint in Equation 2.3.

This optimal control problem of Kushner differs from the ordinary (deterministic) optimal control problem by the white noise term in the state equation and the expectation $E_*[\ ]$. Deterministic optimal control can be treated by dynamic programming techniques of other techniques derived from Pontryagin [25]; stochastic optimal control cannot be formulated with Pontryagin's method due to the nondeterministic nature of the system state equation.

For the above stochastic problem dynamic programming defines an optimal value cost function

$$V(\underline{x},t) = \min_{\underline{u}\varepsilon\Omega} E_*\ [\ \int_t^T dt\ L(\underline{x},\underline{u}) + K(x(T))\ \ ] \tag{2.10}$$

$V(\underline{x},t)$ is the minimum cost of starting in state $\underline{x}$ at time t and proceeding to the end of the interval. Kushner [12] shows that the function $V(\underline{x},t)$ must satisfy the partial differential equation

$$0 = \min_{\underline{u}\varepsilon\Omega} E_*\{\ \frac{\partial V}{\partial t} + \langle\frac{\partial V}{\partial \underline{x}},\ \underline{f}\rangle + L(\underline{x},\underline{u}) + \frac{1}{2}\ Tr[\underline{\underline{S}}(t)\ \frac{\partial^2 V}{\partial x^2}]\} \tag{2.11}$$

where ordinary matrix notation has been employed to simplify the equation. The boundary condition for the partial differential equation is

$$V(\underline{x},T) = K(\underline{x}(T)) \tag{2.12}$$

The solution of Equation 2.11 for $V(\underline{x},t)$ and $\underline{u}$ is not easy. No general techniques are known for solving such systems just as similar techniques are not available for deterministic control problems.

Proceeding with the parallel development of the communications problem, define the cost functional

$$V(x,t) = \min_{m(t)} E_*[\lambda \int_t^T dt\, m^2(t) + (x(T)-\theta)^2] \tag{2.13}$$

It follows from above that $V(x,t)$ satisfies

$$0 = \min_{m(t)} E_* \{ \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} [\phi(t)x(t)+g(t)m(t)] + \lambda\, m^2(t) + \frac{N_0}{4} g^2(t) \frac{\partial^2 V}{\partial x^2}\} \tag{2.14}$$

subject to the boundary condition

$$V(x,T) = (x-\theta)^2 \tag{2.15}$$

Note that the quantity of interest (J in Equation 2.9) is actually $V(0,0)$. By finding $V(x,t)$ first, $V(0,0)$ follows easily.

Observe that although the control $m(t)$ is written with only a time argument, the control is actually a function of t, $x(t)$, and θ. This will become apparent when $E_*[\ ]$ is evaluated. Also Equation 2.15 implies that θ is fixed and known. Later the results will be averaged over θ to obtain the system performance.

2.3 Solution of Noiseless Feedback System

In this section the solution to the stochastic optimal control problem will be found to determine the dependence of $m(t)$ on $x(t)$ and θ. Following this, the functions $\phi(t)$ and $g(t)$ will be optimized to complete the system.

The conditional expectation $E_*[\ ]$ is conditioned on the fact that the transmitter knows $x(t)$; hence, $E_*[x] = x$. Inserting this fact in Equation 2.14 allows $E_*[\ ]$ to be evaluated, leaving

$$0 = \min_{m(t)} \left\{ \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x}[\phi(t)x(t)+g(t)m(t)] + \lambda m^2(t) + \frac{N_0}{4} g^2(t) \frac{\partial^2 V}{\partial x^2} \right\} \tag{2.16}$$

The minimization over $m(t)$ is just a minimization of a quadratic form in $m(t)$ (for a fixed t). Evaluating Equation 2.16 at its minimum gives

$$0 = \frac{\partial V}{\partial t} - \frac{g^2(t)}{4\lambda} \left(\frac{\partial V}{\partial x}\right)^2 + \frac{N_0}{4} g^2(t) \frac{\partial^2 V}{\partial x^2} + \frac{\partial V}{\partial x} \phi(t)x(t) \tag{2.17}$$

where the minimizing choice of m(t) is

$$m(t) = -\frac{g(t)}{2\lambda}\frac{\partial V}{\partial x} \tag{2.18}$$

Equation 2.18 expresses m(t) in terms of the as yet unknown V(x,t).

Several comments can be made at this point relating stochastic optimal control and deterministic optimal control. In this problem the optimal control m(t) is not affected directly by the noise term in Equation 2.16, namely, $(N_0/4)g^2(t)\partial^2V/\partial x^2$ is independent of m(t). Thus, the solution for m = m(x,t) is the same as would be obtained with no noise present; this problem corresponds to optimal control of linear systems treated similarly in Athans and Falb [14]. In general the addition of the noise term in Equation 2.15 is the only difference in the dynamic programming formulation of stochastic problems. For many problems the solution to the deterministic problem will also be the solution to the stochastic problem <u>if</u> the control is given as a function of the state, not just a function of time. The techniques of Pontryagin are not applicable to stochastic problems since they do not explicitly obtain the control as a function of state.

Returning to the partial differential equation for V(x,t) in Equation 2.17, the solution is not at all obvious for this or most other partial differential systems. Since there is a quadratic cost imbedded in the problem, it is perhaps not unreasonable to expect that $V(x,t)$ is also a quadratic form. Therefore, try a solution of the form

$$V(x,t) = P(t)\,[x - y(t)]^2 + r(t) \qquad (2.19)$$

where $P(t)$, $y(t)$, and $r(t)$ need to be determined. Inserting the above expression for $V(x,t)$ into Equation 2.17 and equating the coefficients in front of x^2, x, and x^0 to zero, there result differential equations which $P(t)$, $y(t)$, and $r(t)$ must satisfy if $V(x,t)$ in Equation 2.19 is to be the solution. The differential equations and boundary conditions for these three functions are

$$\frac{d}{dt} P(t) = \frac{g^2(t)P^2(t)}{\lambda} - 2\phi(t)P(t) \qquad P(T) = 1 \qquad (2.20)$$

$$\frac{d}{dt} y(t) = \phi(t)y(t) \qquad y(T) = \theta \qquad (2.21)$$

$$\frac{d}{dt} r(t) = -\frac{N_0}{2} g^2(t)P(t) \qquad r(T) = 0 \qquad (2.22)$$

By solving these three equations, $V(x,t)$ is determined by Equation 2.19, implying that $m(t)$ is (from Equation 2.18)

$$m(t) = -\frac{g(t)P(t)}{\lambda}\,(x(t) - y(t) \qquad (2.23)$$

which is the desired optimal transmitter structure. Observe that Equations 2.20-22 are easily solved numerically by integrating backwards from $t=T$ where the boundary conditions are given.

Actually in this problem the equations can be integrated analytically. Starting with Equation 2.21, define

$$\Phi(t,\tau) = \exp\left[\int_{\tau}^{t} dv\ \phi(v)\right] \qquad (2.24)$$

as the transition function of Equation 2.21. Applying the boundary condition on y(t) gives

$$y(t) = \theta\ \Phi(t,T) = \frac{\theta}{\Phi(T,t)} \qquad (2.25)$$

as the solution for y(t) in the interval. y(t) represents a type of tracking function for the transmitter; whenever x(t) (the feedback signal) happens to equal y(t), the transmitted signal m(t) is zero. y(t) is that value of x which will cause the receiver to "relax" to $x(T) = \theta$ with no further input starting at state $x = y(t)$ at time t. The additive channel noise will always disturb the receiver state so that the transmitted signal will never be zero for any measurable length of time.

Equation 2.20 is a Ricatti equation for P(t) (without a driving term). The solution can be written

$$P(t) = \frac{\phi^2(T,t)}{1 + \frac{1}{\lambda} \int_{t}^{T} d\tau\ g^2(\tau)\Phi^2(T,\tau)} \qquad (2.26)$$

by employing the boundary condition. Finally the solution of Equation (2.22) yields

$$r(t) = -\frac{N_0}{2}\lambda \ln[P(t)\Phi^2(t,T)] \tag{2.27}$$

Now that $V(x,t)$ has been determined the initial point $V(0,0)$ can be evaluated to give the original functional as

$$V(0,0) = \min_{m(t)} E_*\{ \int_0^T dt\, m^2(t) + (x(T)-\theta)^2$$

$$= P(0)y(0)^2 + r(0)$$

$$= \theta^2 s_0 - \frac{N_0}{2}\lambda \ln(s_0) \tag{2.28}$$

where s_0 is defined as

$$s_0 \equiv P(0)\Phi^2(0,T) \tag{2.29}$$

The overall minimum cost (minimum over the transmitter structure only is only a function of

$$\frac{1}{s_0} = 1 + \frac{1}{\lambda}\int_0^T d\tau\, g^2(\tau)\Phi^2(T,\tau) \tag{2.30}$$

and $N_0/2$, λ, and θ^2. Recall that θ is assumed known until the average over θ is taken later.

The next step in the optimization is to determine $g(t)$ and $\phi(t)$ by minimizing the cost in Equation 2.28 over all possible $g(t)$ and $\phi(t)$. Assuming no constraints on these functions, from Equation 2.30 $g(t)$ and $\phi(t)$ (or equivalently $\Phi(T,t)$) enter together in the cost. Therefore, either $g(t)$ or $\Phi(T,t)$ can be set to 1 without any loss in generality. Set $\Phi(T,t) = 1$ (which implies $\phi(t) = 0$) as the choice leaving the above equation as

$$\frac{1}{s_0} = 1 + \frac{1}{\lambda} \int_0^T d\tau \; g^2(\tau) \tag{2.31}$$

Setting $\phi(t) = 0$ implies that the receiver structure is only a multiplication (or correlation) of the received signal by $g(t)$ and integration of the product; the arbitrary memory allowed originally is not needed. Equivalently the multiplicative function could be set to unity and the receiver structure would only involve the memory term $\phi(t)$. Also one could keep both $g(t)$ and $\phi(t)$ if, say, $\phi(t)$ is a given fixed part of the receiver. All of these variations of the receiver have the same overall performance if their values of s_0 in Equation 2.30 are identical. If a higher state receiver had been assumed initially, the unnecessary redundancy of the extra states would appear at this point in the analysis.

Reverting to $\phi(t) = 0$ and s_0 given by Equation 2.31, $V(0,0)$ is still a functional of the arbitrary function $g(t)$. The optimal $g(t)$ can be found by ordinary calculus of variations. Perturbing Equation 2.28 gives

$$0 = V(0,0) = [\theta^2 - \frac{N_0\lambda}{2s_0}] \; \delta s_0 \tag{2.32}$$

The right hand side is zero if $\delta s_0 = 0$ or if the bracketed term is 0. $\delta s_0 = 0$ implies that $g(t) = 0$ which is an impossible solution; therefore; setting the bracketed term to 0 implies

$$s_0 = \frac{N_0\lambda}{2\theta^2} \tag{2.33}$$

or

$$\int_0^T d_\tau\, g^2(t) = \frac{2}{N_0}^2 - \lambda = \text{constant} \tag{2.34}$$

which is the only restriction on the optimal g(t). Note that no solution for g(t) came out of the perturbation, only the above constraint on the integrated square of g(t). This singular solution implies that there are an infinite number of possibilities for g(t), all of which have some performance as long as Equation 2.34 holds.

At this point there are still several steps remaining to obtain the overall system structure. The multiplier λ needs to be determined such that the average transmitted energy is E_0. This averaging involves averaging over θ also. However, the transmitter structure has been determined as

$$m(t) = -\frac{g(t)P(t)}{\lambda}(x(t) - \theta) \tag{2.35}$$

where P(t) is known (Equation 2.26), λ is an unknown constant to be determined, and g(t) is (almost) arbitrary. The transmitter sends a multiple of the instantaneous error between the receiver state x(t) and the desired state θ.

2.4 Evaluation of the Performance of Feedback System

The previous section found the solution to the noiseless feedback problem in terms of the (almost) arbitrary g(t) and the constants λ and θ. Frequently in optimization problems the solution for the optimum is relatively straightforward, but the actual evaluation of the performance is more difficult; this problem is no exception.

Using the optimal transmitter structure in Equation 2.35, the state equation of the overall system (Equation 2.6)

$$\frac{d}{dt}\, x(t) = -\frac{g^2(t)P(t)}{\lambda}\,(x(t) - \theta) + g(t)n(t) \tag{2.36}$$

$$x(0) = 0$$

for $\phi(t) = 0$. Define the instantaneous error given θ as

$$K(t) \equiv E[(x(t)-\theta)^2] \tag{2.37}$$

The differential equation for K(t) is

$$\frac{d}{dt}\, K(t) = -\frac{2g^2(t)P(t)}{\lambda}\, K(t) + \frac{N_0}{2}\, g^2(t) \tag{2.38}$$

$$K(0) = \theta^2$$

Therefore, the conditional performance (mean square error) of the feedback system is the final value of K(t), namely

$$\sigma^2\big|_\theta = E[(x(T)-\theta)^2\big|_\theta\,] = K(T) \tag{2.39}$$

To relate this performance to the energy used, define

$$E_0(t) \equiv \int_0^t d\tau \; E[m^2(\tau)] \tag{2.40}$$

as the energy used after t seconds. In differential equation form

$$\frac{d}{dt} E_0(t) = \frac{g^2(t)P^2(t)}{\lambda} K(t) \tag{2.41}$$

$$E_0(0) = 0$$

The remaining differential equation to specify the performance calculation is that for P(t) given in Equation 2.20 (for $\phi(t)=0$). In order to have all boundary conditions at $t=0$, integrating backwards in Equation 2.20 gives

$$P(0) = s_0 = \frac{N_0 \lambda}{2\theta^2} \tag{2.42}$$

Solution of the three equations for K(t), $E_0(t)$, and P(t) can take many forms. Since g(t) is arbitrary except for the integral square constraint in Equation 2.34, a fixed g(t) could be selected and the equations integrated numerically or analytically. For this problem analytical integration of these equations for an arbitrary g(t) is possible; this procedure was the original solution technique.

In view of the answer obtained, the following derivation is shorter. Recall that the analysis is still conditioned on a fixed known θ. Multiplying Equation 2.38 by $P^2(t)$ and rearranging terms yields

$$\frac{dK}{dt} P^2(t) + 2 \frac{g^2(t)P^3(t)}{\lambda} K(t) = \frac{N_0}{2} g^2(t)P^2(t) \tag{2.43}$$

Inserting the expression for dP/dt gives

$$\frac{d}{dt} [K(t)P^2(t)] = \frac{N_0}{2}\lambda \frac{d}{dt} P(t) \tag{2.44}$$

Integrating and using the initial conditions implies

$$K(t)P^2(t) = \frac{N_0}{2} \lambda P(t) \tag{2.45}$$

which further implies that Equation 2.41 can be written

$$\frac{d}{dt} E_0(t) = - \frac{N_0}{2} \frac{1}{K(t)} \frac{d}{dt} K(t) \tag{2.46}$$

Now the energy and performance are directly related; integrating gives

$$K(t) = \theta^2 \exp[-2E_0(t)/N_0] \tag{2.47}$$

If λ is chosen properly, then $E_0(T) = E_0$ and

$$\sigma^2\Big|_\theta = \theta^2 \exp[-2E_0/N_0] \tag{2.48}$$

as the conditional performance in terms of the allowed energy. The absolute performance is obtained by averaging over θ (the message space) to give the value of the minimum of Equation 2.2 as

$$\sigma^2 = E[\theta^2] \exp[-2E_0/N_0) \qquad (2.49)$$

Unfortunately the solution leading to Equation 2.49 does not contain some of the details of the system structure, such as the value of λ and the constraint on $g(t)$. The most direct (although cumbersome) way to obtain the value of λ is to integrate Equation 2.41, equate $E_0(T)$ to E_0, and evaluate λ as

$$\lambda = \frac{2\ E[\theta^2]}{N_0} \exp[-2E_0/N_0] = \frac{2\ E[\theta^2]}{N_0} s_0 \qquad (2.50)$$

which implies the only constraint on $g(t)$ is

$$\int_0^T dt\ g^2(t) = \frac{2\ E[\theta^2]}{N_0} (1 - \exp[-2E_0/N_0]) \qquad (2.51)$$

Observe that the parameter which was defined in Equation 2.29 is the fractional mean square error (or normalized mean square error)

$$\frac{\sigma^2}{E[\theta^2]} = s_0 = \exp[-2E_0/N_0] \qquad (2.52)$$

The performance in Equation 2.52 is the fractional mean square error for estimating any random variable θ since the probability density of θ has not entered the analysis. Implicitly θ has a zero mean and a finite variance.

2.5 Probability of Error for Linear Coding of Messages

In Section 2.1 the mapping from the message space (M equiprobable messages) to the random variable θ was outlined. Here this mapping will be used to calculate the probability of error (P_e) for the digital signalling scheme.

The receiver decodes the terminal state x(T) into whichever message is most probable. For M equiprobable messages the output space for x(T) (the values X(T) may take) can be broken into uniform width cells (except for the end cells near ± .5) corresponding to the M possible messages. If x(T) falls into the i-th cell, θ_i is the most probable value of θ and the i-th message is the most probable message.

Assume that a particular θ_i is sent. The probability of error given θ_i sent is approximately the average (over all messages) P_e of the whole system; the only difference is that the endpoint messages $\theta_i \pm .5$ have slightly lower conditional probability of error. Henceforth, this conditional P_e given θ_i sent will be treated as the average P_e for the system; it is negligibly higher than the true average P_e.

For a particular θ_i if n(t) is Gaussian, then x(T) is a Gaussian random variable. From the previous section given θ_i, then

$$E[(x(T)-\theta_i)^2] = \theta_i^2\, s_0 = \theta_i^2 \exp[-2E_0/N_o] \qquad (2.53)$$

By appropriate manipulation of the system differential equations the mean value of the difference is

$$E[(x(T)-\theta_i)] = - \theta_i \, s_0 \tag{2.54}$$

which implies that x(T) is a biased estimate of θ_i. Combining the above two equations gives the variance of the Gaussian random variable x(T) as

$$\mathrm{Var}[x(T)|_{\theta_i}] = \theta_i^2 \, s_0(1 - s_0) \tag{2.55}$$

On the average the variance is

$$\mathrm{Var}[x(T)] = E[\theta^2] \, s_0(1 - s_0) \equiv \sigma_1^2 \tag{2.56}$$

Although the Var[x(T)] really is not the same for each i, for purposes of analysis it will be assumed to be the constant σ_1^2 above. Another approach would be to upper bound θ_i^2 by its maximum value of .25 to remove the i dependence. The resulting P_e would be an upper bound not significantly different from P_e calculated using σ_1^2.

The transmitter message space [-.5,.5] is compressed by a bias factor $(1 - s_0)$ at the receiver; thus, whereas the message points are 1/(M-1) apart at the transmitter, they are only $(1-s_0)/(M-1)$ apart in the receiver space. If θ_i is sent, the receiver will make the correct decision if

$$(1-s_0)(\theta_i - \frac{1}{2(M-1)}) \quad < \quad x(T) \quad < (\theta_i + \frac{1}{2(M-1)})(1-s_0) \qquad (2.57)$$

The probability of error is the probability of exceeding the above cell and can be written

$$P_e = P_e\Big|_i = 2 \int_{\frac{1-s_0}{2M-2}}^{\infty} dz \; \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp[-z^2/2\sigma_1^2] \qquad (2.58)$$

Defining

$$Q(v) = \int_v^{\infty} dz \; \frac{1}{\sqrt{2\pi}} \exp[-z^2/2] \qquad (2.59)$$

P_e becomes

$$P_e = 2\, Q(\frac{1}{2(M-1)\; E[\theta^2]} (\frac{1}{s_0} - 1)^{1/2}) \qquad (2.60)$$

or using Equation 2.52

$$P_e = 2\, Q(\frac{1}{2(M-1)\; E[\theta^2]} \; (\exp[2E_0/N_0] - 1)^{1/2}) \qquad (2.61)$$

If the variance $E[\theta^2]$ is approximated by the variance of a random variable with uniform density in the interval [-.5,.5], then $E[\theta^2] = 1/12$ and P_e can be evaluated. A better choice for $E[\theta^2]$ would be the actual variance of θ for the message space assumed; this is done in the next section.

A comparison of P_e above with that obtained by Omura shows that Omura assumes $E[\theta^2] = 1/12$ and that, if the bias $(1-s_0)$ can be ignored, his discrete system has the same P_e when evaluated in the continuous-time limit. The bias can only be ignored for large signal-to-noise ratios $(2E_0/N_0)$; Omura fails to note this fact.

Schalkwijk and Kailath's system output is an unbiased estimate of θ (by their arbitrary choice) and has no such restrictions. Their performance, however, is inferior in the limit. If Schalkwijk and Kailath allowed a biased estimate to be fedback and optimized their system, better performance could be obtained. Recall, however, that they make no optimization attempts in their application of a stochastic approximation theorem.

As noted by Omura, Schalkwijk, and Kailath, P_e goes to zero in a doubly exponential manner for feedback systems. To relate this P_e to channel capacity, using Equations 1.4 and 1.5 to define capacity and rate, the probability of error is (approximately)

$$P_e \simeq 2\, Q\left(\frac{1}{2}\, E[\theta^2]^{-1/2} \exp[(C-R)T]\right) \qquad (2.62)$$

which is Omura's result for unbiased partitioning (P_e when the bias is ignored). A nofeedback system employing block orthogonal coding also has a P_e which goes to zero (for increasing T and

and R<C), but the error is only singly exponential in T. The doubly exponential dependence of the feedback systems implies that for finite T the feedback system will have a lower P_e than the block orthogonal system without feedback. Schalkwijk and Kailath [4] have some curves which indicate the improvement of the feedback system over the block orthogonal system.

As noted earlier, Schalkwijk and Kailath do not use a biased system. From their paper the P_e obtained by them is

$$P_e \simeq 2\, Q(\tfrac{1}{2}\, E[\theta^2]^{-1/2} \exp[(C-R)T]\ .454\ ) \qquad (2.63)$$

which is different from Equation (2.62) by the factor .454. For finite T their system will perform substantially worse than the feedback scheme of this chapter. For example, if the continuous-time feedback system has an error $P_e \simeq 10^{-10}$, the unbiased system (Equation 2.63) would have $P_e \simeq 10^{-2}$ for the same C, T, and R. This difference is a consequence of the fact that feedback signal of Schalkwijk and Kailath is an unbiased estimate of θ.

2.6 Comparison of Feedback System Performance with Butman's Results

Butman [23] assumes a general linear receiver and linear transmitter for the discrete-time feedback problem. His solution for the optimal linear system has the same limiting (discrete to continuous) performance as Equation 2.61. This result is further verification of the fact that the simple one-state receiver assumed in this chapter performs as well as any higher dimensional arbitrary linear receiver.

To rewrite Equation 2.61 so that it conforms to Butman's result requires only the evaluation of $E[\theta^2]$. For the random variable θ as described in Section 2.1, the variance is

$$E[\theta^2] = \frac{M+1}{12(M-1)} \tag{2.64}$$

which for large M is 1/12. Inserting the above expression for $E[\theta^2]$ into Equation 2.61 gives

$$P_e = 2\, Q\left(\left[\frac{3(\exp[2E_0/N_0] - 1)}{M^2 - 1}\right]^{1/2}\right) \tag{2.65}$$

which is Butman's result in the continuous-time limit.

2.7 Performance of Linear Receiver without Feedback

Some idea of the advantage and improvement of the feedback system can be gained by examining the same problem without the feedback link. Given a linear receiver, energy constraint, and cost function (Equation 2.13), determine the best transmitter structure and optimal receiver parameters for minimizing the cost. The solution follows using ordinary calculus of variations. The best transmitter structure is linear in θ, that is,

$$m(t) = \theta\, h(t) \tag{2.66}$$

where h(t) is arbitrary except for energy normalization. The linear receiver is "matched" to the waveform h(t). Rather than demonstrate the approach just outlined for obtaining the solution to the nofeedback problem, the preceding results of the noiseless feedback system can be extended to the nofeedback problem, the preceding results of the noiseless feedback system can be extended to the nofeedback system.

The solution assuming no feedback implies that $E_*[\ ]$ is a different conditional expectation. $E_*[\ ]$ is conditional on the information available at the transmitter; now, without a feedback channel, there are no conditions, namely

$$E_*[x(t)] = \overline{x(t)} = \text{mean of } x(t) \tag{2.67}$$

replaces the previous definition of $E_*[\ ]$. Equation 2.16 now becomes

$$0 = \min_{m(t)} \{E_*[\frac{\partial V}{\partial t}] + g(t)m(t)\ E_*[\frac{\partial V}{\partial x}] + \phi(t)\ E_*[x(t)\ \frac{\partial V}{\partial x}] + \lambda m^2(t) + \frac{N_0}{4}\ g^2(t)\ E_*[\ \frac{\partial^2 V}{\partial x^2}]\} \tag{2.68}$$

The minimization over m(t) proceeds as before giving

$$m(t) = -\frac{g(t)}{2\lambda}\ E_*[\frac{\partial V}{\partial x}] \tag{2.69}$$

$$0 = E_*[\frac{\partial V}{\partial t}] - \frac{g^2(t)}{4\lambda} E_*^2[\frac{\partial V}{\partial x}] + \phi(t) E_*[x(t) \frac{\partial V}{\partial x}]$$

$$+ \frac{N_0}{4} g^2(t) E_*[\frac{\partial^2 V}{\partial x^2}] \tag{2.70}$$

which correspond to Equation 2.18 and 2.17 respectively. The same quadratic form solution will satisfy Equation 2.70 with exactly the same P(t), y(t), and r(t); however, the E_* operation removes the variable x and leaves Equation 2.70 as an ordinary differential equation. V(x,t) has no meaning any more since x is not available to the transmitter. The transmitter structure implied by Equation 2.69, however, is the optimal one which minimizes Equation 2.9.

Inserting the quadratic form for V(x,t) into Equation 2.69 gives

$$m(t) = - \frac{g(t)P(t)}{\lambda} \overline{(x(t)} - \theta\Phi(t,T)) \tag{2.71}$$

as the optimal transmitter structure. At this point in the feedback problem the optimization for g(t) and $\phi(t)$ was carried out by minimizing V(0,0). Here, using the above definition for m(t), the performance and energy of the system can be calculated to form the functional equivalent to V(0,0) for minimization.

The mean value $\overline{x(t)}$ satisfies

$$\frac{d}{dt} \overline{x(t)} = [\phi(t) - \frac{g^2(t)P(t)}{\lambda}] \overline{x(t)} + \theta \frac{g^2(t)P(t)\Phi(t,T)}{\lambda}$$

$$\overline{x(0)} = 0 \tag{2.72}$$

which follows from Equations 2.1 and 2.71. By solving this equation for $\overline{x(t)}$ and using it in the expression for m(t), the overall performance of the nofeedback system can be evaluated as

$$\lambda E_0 = \theta^2 (s_0 - s_0^2) \tag{2.73}$$

$$E[(x(T)-\theta)^2] = \theta^2 s_0^2 + \frac{N_0}{2}\lambda[\frac{1}{s_0} - 1] \tag{2.74}$$

with exactly the same definition of s_0 as before (Equation 2.31). Forming the function J in Equation 2.9 in order to optimize over $\phi(t)$ and g(t) gives

$$\begin{aligned} J &= E[(x(T)-\theta)^2] + \lambda E_0 \\ &= \theta^2 s_0 + \frac{N_0}{2} \lambda[\frac{1}{s_0} -1] \end{aligned} \tag{2.75}$$

The dependence of J on $\phi(t)$ and g(t) is again only through s_0; hence, one can take $\phi(t) = 0$ without loss of generality. Using calculus of variations to determine g(t), the perturbation of J is

$$\delta J = 0 = [\theta^2 - \frac{N_0\lambda}{2s_0^2}] s_0 \tag{2.76}$$

For a meaningful solution the bracketed term must be 0 so that

$$s_0 = (\frac{N_0}{2\theta^2})^{1/2} \tag{2.77}$$

is the optimal value of s_0; g(t) again is almost arbitrary. g(t) is analogous to the matched filter impulse response in that approach to this problem.

The optimal performance is found by solving Equations 2.73 and 2.77 to eliminate λ and s_0. The average performance of the nofeedback system is then

$$E[(x(T)-\theta)^2] = \frac{E[\theta^2]}{1 + (2E_0/N_0)} = E[\theta^2]\, s_0 \qquad (2.78)$$

Proceeding as in the noiseless feedback case to calculate P_e,

$$E[x(T)] = \theta\,(1 - s_0) \qquad (2.79)$$

and

$$\mathrm{Var}[x(T)] = E[\theta^2]\, s_0(1 - s_0) \qquad (2.80)$$

These equations are exactly the same as Equations 2.54-56 in the noiseless feedback problem; the value of s_0 is different, though. All of the arguments for P_e are exactly the same; hence,

$$\text{(nofeedback)} \quad P_e = 2\, Q\left(\frac{E^{-1/2}[\theta^2]}{2(M-1)}\, (2E_0/N_0)^{1/2}\right) \qquad (2.81)$$

Comparing the performance of the two systems with and without feedback, the performance of the nofeedback system is much less than that of the noiseless feedback system except when

$2E_0/N_0$ is small. For these values of the signal-to-noise ratio noiseless feedback offers no improvement (more correctly, negligible) over nofeedback. The lack of exponential dependence of the argument of the Q function on $2E_0/N_0$ in Equation 2.81 implies that the nofeedback system cannot transmit error-free at nonzero information rates.

The purpose of this diversion to the nofeedback system is twofold. First it demonstrates that for very small signal-to-noise ratios ($2CT = 2E_0/N_0 << 1$) feedback is no improvement over no feedback. In this region of operation one need not bother with a feedback system even if the feedback link is available. Second, it demonstrates how to interpret $E_*[\ ]$ to solve another problem (the nofeedback problem) which is very similar to the original noiseless feedback problem. Many equations turned out to be identical except that s_0 (the fractional estimation error) took on different values for the two problems. This technique will be used later to investigate the noisy feedback problem.

2.8 Operational Characteristics of Feedback Systems

Previously the performance of the feedback system has been the only concern. In practice other operational characteristics (e.g., power distribution, bandwidth) are also important in physical systems. Many feedback schemes point out that an infinite peak power is required to achieve capacity (or that a very large peak power is required to achieve some given P_e);

this fact is a severe limitation on any physical communications system. As will be shown shortly, such large peak powers can be avoided by choosing the free function g(t) properly.

Consider the noiseless feedback system for which $\phi(t) = 0$. The transmitted signal is

$$m(t) = -\frac{g(t)P(t)}{\lambda}(x(t) - \theta) \tag{2.82}$$

as before. Since x(t) is a random process, m(t) is also; hence, the instantaneous power $m^2(t)$ is a random variable at any instant of time. Since x(t) is Gaussian, $m^2(t)$ can be arbitrarily large (with some probability) if $E[m^2(t)]$ is large or if $E[m^2(t)]$ is not large, but $m^2(t)$ just happens to fall at a large value. The former case represents a serious problem to a physical transmitter; if the average instantaneous power is large, then with high probability $m^2(t)$ will also be large necessitating frequent power peaks for the transmitter. Even if the average instantaneous power $E[m^2(t)]$ is small, power peaks can occur since $m^2(t)$ can deviate from its mean. Such occurences are unavoidable if feedback is used; if the forward channel noise is statistically unlikely, the receiver will tend toward the wrong message, causing the transmitter (because it knows this) to increase its power in an effort to combat the bad noise sample. For the most part the forward channel noise will be statistically good, causing the transmitted power to be close to its mean. The conclusion is that transmitter power peaks caused by unnatural forward channel noise cannot be avoided, but that power peaks caused by $E[m^2(t)]$ being large should be avoided if possible.

Reverting to the results of Section 2.4 the average value of m(t) given the i-th message sent is

$$E[m(t)\big|_{\theta_i}] = \theta_i \frac{s_0}{\lambda} g(t) = 6N_0\theta_i\, g(t) \qquad (2.83)$$

where

$$s_0 = \frac{N_0}{2E[\theta^2]} = 6N_0\lambda \qquad (2.84)$$

for $E[\theta^2] = 1/12$. Similarly the conditional instantaneous power is

$$E[m^2(t)\big|_{\theta_i}] = (6N_0\theta_i\, g(t))^2 \frac{P(t)}{P(0)} \qquad (2.85)$$

which is larger than the square of the mean by the factor $P(t)/P(0)$. Observe that the choice of g(t) essentially determines the time dependence of the mean of m(t), but that $g^2(t)P(t)$ determines the mean square value. The preceding paragraph indicates that peaks in $g^2(t)P(t)$ are to be avoided if possible; P(t) depends on g(t) through the differential equation for P(t).

If g(t) is selected as a constant (that constant which satisfies the integral square constraint in Equation 2.31), then the average instantaneous power ($E[m^2(t)]$) is proportional to P(t), a steadily increasing function with a sharp peak at $t = T$. The ratio of this peak at $t = T$ to the average power at the start of the interval is

$$\frac{P(T)}{P(0)} = \exp[2E_0/N_0] \qquad (2.86)$$

which could be quite a large peak power for even reasonable values of the signal-to-noise ratio. In a limiting argument showing that the feedback system will achieve capacity the peak power becomes infinite.

Schalkwijk and Kailath's scheme chooses a transmitted signal which is a constant multiple of the error waveform, namely $m(t) \propto (x(t)-\theta)$; hence, in order for this to be the transmitted signal, $g(t)P(t) =$ constant. Equation 2.85 implies then that the instantaneous average power is proportional to $1/P(t)$ (for another P(t)). This system has roughly the same peak power ratio given in Equation 2.86 except that the peak occurs at the beginning of the interval instead of the end. As Schalkwijk and Kailath noted, the peak power becomes infinite as channel capacity is achieved.

Omura and Butman have shown that the optimal discrete-time system produces a constant average instantaneous power. If g(t) is chosen so that

$$E[m^2(t)] = \frac{E_0}{T} = P_{ave} \qquad (2.87)$$

then this implies that $g^2(t)P(t)$ = constant. The solution for $g^2(t)$ yields

$$g^2(t) = \frac{E_0}{3N_0^2\ T} \exp[-2E_0 t/N_0 T] \tag{2.88}$$

Using this choice of g(t), the complete system is determined; it is drawn in Figure 2-2.

The above choice of g(t) does away with all peaks of $E[m^2(t)]$ and is therefore the best one can do. Power peaks can still occur since $m^2(t)$ is random, but as argued earlier these occur with low probability.

Another aspect of feedback system is the fact that the transmitted energy in any T second interval is also a random variable; unlike transmitting fixed deterministic signals each transmission of a message has some energy which fluctuates about the mean energy E_0. The transmitter as designed here must be able to handle "energy peaks" from time to time. Any transmitter sending a random process must be able to do this.

Wyner [26] has analyzed Schalkwijk and Kailath's system with the constraint that the transmitter can never send more than E_0 energy; the transmitter is turned off if E_0 joules are used before T seconds are up. The performance suffers considerably with this constraint; P_e is no longer doubly exponential in (C-R)T, but only singly exponential like block orthogonal coding

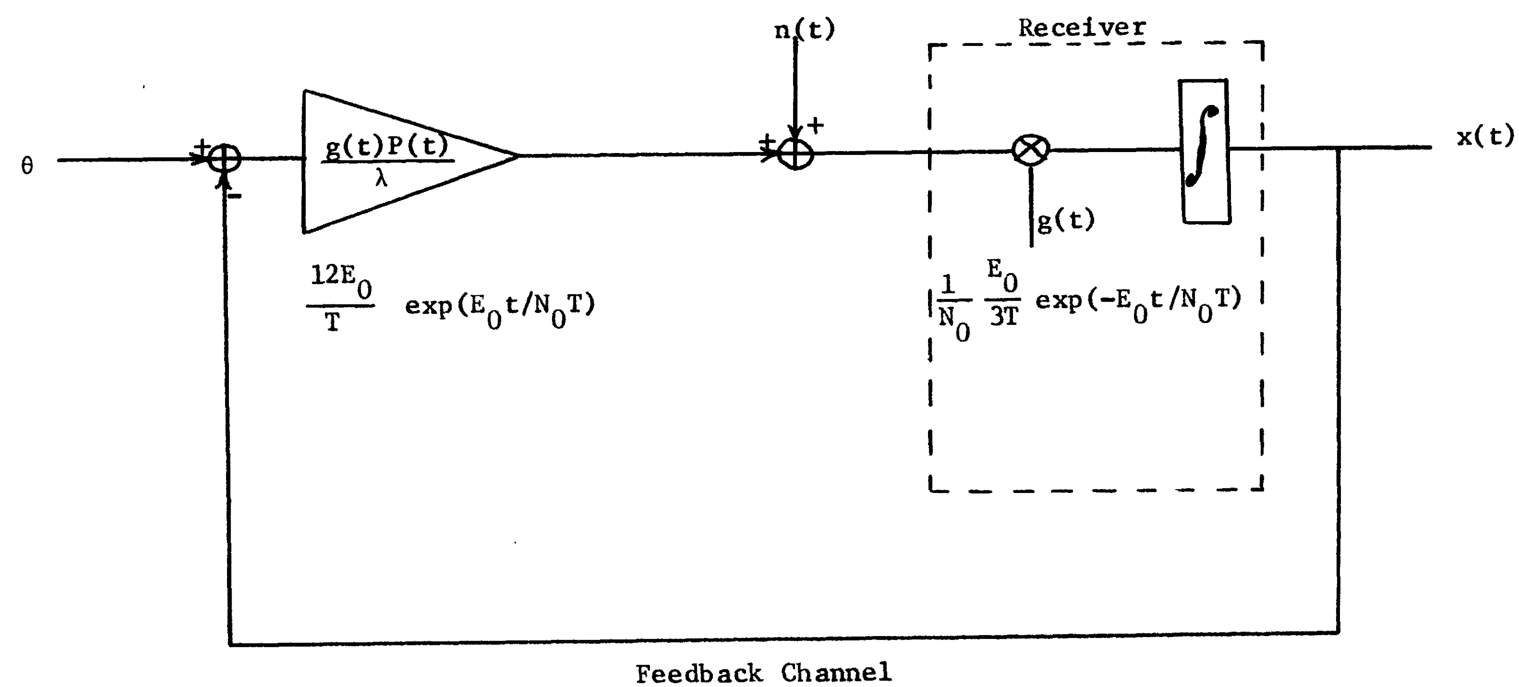


Figure 2-2. Constant average power feedback system

without feedback. As noted, Schalkwijk and Kailath's scheme is suboptimal in many ways. Particularly with the power peak right at the start of the time interval the transmitter uses up most of its allowed energy early in the interval. The transmitted energy of the optimal constant power system would have less tendency to be used up before the end of the interval and, hence, perform better under a strict energy constraint.

The transmitted signal has a bandwidth on the order of the bandwidth of x(t), the receiver state. Equation 2.82 shows that m(t) is a time-varying multiple of the instantaneous error. Since the waveforms are not strictly bandlimited, the bandwidths to be discussed are only approximations in the sense that signal power does actually exist outside the bandwidth of the signal. Most of the signal power, however, is within the indicated bandwidth. A nofeedback system has a bandwidth (at the transmitter) on the order of 1/T. The analogous feedback system will generally have a larger transmitter bandwidth; feedback is analogous to other bandwidth expansion schemes which trade increased bandwidth for improved performance.

The feedback system state x(t) is the output of a one state linear filter driven by white noise. The location of the pole (even though it may change with time) gives an approximate idea of the bandwidth of the process x(t). The transmitted signal m(t) has roughly this same bandwidth.

Consider now the various choices of g(t) and the effects on the transmitter bandwidth. If g(t) = constant is selected, the pole of the x(t) process starts at $(1 - \exp[-2E_0/N_0])/T$ at t=0 and increases to $(1 - \exp[-2E_0/N_0])/T$ times $\exp[2E_0/N_0]$ at t=T. At the beginning of the interval the bandwidth is essentially that of the nofeedback system, but the bandwidth increases rapidly at the end of the interval (at the same time that the power peaks). For Schalkwijk and Kailath's system exactly the time reverse happens, large bandwidths and power at the start of the interval.

By choosing g(t) such that the average transmitted power is constant, the pole of the x(t) process remains constant at $2E_0/N_0T$. Although the feedback system uses $2E_0/N_0$ times the bandwidth of the nofeedback system, this increased bandwidth is still much less than that required for other choices of g(t). Also the bandwidth for this constant power system does not tend to infinity for channel capacity arguments. All of the operational properties of g(t) in Equation 2.88 and in the system diagram Figure 2-2 make it the best choice even though almost any choice of g(t) will have the same P_e.

Rather than setting $\phi(t) = 0$, suppose g(t) = 1 and $\phi(t)$ is the arbitrary function. In this case similar variations are possible for different choices of $\phi(t)$ (or equivalently $\Phi(T,t)$). $\Phi(T,t)$ is the impulse response of the receiver, but not all choices of $\Phi(T,t)$ yield some $\phi(t)$ (i.e., not all linear filters are state realizable). If the particular impulse response

$\Phi(T,t)$ can be achieved with some other realization (other than a one-state system), the receiver and transmitter of the feedback system would perform as desired. In other words any linear receiver structure will make an acceptable feedback system; the receiver need not be a finite state filter as has been assumed.

This completes the analysis of the noiseless feedback system. The continuous-time system analyzed here is related to the many discrete-time systems studied by others. The continuous-time system is very much unrestricted in receiver structure; many realizations are possible all of which have the same performance. The various realizations, however, differ in such characteristics as power distribution and bandwidth.

Throughout a noiseless feedback link has been assumed to be available. It remains to be shown in Chapter 4 exactly how critical this noiseless assumption is. In the next chapter the application of a noiseless feedback link to an arbitrary communications system (as opposed to the particular model studied in this chapter) is made. The utilization is motivated by the results of this chapter.

CHAPTER 3

Noiseless Feedback -- Analog

The previous chapter treated a digital signalling problem for transmitting one of M equiprobable messages over a channel employing feedback. The actual analysis took the form of a parameter (θ) estimation problem by relating the message space to a set of message points θ_i. Conceivable the original communications problem could have been that of transmitting the value of a continuous random variable θ (with some probability density) over the channel. This new problem is simply an extension of Chapter 2. A more general problem would be that of transmitting a random process over the channel using feedback. This problem is the subject of this chapter, that of using noiseless feedback to convey an analog message through the channel. Essentially an analog system is one in which the criterion is a mean square error rather than probability of error. Analog messages (processes) can be transmitted in continuous time (e.g., angle modulation), sampled in time with continuous amplitudes (e.g., pulse amplitude modulation), or sampled, quantized and relayed over a digital channel.

One approach to the analog estimation problem would be a structured approach similar to the solution technique of Chapter 2. By choosing a communications system structure which employs the feedback channel and then optimizing over any free functions available, a feedback system results. Unfortunately such system, while they do give improved performance over nofeedback systems, give little insight into the effect of feedback on the system or its performance. Rather than approach analog estimation in this manner, a feedback scheme which is independent of the communication problem will be presented and then applied to several systems. Basically given a complete nofeedback system (which could be digital or analog with any appropriate transmitter/receiver), a procedure for adding a feedback channel to improve the performance without significantly changing the system modulation/demodulation is presented.

In Chapter 2 the optimal transmitted signal was found to be a multiple of the instantaneous error, the difference between the receiver state x(t) and the desired receiver state θ. Perhaps all communications systems using noiseless feedback should transmit some type of "error" waveform. For many reasons this is a logical choice for a transmitted signal using feedback. In transmitting the error the transmitter does not re-transmit what the receiver has already determined; thus, more power is available for transmitting what the receiver needs to know, namely, the error the receiver is making. For example, in a block coded system the feedback channel could inform the transmitter of the current status of the decoded message. The transmitter would then delete the remainder of the block bits if the receiver had already decided on the correct message and proceed to wait until it is time for the next message. If the receiver has not decoded the message correctly, the transmitter would continue to transmit bits which "drive" the decoder towards the correct message. Considerable savings in power are possible by not having to transmit the "correction" bits whenever the message bits are decoded properly. As will be shown, there are many cases in which the transmission of the error waveform is in fact optimal. Much of the next two sections is treated in Cruise [27].

3.1 Application of Feedback to Arbitrary Nofeedback Systems

Consider the following nofeedback system to which feedback will be applied to improve the performance. The transmitted signal is m(t) and is somehow related to the information being transmitted. For example if a process a(t) is being transmitted, m(t) could be of the form

$$m(t) = m(t,a(t)) \tag{3.1}$$

The exact dependence of the transmitted signal on the message is unimportant at this point. The channel is assumed to be a white noise channel. The received signal r(t) is

$$r(t) = m(t) + n(t) \tag{3.2}$$

Again, exactly what the receiver does to decode the message is not important.

For these signals a feedback channel has been added as shown in Figure 3-1a. Note that the transmitter which generates m(t) and the receiver which processes r(t) are not shown in the figure; only the feedback elements are shown. The box H_1 represents a time-varying realizable (possibly nonlinear) filter to be determined shortly. K(t) is a gain also to be determined. For the feedback system the transmitted signal is now

$$m'(t) = K(t)\,[m(t) - \hat{m}(t)] \tag{3.3}$$

In order to preserve the characteristics of the transmitted signal of the nofeedback system, let K(t) be such as to maintain exactly the same average instantaneous power in the feedback system as in the nofeedback system. This choice of K(t) will not guarantee that all the characteristics of the two transmitted signals will be the same, but at least the transmitter power/energy constraints will be identical. The appropriate choice of K(t) is

$$E[m^2(t)] = K^2(t)\,E[(m(t) - \hat{m}(t))^2] \tag{3.4}$$

Equation 3.4 determines K(t) in terms of the statistics of m(t) (the nofeedback transmitted signal) and $\hat{m}(t)$ (the signal fedback).

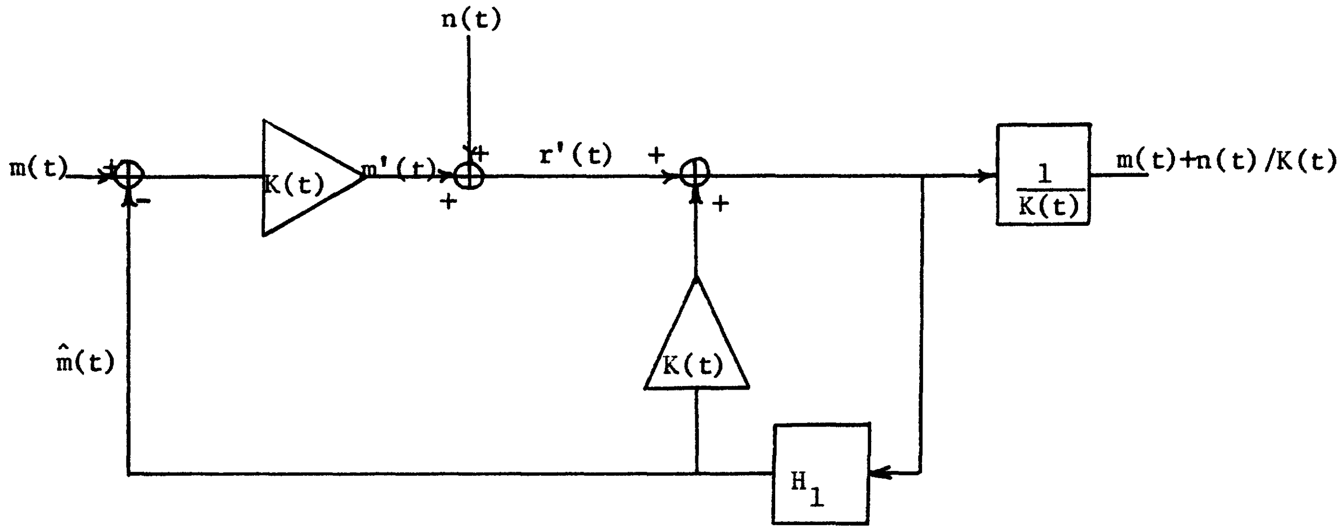


Figure 3-1a. General feedback structure

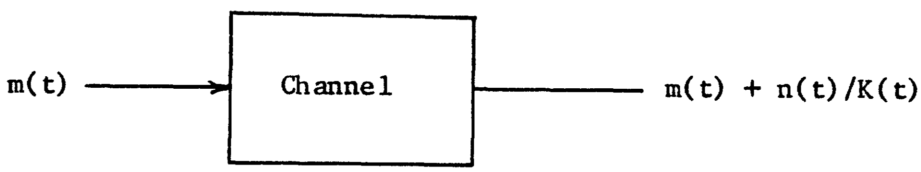


Figure 3-1b. Equivalent channel for feedback system

$\hat{m}(t)$ is determined by the choice of the filter H_1. As implied by the notation, H_1 is chosen to make $\hat{m}(t)$ an estimate of m(t), namely the minimum mean square error estimate of m(t) given the received signal up to time t. This estimate is also the same as the conditional mean of m(t) given $r(\tau)$ for $\tau < t$. Note that the estimate is of the transmitted signal, not of the message itself. If m(t) is as given in Equation 3.1, then

$$\hat{m}(t) = \hat{m}(t,a(t)) \tag{3.5}$$

not $m(t,\hat{a}(t))$ which would be another type of feedback system. For the most part in the examples these two types of estimates are equal due to the linearity of the modulation.

In Figure 3-1b the system of Figure 3-1a has been redrawn lumping all of the feedback parts into a new channel which appears very much like the original additive noise channel. Whichever transmitter/receiver structure is present for the nofeedback system can be used directly in the feedback system of Figure 3-1b with a slight modification for the time-varying white noise. The receiver observes the transmitted signal m(t) in a noise n(t)/K(t) rather than just n(t) in the nofeedback case. Since $K(t) \geq 1$, this is always an improvement since the noise is reduced in amplitude.

In conclusion a noiseless feedback system has been designed which effectively reduces the channel white noise density, regardless of the particular communications application of the channel. The ordinary nofeedback structure can be added at both ends of the feedback channel in Figure 3-1b. The feedback structure, however, does depend on the type of modulation being used. Observe that the feedback structure does not really leave an additive noise channel with smaller noise than the original additive noise; this would mean an increased channel capacity and a violation of Shannon's [5] result. A nofeedback system which operates at channel capacity will not operate above capacity with feedback; it will, however, perform better for finite T and approach $P_e = 0$ faster.

3.2 Parameter Estimation

In this section a feedback system similar to the system in Chapter 2 will be developed based on the ideas presented in Section 3.1. The approach is quite different here although the feedback system here is almost identical to that discussed in Chapter 2.

For the nofeedback system assume that the value of a Gaussian random variable θ (zero mean, variance σ^2) is to be conveyed across an additive white noise channel with spectral density $N_0/2$. θ might represent a voltage to be transmitted or perhaps part of a more complicated message. In Chapter 2 θ was related to the message points and did not have a Gaussian density.

Assume that the transmitter uses pulse amplitude modulation with the height of the pulse being proportional to θ. Then the transmitted signal m(t) is

$$m(t) = \sqrt{\frac{E_0}{\sigma^2 T}}\,\theta \tag{3.6}$$

in order to maintain an average transmitted energy of E_0 in the transmission interval [0,T]. The receiver is assumed to be the minimum mean square error estimator of θ; therefore, the output of the receiver at time T (when the estimate of θ is generated) is

$$\hat{\theta}(T) = \int_0^T dt \frac{\sqrt{\frac{\sigma^2 E_0}{T}}\; r(t)}{N_0/2 + E_0} \tag{3.7}$$

The normalized (or fractional) variance of the estimate $\hat{\theta}(T)$ for this nofeedback system is

$$\xi_{\substack{\text{no}\\ \text{feedback}}} = \frac{1}{1 + (2E_0/N_0)} = \frac{1}{\sigma^2} E[(\hat{\theta}(T) - \theta)^2] \tag{3.8}$$

which is identical to Equation 2.78, the nofeedback system of Chapter 2. This performance is independent of the density of θ and is optimal for the constraints of a simple linear receiver and a Gaussian density for θ. Certainly a nonlinear scheme could be devised which would convey the value of θ through the channel with a smaller normalized mean square error.

For this simple communications system feedback will be added as outlined in Section 3.1. One component of the feedback structure in Figure 3-1a which must be determined is the filter H_1. The input to H_1 is $K(\tau)m(\tau) + n(\tau)$ and the output is the minimum mean square error estimate of $m(\tau)$. For this problem with linear modulation and Gaussian statistics the best filter is also linear. The realization of this filter is just a Kalman [30] filter which, if the input is

$$r'(t) = K(t)m(t) + n(t)$$

$$= K(t)\sqrt{\frac{E_0}{\sigma^2 T}}\,\theta + n(t) \tag{3.9}$$

the minimum mean square error estimate of m(t) is

$$\hat{m}(t) = \sqrt{\frac{E_0}{\sigma^2 T}}\;\hat{\theta}(t) \tag{3.10}$$

where $\hat{\theta}(t)$ is the output of the Kalman filter

$$\frac{d}{dt}\,\hat{\theta}(t) = -K^2(t)P(t)\,\frac{2E_0}{\sigma^2 N_0 T}\,\hat{\theta}(t) + \frac{2}{N_0}\sqrt{\frac{E_0}{\sigma^2 T}}\,K(t)P(t)r'(t)$$

$$\hat{\theta}(0) = 0 \tag{3.11}$$

P(t) is the covariance of the estimate $\hat{\theta}(t)$ and is defined

$$P(t) = E[(\hat{\theta}(t) - \theta)^2] \tag{3.12}$$

P(t) is the solution of the Ricatti equation associated with the Kalman filter, namely

$$\frac{d}{dt} P(t) = \frac{2E_0}{\sigma^2 N_0 T} K^2(t) P^2(t) \qquad (3.13)$$

$$P(0) = \sigma^2$$

Applying Equation 3.4 the gain K(t) is evaluated from

$$K^2(t) E[(m(t) - \hat{m}(t))^2] = \frac{E_0}{\sigma^2 T} K^2(t)P(t) = \frac{E_0}{T} \qquad (3.14)$$

For this choice of K(t) the covariance P(t) satisfies

$$\frac{d}{dt} P(t) = -\frac{2E_0}{N_0 T} P(t) \qquad (3.15)$$

which is easily solved using the initial condition in Equation 3.11. The formulation of the Kalman filter in terms of $\hat{\theta}(t)$ (rather than $\hat{m}(t)$) is a convenience because the mean square error of $\hat{\theta}(T)$ is just P(T). The normalized mean square error of the feedback system is therefore

$$\xi_{feedback} = \exp[-2E_0/N_0] = \frac{P(T)}{\sigma^2} \qquad (3.16)$$

Observe that Equation 3.16 is identical to Equation 2.52 (if the notational differences are accounted for). The system derived in this section is the constant power system discussed in Section 2.8 and drawn in Figure 2-2. This system has a constant average power because the nofeedback system from which the feedback system was designed has a constant average power.

Another way to consider the addition of feedback is to consider the nofeedback receiver (Equation 3.7) redrawn as a Kalman filter in Figure 3-2a. Instead of just estimating θ at the end of the interval, a continuous estimate $\hat{\theta}(t)$ is generated. At $t = T$ both systems produce the same estimate and have the same mean square error. The advantage of this new realization of the nofeedback system is that the system with feedback looks very similar as shown in Figure 3-2b. Basically the feedback path in the nofeedback Kalman filter becomes the feedback signal in the noiseless feedback system. $K(t)$ adjusts the transmitter power, and the covariance $P(t)$ of the feedback system is now different from that of the no feedback system.

This derivation of the same system of Chapter 2 gives considerably more insight into exactly which part of the feedback link plays in the improved performance of feedback systems over nofeedback systems. Feedback allows the transmitter power to be reduced without changing the basic structure of the system or its performance. By inserting a gain ($K(t)$) to raise the transmitter power back to its allowed level, the overall performance is substantially improved.

3.3 Rate-Distortion Bound -- Parameter

Previously the linear feedback system was shown to be capable

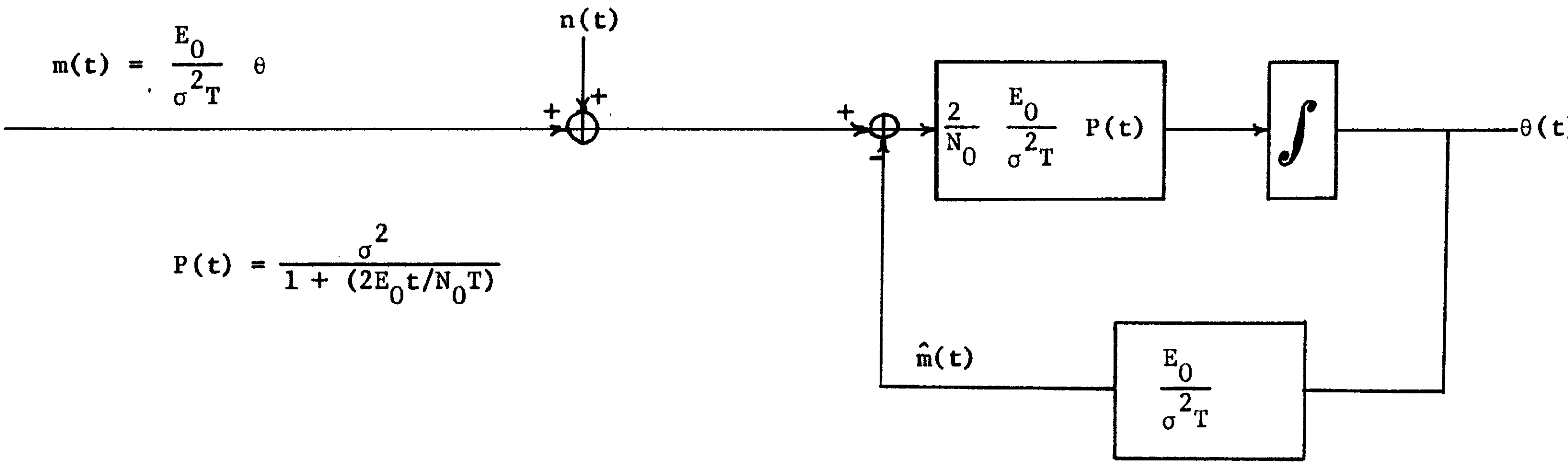


Figure 3-2a. Nofeedback system

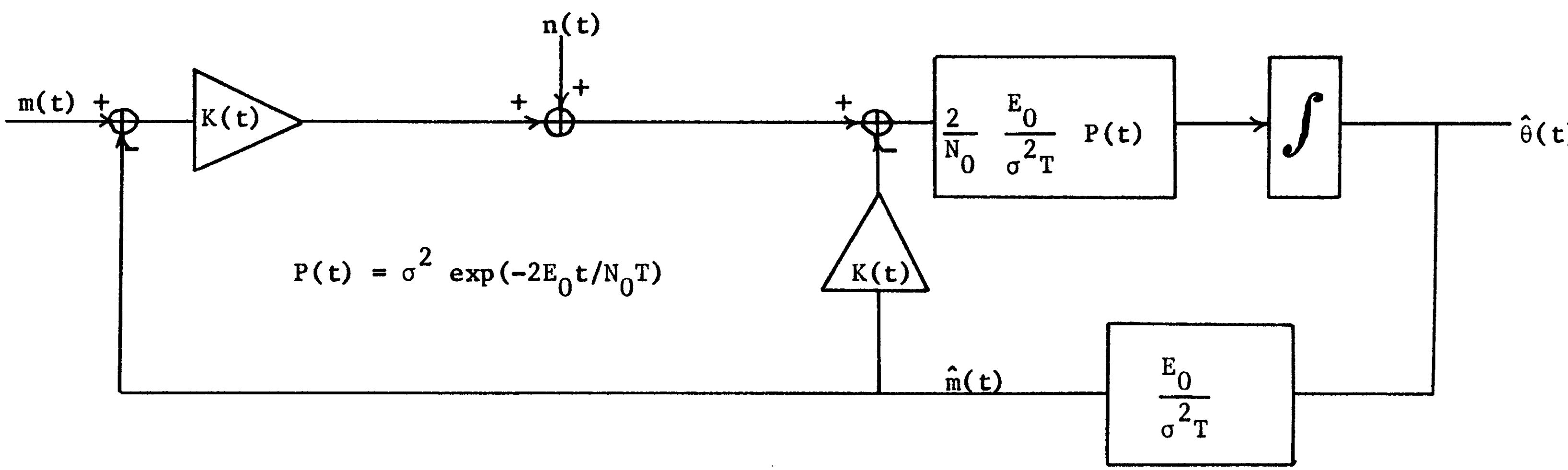


Figure 3-2b. Feedback system

transmitting messages at rates up to channel capacity with arbitrarily low probability of error. Similarly the linear feedback system can be shown to achieve the rate-distortion bound on mean square error. For digital systems channel capacity indicates the ultimate performance achieveable; for analog systems the rate-distortion bound indicates the minimum mean square error attainable.

Shannon [5] has derived an expression for the minimum channel capacity required to transmit a Gaussian variable θ (zero mean, variance σ^2) with a mean square error (distortion) of ε ($\leq\sigma^2$). This rate is given by

$$R(\varepsilon) = \frac{1}{2}\ln\left(\frac{\sigma^2}{\varepsilon}\right) \text{ nats} \tag{3.17}$$

$R(\varepsilon)$ is the minimum amount of information require to estimate θ with a mean square error no greater than ε.

The additive white Gaussian noise channel of the communications system has a capacity given in Equation 1.4. If this channel is used for T seconds (with or without feedback), the maximum information transmitted from the transmitter to the receiver is

$$C(T) = \frac{E_0}{N_0} \text{ nats} \tag{3.18}$$

if the channel is operated at channel capacity. By equating this maximum information with the amount of information required for a given error (Equation 3.17)

$$C(T) = R(\varepsilon) = \frac{E_0}{N_0} = \frac{1}{2}\ln(\frac{\sigma^2}{\varepsilon}) \tag{3.19}$$

the minimum mean square error attainable is

$$\frac{\varepsilon}{\sigma^2} = \exp[-2E_0/N_0] \tag{3.20}$$

Equation 3.20 is the rate-distortion bound on transmitting a Gaussian random variable over an additive white noise channel in T seconds; no modulation/demodulation system can achieve a normalized error less than that of Equation 3.20.

The linear nofeedback system analyzed in Section 3.2 has a normalized error given by Equation 3.8 which is substantially above the rate-distortion bound. The addition of feedback altered the system performance so that the normalized mean square error became that of Equation 3.16 which is precisely the rate-distortion bound in Equation 3.20, indicating that the linear feedback system is optimal. No other modulation scheme could possibly do better.

3.4 Rate-Distortion Bound -- Process, Finite Interval

Having shown that the linear feedback system is the best system in the sense that no other system can have a lower mean square error in transmitting a single random variable, the feedback system can be modified to transmit optimally a finite set of random variables or a T second segment of a random process. The solution to the former problem will be apparent from the solution of the latter. In order to send a Gaussian random process with a minimum mean square error given by the rate-distortion bound, the process is decomposed into its Karhunen-Loeve coordinates and each coordinate is transmitted exactly like θ in Section 3.2.

A Gaussian random process (zero mean) a(t) can be expanded

$$a(t) = \sum_{i=1}^{\infty} a_i \phi_i(t) \qquad 0 \le t \le T \tag{3.21}$$

where

$$a_i = \int_0^T dt\, a(t)\phi_i(t) \tag{3.22}$$

$$\int_0^T du\, \phi_i(u)\phi_j(u) = \delta_{ij} \tag{3.23}$$

The infinite set of a_i's are independent zero mean Gaussian variables with variance λ_i where

$$\lambda_i\, \phi_i(t) = \int_0^T du\, R_a(t,u)\phi_i(u) \tag{3.24}$$

$$E[a(t)a(u)] = R_a(t,u) \tag{3.25}$$

These equations are just the statement of the Karhunen-Loeve expansion (see Davenport [28]).

In order to transmit the T second segment of a(t), the coefficients a_i are first evaluated at the transmitter. This operation requires an initial delay of T seconds so that the a_i in Equation 3.22 can be calculated. Then these a_i are transmitted one at a time over the noiseless feedback system using a subinterval of length τ_i and a reduced energy E_i for the i-th coefficient. The exact division of time and energy to optimize the system must be determined.

The criterion of the system is the integrated mean square error

$$\varepsilon = \int_0^T dt\, E[(\hat{a}(t) - a(t))^2] \tag{3.26}$$

which is to be minimized. If ε_i is the estimation error in estimating each of the a_i at the receiver, then the total error is

$$\varepsilon = \sum_{i=1}^{\infty} \varepsilon_i \tag{3.27}$$

Using the feedback system of Section 3.2, Equation 3.16 implies that ε_i is $\lambda_i \exp[-2E_i/N_0]$ and therefore

$$\varepsilon = \sum_{i=1}^{\infty} \lambda_i \exp[-2E_i/N_0] \tag{3.28}$$

is the expression to be minimized by selecting E_i. The E_i are not completely arbitrary since there sum is the average energy transmitted by the transmitter of the feedback system. Therefore, the E_i must satisfy the energy constraint

$$E_0 = \sum_{i=1}^{\infty} E_i \tag{3.29}$$

The minimization of Equation 3.28 subject to the constraint in Equation 3.29 can be handled easily with ordinary calculus. The resulting necessary condition for a minimum is

$$\lambda_i \exp[-2E_i/N_0] = \text{constant} = \beta \qquad \text{for } E_i \neq 0 \tag{3.30}$$

The solution for the optimal energy distribution is such that only a finite number of the a_i's are transmitted. No energy is used to convey the lower energy eigenvalues of the process.

Assume that K of the a_i are transmitted with nonzero energy; Equation 3.31 implies that these K variables correspond to the K largest eigenvalues λ_i. The mean square error is

$$\varepsilon = \beta K + \sum_{i=K+1}^{\infty} \lambda_i \tag{3.31}$$

where β is such that

$$E_0 = \frac{N_0}{2} \sum_{i=1}^{K} \ln\left(\frac{\lambda_i}{\beta}\right) \tag{3.32}$$

Equations 3.31-2 are precisely the rate-distortion equations as derived in Goblick [16] for a random process. Note that the individual energies are selected so that each coordinate is estimated with exactly the same (β) mean square error; the error in the remaining coordinates is just the variance.

The selection of the lengths (τ_i) of the subintervals determines the power distribution in the interval. In order to achieve a constant average instantaneous power, the intervals should be chosen proportional to the energies E_i.

As noted earlier, a delay of T seconds is necessary at the transmitter in order to calculate the K a_i to be sent. Similarly another T seconds of delay is required at the receiver to reconstruct the estimate a(t).

In conclusion a technique for transmitting an arbitrary Gaussian random process a(t) over a white noise channel in the time interval [0,T] with the minimum mean square error attainable has been demonstrated by utilizing a noiseless feedback channel. In the next section the identical problem is treated for the case when the time interval is infinite.

3.5 Rate Distortion Bound -- Stationary Process

For a stationary process a(t) ($-\infty<t<\infty$) several concepts from the finite time interval case of Section 3.4 need to be altered to achieve the appropriate rate-distortion bound. Basically in the infinite interval an integrated error is meaningless; the new criterion is mean square error. Unless the process is stationary, the criterion is a function of time; hence, a(t) is assumed to be a sample function from a stationary random process. During an infinite interval the transmitted energy can be infinite and the information conveyed can be infinite; hence, a transmitter power constraint is appropriate rather than energy, and a channel capacity in nats/second replaces the previous capacity in nats. The eigenfunctions of the finite interval sample function tend to sinusoids as $T \to \infty$. The eigenvalues become a continuum, namely the spectral density of the process. Given these changes, the transmission scheme is very similar to finite time interval case of the preceding section. The lower amplitude eigenvalues are neglected and the higher amplitude eigenvalues are scaled in power before transmission. These operations are easily done with linear filters designed in the frequency (eigenvalue) domain.

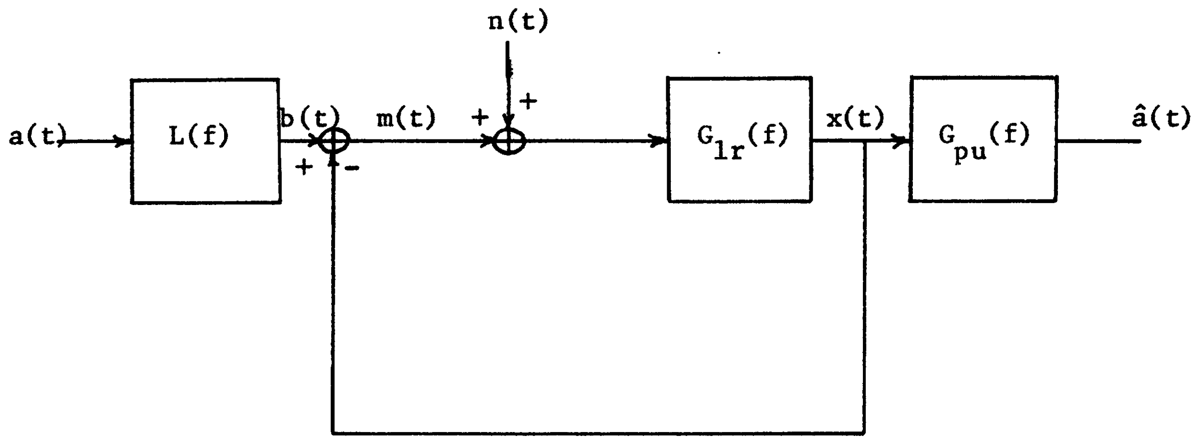
Cruise [15] describes a noiseless feedback system which achieves the rate-distortion bound on the transmission of analog signals over additive white Gaussian noise channels. The performance of the feedback system is derived by comparison of the noiseless feedback system with a phase locked loop model. In this thesis these same results will be derived without making reference to phase locked loops. A linear filter structure for the feedback system is assumed and the filters chosen optimally; the resulting performance is that given by the rate-distortion bound.

Figure 3-3 shows a linear noiseless feedback system which for suitable choices of the filters L(f), $G_{1r}(f)$, and $G_{pu}(f)$ will achieve the rate-distortion bound on performance. The preemphasis filter L(f) and the postloop filter $G_{pu}(f)$ are allowed to be unrealizable filters because they can be realized with some delay. The loop filter $G_{1r}(f)$, however, must be realizable if the feedback channel is to be realistic. The overall system is realizable-with-delay, just as a coded digital system is realizable with coding and decoding delays.

Consider the design of $G_{1r}(f)$, the loop realizable filter. Choose this filter such that x(t) (see Figure 3-3) is the minimum variance (realizable) estimate of b(t), the output of the preemphasis filter. The solution for $G_{1r}(f)$ involves the solution for a realizable Wiener-Hopf filter. Without actually solving for $G_{1r}(f)$ the performance or variance of the estimate x(t) of b(t) is given by

$$\xi_b = E[(x(t)-b(t))^2] = \frac{N_0}{2}\int_{-\infty}^{\infty} df\ \ln[1 + \frac{2}{N_0} S_b(f)] \quad (3.33)$$

$$= \frac{N_0}{2}\int_{-\infty}^{\infty} df\ \ln[1 + \frac{2}{N_0} |L(f)|^2 S_a(f)]$$



a(t) = message process with spectral density $S_a(f)$

m(t) = transmitted signal

x(t) = feedback signal

$G_{lr}(f)$ = realizable linear filter

$G_{pu}(f)$ = unrealizable linear filter

L(f) = unrealizable linear filter

â(t) = realizable-with-delay estimate of a(t)

Figure 3-3. Noiseless feedback system achieving the rate-distortion bound on performance

Observe that the error ξ_b is also the transmitted power of the feedback system in Figure 3-3 upon which there is a constraint. Also, the calculation of $G_{1r}(f)$ has been bypassed; it is not needed and will not be calculated here.

Since the postloop filter is allowed to be unrealizable and arbitrary, any effects of the loop filter and feedback link can be removed by inverse filtering, leaving an unrestricted unrealizable filtering problem. The overall system appears as shown in Figure 3-4 where the realizable loop has been redrawn as the realizable filter $H_r(f)$ with spectrum given by

$$H_r(f) = \frac{G_{1r}(f)}{1 + G_{1r}(f)} \tag{3.34}$$

In order to cause the output of the system a(t) to be a minimum variance estimate of a(t), $G_{pu}(f)$ is chosen to be that function which makes $H_r(f)G_{pu}(f)$ the unrealizable Wiener filter for the problem shown in Figure 3-4. The variance of this estimate can be expressed (again without actually computing the optimal filter) as

$$\varepsilon = E[(\hat{a}(t) - a(t))^2] = \int_{-\infty}^{\infty} df \, \frac{S_a(f)}{1 + \frac{2}{N_0}|L(f)|^2 S_a(f)} \tag{3.35}$$

The transmitted power (average) $E[m^2(t)]$ is given by ξ_b in Equation 3.33 as a function of the message spectrum $S_a(f)$ and the power spectrum $|L(f)|^2 \equiv S_L(f)$. The performance (mean square error)

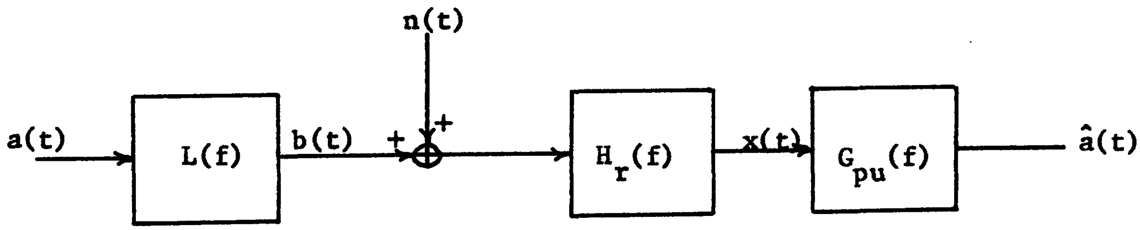


Figure 3-4. Simplification of feedback system in Figure 3-3

of the feedback system is given in Equation 3.35 in terms of the same quantities. The two filters $G_{lr}(f)$ and $G_{pu}(f)$ have been determined in deriving these two equations, leaving $S_L(f)$ as the remaining function to be selected in the system. $S_L(f)$ is adjusted so that

$$\xi_b = P_{ave} = \text{transmitted power constraint} \qquad (3.36)$$

and the performance ε is minimum. $S_L(f)$ is an unrestricted function except that it is a power spectrum and must be nonnegative. The minimization can be carried out by forming a functional J which is

$$J = \varepsilon + \frac{2\beta}{N_0}\xi_b$$

$$= \int_{-\infty}^{\infty} df \, [\frac{S_a(f)}{1 + (2/N_0)S_a(f)S_L(f)}$$

$$+ \beta \ln(1 + \frac{2}{N_0} S_a(f)S_L(f))] \qquad (3.37)$$

where the Lagrange multiplier is $2\beta/N_0$; β is arbitrary but constant. Perturbation of the above equation yields

$$\delta J = 0 = [-S_L(f) + \frac{N_0}{2\beta} - \frac{N_0}{2S_a(f)}] \, \delta S_L(f) \qquad (3.38)$$

or solving for $S_L(f)$

$$S_L(f) = \text{larger of} \begin{cases} \frac{N_0}{2}\left(\frac{1}{\beta} - \frac{1}{S_a(f)}\right) \\ 0 \end{cases} \tag{3.39}$$

The constant β is determined by setting the power constraint in Equation 3.36 to equality. Denote the frequencies where $S_L(f)>0$ by F and the frequencies where $S_L(f) = 0$ by $\bar{F}$. Then the power constraint in Equation 3.36 becomes

$$\xi_p = P_{ave} = \frac{N_0}{2} \int_F df \ \ln[S_a(f)/\beta] \tag{3.40}$$

and the performance or mean square error of the estimate $\hat{a}(t)$ is

$$\varepsilon = \int_{\bar{F}} df \ S_a(f) + \beta \int_F df \tag{3.41}$$

These two equations are the continuous analogs to the discrete spectrum (finite time interval) problem of Section 3.4. Note the similarity between Equations 3.40-1 and Equations 3.31-2.

A comparison of Equations 3.40-1 with the rate-distortion bound (see Goblick [16]) equations shows that they are identical. Thus, the feedback system of Figure 3-3 can achieve the ultimate performance of all systems (feedback or nofeedback) operating over the white noise channel. No other system can have a lower mean square error than that given by the simultaneous solution of Equations 3.40 and 3.41.

The solution for $S_L(f)$ in Equation 3.39 indicates that L(f) is a strictly bandlimited (perhaps several passbands) unrealizable filter which passes only the frequency regions where $S_a(f)$ is large. No attempt is made to estimate the process a(t) in the frequency range(s) $\bar{F}$. Those frequencies in F which are not attenuated to 0 by L(f) are scaled so that the resulting error spectrum is constant. This form of the error spectrum is indicative of achieving the rate-distortion bound.

The processing of the sample function in this infinite interval is the same type of processing found in the previous section for finite time intervals where only the highest eigenvalue coefficients were transmitted. Observe that a delay is required for L(f) just as a delay was required at the receiver in both systems.

Some idea of how much improvement in performance (mean square error) is afforded by the feedback channel can be gained by considering the one pole message spectrum

$$S_a(f) = \frac{2k\, P_{ave}}{4\pi^2 f^2 + k^2} \qquad (3.42)$$

Define a signal-to-noise ratio for the problem as

$$\lambda = \frac{4P_{ave}}{N_0 k} \tag{3.43}$$

For large values of λ (roughly $\lambda>20$) the noiseless feedback system (also rate-distortion bound) has a normalized mean square error of

$$\frac{\varepsilon}{P_{ave}} \simeq \frac{8}{\pi^2} \frac{1}{1 + \lambda/2} \qquad \text{(rate-distortion)} \tag{3.44}$$

Suppose that the preemphasis filter L(f) is replaced by a constant gain. It will not reject the proper frequencies nor scale properly the rest of the frequencies; therefore, such a system is a suboptimal feedback system. The normalized mean square error for this system is

$$\frac{\varepsilon}{P_{ave}} = \frac{1}{1 + \lambda/2} \qquad (L(f) = \text{constant}) \tag{3.45}$$

This performance is not significantly different from the ultimate given in Equation 3.44. The conclusion is that the preemphasis filter is not extremely critical to the performance of the feedback system.

Suppose the suboptimal system is further degraded by requiring the postloop filter $G_{pu}(f)$ to be constant. Now the system is realizable-without-delay since both unrealizable filters have been removed. The performance drops to

$$\frac{\varepsilon}{P_{ave}} = \frac{1}{1 + \lambda/4} \qquad \text{(realizable-without-delay)} \qquad (3.46)$$

Again the system performance is reduced, but not significantly. The possible advantage of this suboptimal feedback system is that all the filters are realizable.

If the feedback channel is removed, the system is an ordinary linear filtering problem. Allowing delay, the mean square error for this unrealizable Wiener filtering problem is

$$\frac{\varepsilon}{P_{ave}} = \frac{1}{\sqrt{1 + \lambda}} \qquad \text{(realizable-with-delay without feedback)} \qquad (3.47)$$

This is the performance of the nofeedback system to which feedback has been added. With feedback the performance increases from that of Equation 3.47 to that of any of the preceding three equations depending on the choice of system filters. Equation 3.44 is the ultimate performance achieveable with any system (not just linear) with or without feedback.

A comparison of the forms of the feedback system (Equation 3.47) and the feedback system indicates that feedback essentially squares the normalized mean square error. If the fractional error would be .01 for the linear system without feedback, it would be about .0001 if feedback were added. Compared with the exponential improvement that feedback offers when transmitting a single random variable, this improvement of the feedback system is not as dramatic. Yet the rate-distortion bound states this is in fact the maximum improvement possible.

The various feedback systems (optimum and suboptimum) mentioned above indicate that the feedback system is relatively insensitive to the filters in it. A realizable-without-delay feedback system (which is suboptimum) is only 3 db worse than the optimum for large signal-to-noise ratios. The inverse dependence of the normalized error for feedback systems is not lost by restricting the feedback system filters to be realizable; hence, even such suboptimum systems offer almost as much performance gain over the no-feedback system as does the optimum.

Another useful way of comparing nofeedback and feedback systems is in terms of the effective increase in signal-to-noise ratio which the addition of feedback implies. For example, suppose the signal-to-noise ratio (suitably defined for some system) is 10. Without feedback this 10 implies some performance of the system; with feedback this 10 implies a much improved performance. Taking this improved performance with feedback, it implies that a nofeedback system would require a much higher signal-to-noise ratio to achieve this same performance, say 1000. Thus, feedback gives an effective signal-to-noise ratio of 1000 for an actual signal-to-noise ratio in the channel of 10. In other words a nofeedback system with signal-to-noise ratio of 1000 performs as well as a feedback system operating through a channel with signal-to-noise ratio of 10.

For the digital system of Chapter 2 and the parameter system of Section 3.2 the appropriate signal-to-noise ratio is

$$SNR = 2E_0/N_0 \tag{3.48}$$

The use of feedback implies that the effective signal-to-nosie ratio is

$$SNR \xrightarrow{\text{digital}} \exp[SNR] - 1 \tag{3.49}$$

for the digital system. For the analog systems in this section the appropriate signal-to-noise ratio is λ given in Equation 3.43. For this definition the addition of feedback implies an effective signal-to-noise ratio

$$SNR \xrightarrow{\text{analog}} SNR(1 + SNR/4) \tag{3.50}$$

Comparing Equations 3.49 and 3.50 for the improvement which feedback offers for these two types of systems, the improvement of analog process systems is much less signal-to-noise ratio-wise than that of the digital (or random variable) system. For the digital system as time progresses in the interval the receiver is able to improve its estimate of θ continually (and hence reduce the effective transmitted signal power via the feedback link). The analog process system has a message which changes continuously with time so that the error in the estimate of the message cannot be arbitrarily small with increasing time; hence, the feedback link is not able to reduce the effective transmitted power as much.

3.6 Kalman Filtering with Noiseless Feedback

In many applications of message estimation the Kalman [17] formulation is appropriate. These instances are when the message process is suitably represented by a finite dimensional vector random process. For such a process observed in white noise the minimum variance estimate of the process is the output of a Kalman filter defined by a linear differential equation. The Kalman filter is the realizable Wiener-Hopf filter for the problem. Since the filter is realizable, the results of Section 3.4 imply that the addition of a feedback link to the system will not allow the overall system to achieve the rate-distortion bound; delays at the transmitter and receiver are necessary to achieve the bound. Nevertheless, it does offer improvement even when delays are not allowed. The example of Section 3.4 indicates that omitting the delays is not critical to the improved performance of feedback systems.

The general vector formulation of Kalman filtering implies a vector or diversity channel, that is, the communications channel is actually several parallel white noise channels. The transmitter power constraint can take any number of forms in this situation depending on how the composite channel is constructed. For example, the total transmitted power (sum over all separate channels) might be limited or perhaps each separate channel has a maximum transmitted power. Many other possibilities exist for appropriate transmitter(s) power constraint.

Rather than choose a particular transmitter power constraint, for analysis purposes the addition of a feedback link will be made without actually calculating the improvement which feedback offers. This improvement depends exactly on the definition of the transmitter power constraint.

Following Kalman's [17] notation define the vector message process by the vector differential equation

$$\frac{d}{dt}\underline{x}(t) = \underline{\underline{F}}(t)\underline{x}(t) + \underline{\underline{G}}(t)\underline{u}(t) \tag{3.51}$$

where the white noise driving the equation satisfies

$$E[\underline{u}(t)\underline{u}'(s)] = \underline{\underline{Q}}(t)\,\delta(t-s) \tag{3.52}$$

The transmitted signal is

$$\underline{y}(t) = \underline{\underline{H}}(t)\,\underline{x}(t) \tag{3.53}$$

which need not be of the same dimension as the process $\underline{x}(t)$. The channel adds a vector white noise $\underline{w}(t)$ to $\underline{y}(t)$ to form the received signal

$$\underline{z}(t) = \underline{\underline{H}}(t)\, \underline{x}(t) + \underline{w}(t) \tag{3.54}$$

where the remaining noise correlations are

$$E[\underline{w}(t)\underline{w}'(s)] = \underline{\underline{R}}(t)\, \delta(t-s) \tag{3.55}$$

and

$$E[\underline{w}(t)\underline{u}'(s)] = \underline{\underline{0}}$$

$\underline{\underline{R}}^{-1}(t)$ is assumed to exist. The receiver structure is the linear Kalman filter specified by

$$\frac{d}{dt}\,\hat{\underline{x}}(t) = \underline{\underline{F}}(t)\hat{\underline{x}}(t) + \underline{\underline{P}}(t)\underline{\underline{H}}'(t)\underline{\underline{R}}^{-1}(t)[\underline{z}(t) - \underline{\underline{H}}(t)\hat{\underline{x}}(t)] \tag{3.57}$$

where

$$\underline{\underline{P}}(t) = E[(\underline{x}(t) - \hat{\underline{x}}(t))(\underline{x}(t) - \hat{\underline{x}}(t))'] \tag{3.58}$$

The differential equation for $\underline{\underline{P}}(t)$ is

$$\frac{d}{dt}\underline{\underline{P}}(t) = \underline{\underline{F}}(t)\underline{\underline{P}}(t) + \underline{\underline{P}}(t)\underline{\underline{F}}'(t) + \underline{\underline{G}}(t)\underline{\underline{Q}}(t)\underline{\underline{G}}'(t)$$

$$- \underline{\underline{P}}(t)\underline{\underline{H}}'(t)\underline{\underline{R}}^{-1}(t)\underline{\underline{H}}(t)\underline{\underline{P}}(t) \qquad (3.59)$$

All of the above differential equations have associated initial conditions which are not stated explicitly.

The transmitted power matrix for this nofeedback system is

$$\underline{\underline{S}}(t) = E[\underline{y}(t)\underline{y}'(t)] = \underline{\underline{H}}(t)\, E[\underline{x}(t)\underline{x}'(t)]\, \underline{\underline{H}}'(t)$$

$$= \underline{\underline{H}}(t)\, \underline{\underline{T}}(t)\, \underline{\underline{H}}'(t) \qquad (3.60)$$

The appropriate power constraint will depend on $\underline{\underline{S}}(t)$. For example, if the total power transmitted over the several diversity white noise channels is limited, then this constraint involves only $Tr[\underline{\underline{S}}(t)]$.

The nofeedback system above has a performance given by Equation 3.59 and a transmitted power given by Equation 3.60. Consider adding a noiseless feedback channel to this system. The development associated with the parameter system shown in Figure 3-2 indicates that the formulation of the nofeedback system in terms of a Kalman filter has a natural feedback structure by feeding back the Kalman filter output estimate of the transmitted signal. In the case here the estimate of the transmitted signal is $\underline{\underline{H}}(x)\ \underline{x}(t)$ which is returned to the transmitter via the feedback channel. The new transmitted signal (of the feedback system) is the difference

$$\underline{y}(t)\Big|_{\text{feedback}} = \underline{\underline{H}}(t)\ [\underline{x}(t) - \hat{\underline{x}}(t)] \tag{3.61}$$

The channel adds white noise to the transmitted signal to form the received signal

$$\underline{z}(t)\Big|_{\text{feedback}} = \underline{\underline{H}}(t)[\underline{x}(t) - \hat{\underline{x}}(t)] + \underline{w}(t) \tag{3.62}$$

The receiver structure is only slightly altered

$$\frac{d}{dt}\hat{\underline{x}}(t)\Big|_{\text{feedback}} = \underline{\underline{F}}(t)\hat{\underline{x}}(t)\Big|_{\text{feedback}} + \underline{\underline{P}}(t)\underline{\underline{H}}'(t)\underline{\underline{R}}^{-1}(t)\underline{z}(t) \tag{3.63}$$

where $\underline{\underline{P}}(t)$ is exactly the same covariance function specified in Equation 3.59. Therefore, the performance of the feedback system is exactly the same as that of the nofeedback, but the transmitted power is reduced to

$$\underline{\underline{S}}(t)\Big|_{\text{feedback}} = \underline{\underline{H}}(t)\underline{\underline{P}}(t)\underline{\underline{H}}'(t) \tag{3.64}$$

The exact improvement of the feedback system depends on the definition of the power constraint as well as the definition of the performance. Perhaps the receiver is only interested in estimating one of the states of $\underline{x}(t)$ or some combination of states. These characteristics all depend on the individual problem treated and can be determined from the above equations. One would expect that the improved performance is on the order of that calculated in Section 3.5, namely the normalized mean square error (suitably defined for the vector problem) of the feedback system is roughly the square of the normalized mean square error of the nofeedback system.

The assumption was made above that the feedback link diversity equalled the forward channel diversity, that is, if there are 3 forward white noise channels, then there are also 3 reverse noiseless channels. It is easily shown that more diversity in the feedback link than the forward is redundant and will not improve the performance. The actual feedback diversity needed can be less than the forward diversity. The required diversity in the feedback channel for the above analysis to be valid is given by

$$\text{feedback diversity} = \text{rank}[\underline{\underline{H}}(t)] \leq \text{forward diversity} \quad (3.65)$$

If the feedback diversity is less than that given in Equation 3.65, then the feedback system falls in the class of noisy feedback systems.

CHAPTER 4

Noisy Feedback Systems

Thus far only noiseless feedback systems have been investigated. The implied definition of a noiseless feedback system is a system in which the transmitter has knowledge of the exact state of the receiver at each instant in time during the transmission interval. A noisy feedback system, then, is one in which the transmitter does not have exact knowledge of the state of the receiver. For example, suppose the feedback channel has a separate additive white noise which is added to the signal transmitted from the receiver back to the transmitter. The transmitter would then observe the receiver state in white noise and would be unable to determine the exact receiver state. Another example of a noisy feedback system would be one in which there is a delay in the feedback path; the transmitter would observe $x(t-t_0)$, but not $x(t)$, the current state. For this system the feedback "noise" is the delay t_0 which prohibits exact knowledge of the current receiver state.

In this chapter these various types of noisy feedback systems are analyzed to evaluate the forward channel improvement using a noisy feedback link. Most physical systems have some type of noise in the feedback link; hence, this chapter is more important from a practical point of view than the preceding chapters. Unfortunately there are very few analytic results for noisy feedback systems.

4.1 Discrete-Time Solution of Noisy Feedback

Before proceeding to the continuous-time formulation of noisy feedback systems, some useful motivation and insight can be gained from the discrete-time systems of Butman [23] and Elias [3]. Both have formulated the general linear feedback system with an additive noise feedback channel, but were unable to solve for the optimal parameters of their respective systems.

In a discrete-time system the transmitter uses the channel N separate times to transmit the message point θ as indicated in Figure 4-1. The transmitted signal m_k depends linearly on θ and the output of the feedback channel. The feedback channel has a delay of one time unit so that it is only used N-1 times. The performance (mean square error in estimating θ) and the transmitter energy constraint are identical in form to those of the continuous-time system analyzed in Chapter 2. Since there is noise in the feedback channel, a feedback transmitter energy constraint is also necessary. The additive Gaussian noises in the forward and feedback channels, n_k and w_k, are assumed independent of each other with zero means and constant variances, $E[n_k^2] = N_0/2$ and $E[w_k^2] = W_0/2$.

The optimization problem is to find the four arbitrary linear filter shown in Figure 4-1 which minimize the performance subject to the two transmitter (forward and feedback) constraints. This is the statement of the general linear additive noise discrete

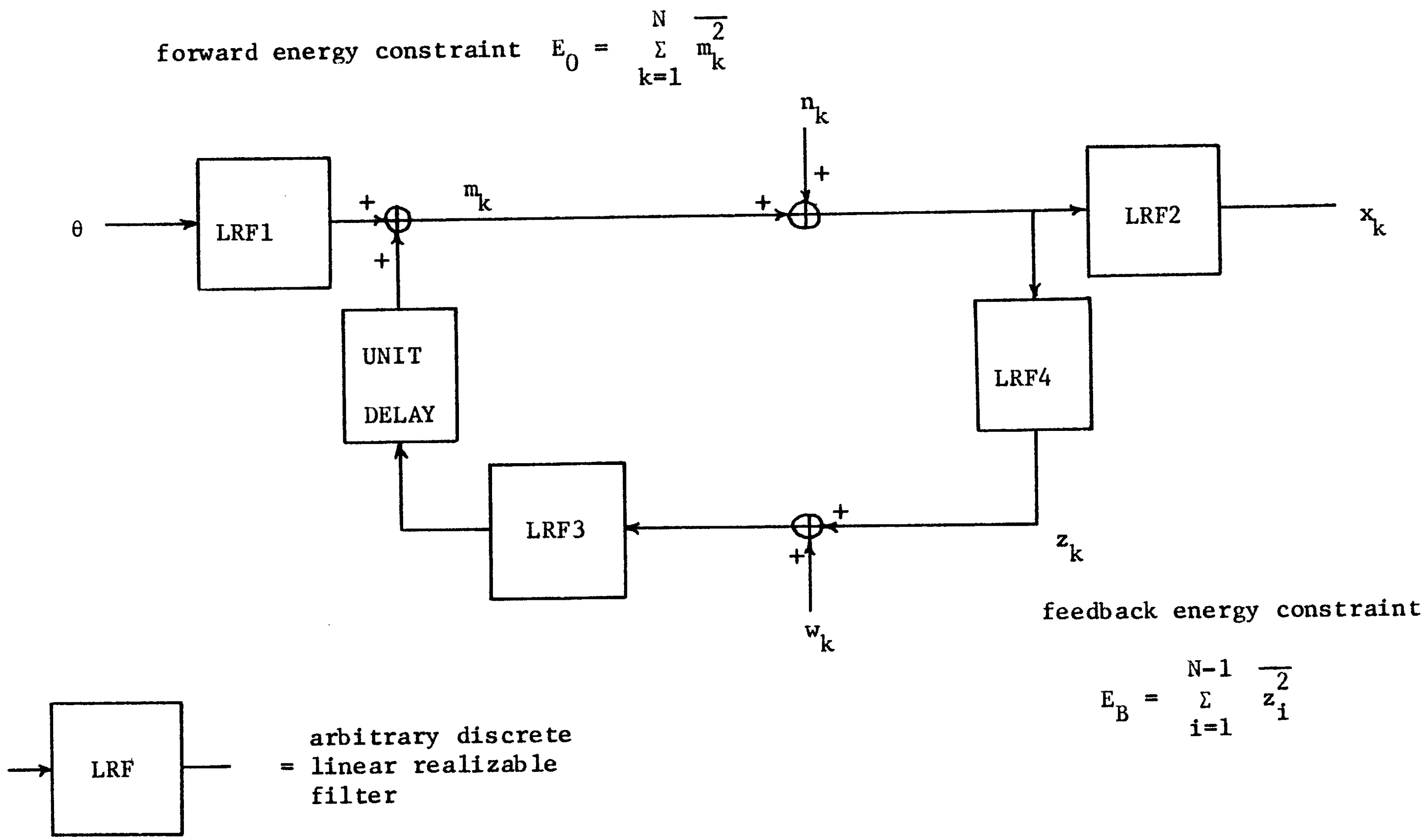


Figure 4-1. General linear discrete-time feedback system

feedback problem. The continuous-time version of this general linear problem is obvious. The solution to either version of the problem is highly desirable, but has not been found. The only optimal solution known is in the discrete-time version for $N = 2$ as solved by Elias [3].

Since Elias' result is the only true solution to any noisy feedback system, it is important to understand it for any assistance it might give in designing suboptimal systems (suboptimal since the optimal is not known) utilizing noisy feedback. Elias' result on feedback systems is a sidelight to his paper on networks of Gaussian channels; he gives very little detail other than the system performance. Here his system will be described in much more detail with emphasis on the structure rather than the performance.

If $N = 2$ for the system in Figure 4-1, then only 2 transmissions forward and 1 transmission backward is made. This special case of Figure 4-1 is redrawn in an expanded fashion in Figure 4-2. The time sequence of operation of the system in Figure 4-2 is the top horizontal channel operates, then the sloping feedback channel operates, then the bottom channel operates, and finally the estimate of θ is generated as a linear combination of the 2 receiver outputs. The constant gains $t_1,\ldots,t_6$ can be identified as the components of the arbitrary filters LRF1,...,LRF4 in Figure 4-1. The noise has been relabeled to correspond to Elias' [3] notation; actually n_1 and n_3 are the two forward channel noise samples and n_2

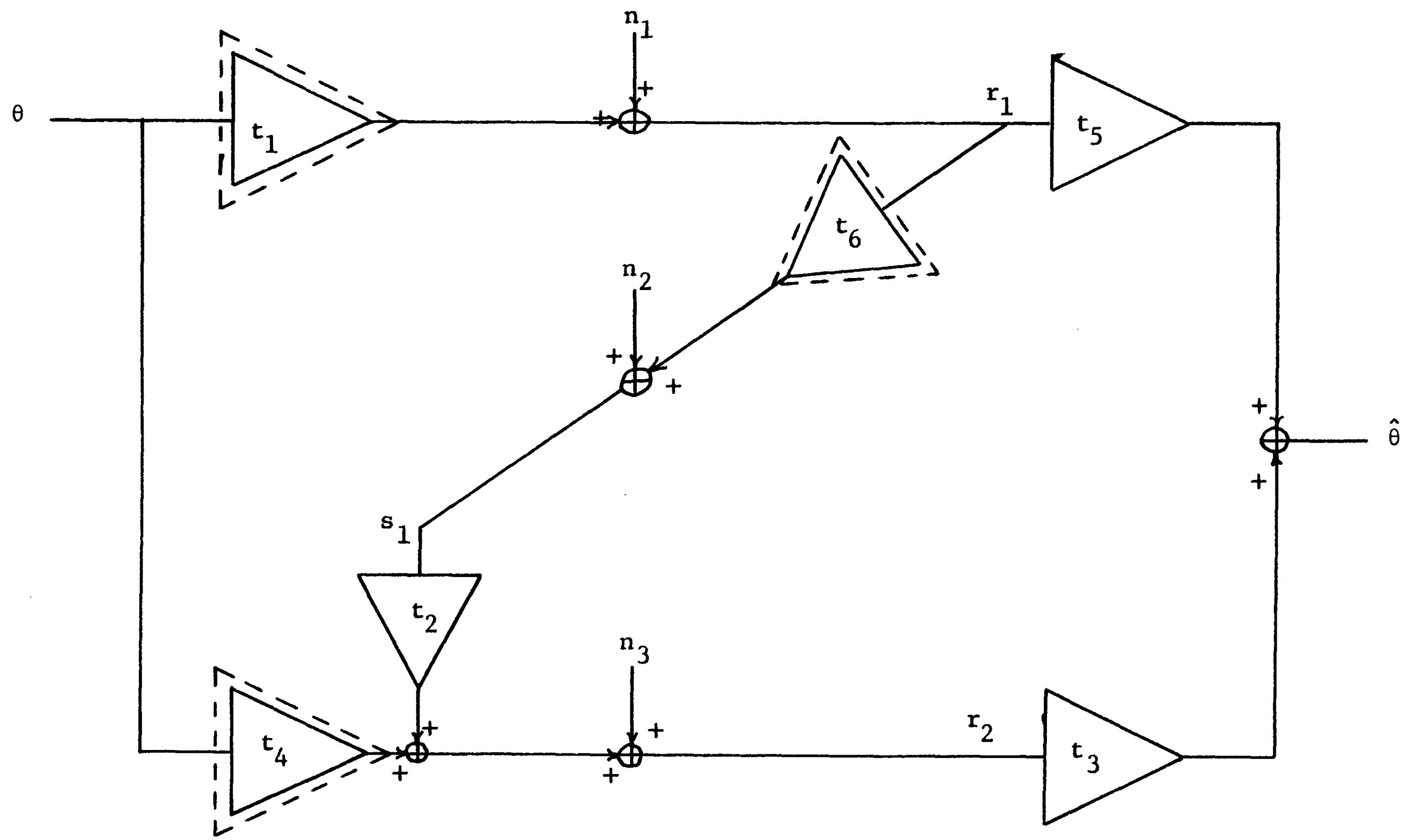


Figure 4-2. Discrete-time feedback system for N=2

is the feedback channel noise sample. The noise variances are defined by Elias so that the specified signal-to-noise ratio of each of the three noisy branches is maintained. Each branch has a signal-to-noise ratio of $1/N_i$ $(i=1,3)$; the forward and feedback energy constraints imply what these three numbers should be. With the noise variances defined so as to maintain the required signal-to-noise ratio, the gains t_1, t_4, and t_6 may be set to unity since they are redundant degrees of freedom in the analysis; these gains have dotted lines around them in Figure 4-2. The remaining 3 parameters t_2, t_3, and t_5 are the constants to be determined optimally.

Summarizing Elias' problem, given branch signal-to-noise ratios $1/N_i$ find the optimal choice of t_2, t_3, and t_5 such that the output signal-to-noise ratio

$$S_{out} = \frac{Var[\theta]}{Var[\hat{\theta}]} - 1 \tag{4.1}$$

is maximized. Clearly maximizing S_{out} is the same as minimizing the mean square error (variance) of the estimate $\hat{\theta}$. The constraint on the forward transmitted energy can be written in terms of the signal-to-noise ratios as

$$\frac{2E_0}{N_0} = SNR_F = \frac{1}{N_1} + \frac{1}{N_3} \tag{4.2}$$

The feedback transmitter energy constraint is

$$SNR_B = \frac{1}{N_2} \tag{4.3}$$

The transmitted variable θ is assumed zero-mean Gaussian; hence, the linear feedback system structure assumed turns out to be optimal of all possible feedback structures.

Given t_2 in Figure 4-2, the optimal choice of t_3 and t_5 is obvious since the structure is a discrete Gauss-in-Gauss problem (see Van Trees [30]). By carrying out the analysis outlined in Elias [3], the weight t_2 and the overall optimal output signal-to-noise ratio are

$$t_2 = -\frac{N_1}{(1+N_1)(N_1 + N_2(1+N_1))} \tag{4.4}$$

and

$$S_{\substack{out\\opt}} = \frac{2}{N_1} + \frac{1}{N_1^2 + N_2(1+N_1)^2} \tag{4.5}$$

if θ is taken as a unit variance random variable. Also the optimal choice of N_3 and N_1 such that Equation 4.2 holds is

$$\frac{1}{N_3} = \frac{1}{N_1} = \frac{1}{2}\ SNR_F \tag{4.6}$$

which implies that each use of the forward channel should have the same mean square transmitted power since the forward channel white noise variance is constant ($N_0/2$).

Consider now exactly what part the feedback channel plays in the overall system. From Figure 4-2 the transmitted signal at the second iteration (bottom channel in Figure 4-2) is

$$m_2 = \theta + t_2 s_1 = \theta + t_2(\theta + n_1 + n_2) \tag{4.7}$$

In order to give the term "$t_2 s_1$" some interpretation, the optimal choice of t_2 in Equation 4.4 can be written in terms of the noise variances (rather than the signal-to-noise ratios) as

$$t_2 = \frac{-1}{1 + \overline{n_1^2} + \overline{n_2^2} + (\overline{n_2^2}/\overline{n_1^2})} \tag{4.8}$$

From the study of noiseless feedback systems the optimal system was found to transmit the error signal (or difference between the message θ and the current receiver estimate). If this were true for the noisy system also, then t_2 would be chosen to make $-t_2 s_1$ the minimum mean square error estimate of the receiver's estimate of θ. After one iteration the receiver has only

$$r_1 = \theta + n_1 \tag{4.9}$$

available and would therefore estimate θ as

$$\hat{\theta}_1 = \frac{r_1}{1 + \overline{n_1^2}} \qquad (4.10)$$

If $-t_2 s_1$ in Equation 4.7 is chosen to be the minimum mean square error estimate of $\hat{\theta}_1$ above, then the choice for t_2 is

$$t_{2_{mse}} = \frac{-1}{1 + \overline{n_1^2} + \overline{n_2^2}} \qquad (4.11)$$

which is different from the optimal t_2 in Equation 4.8. If

$$\frac{\overline{n_2^2}}{\overline{n_1^2}} = \frac{N_2(1 + N_1)}{N_1} = \frac{1 + SNR_F/2}{SNR_B} \qquad (4.12)$$

is sufficiently small, then the optimal weighting in Equation 4.8 does not differ greatly from the suboptimal mean square error weighting in Equation 4.11. The implication is that for sufficiently large signal-to-noise ratios in the feedback path the mean square error system is essentially as good as the optimal. The performance should also indicate this fact and will be calculated now.

After two iterations of the forward channel, the receiver has available the two observations

$$r_1 = \theta + n_1 \qquad (4.13)$$

$$r_2 = (1 + t_2)\theta + t_2(n_1 + n_2) + n_3 \tag{4.14}$$

where the noise powers can be calculated as

$$\overline{n_1^2} = N_1 \tag{4.15}$$

$$\overline{n_2^2} = N_2(1 + N_1) \tag{4.16}$$

$$\overline{n_3^2} = N_1[1 - \frac{1}{(1+N_1)(1+N_2)}] \tag{4.17}$$

The performance of the optimal system is given in Equation 4.5. The suboptimal system performance is obtained if t_2 is given by $t_{2_{mse}}$ in Equation 4.11 and the optimal choices for t_3 and t_5 are selected for this given $t_{2_{mse}}$.

The output signal-to-noise ratio for this choice of t_2 can be written down without directly calculating t_3 and t_5. Define a noise covariance matrix $\underline{\underline{N}}$ based on the observation equations for r_1 and r_2 as

$$\underline{\underline{N}} = \begin{bmatrix} N_1 & N_1 t_{2_{mse}} & \\ & & \\ N_1 t_{2_{mse}} & \overline{n_3^2} + t_{2_{mse}}^2 & (\overline{n_1^2} + \overline{n_2^2}) \end{bmatrix} \tag{4.18}$$

The best choice for t_3 and t_5 yield can estimate of θ for which the output signal-to-noise ratio defined in Equation 4.1 is

$$S_{\substack{out\\mse}} = [1 \quad 1+t_{2_{mse}}] \; \underline{\underline{N}}^{-1} \begin{bmatrix} 1 \\ 1+t_{2_{mse}} \end{bmatrix}$$

$$= \frac{[N_1 + N_2(1+N_1)]\;[2N_1(1+N_2) + 1]}{N_1[\,N_1(1+N_1)(1+N_2)^2 - N_1(1+N_2) + N_2\,]} \tag{4.19}$$

A direct comparison of the two system performances given in Equations 4.5 and 4.19 is difficult. Theoretically it follows that the suboptimal system has a lower output signal-to-noise ratio, or

$$S_{\substack{out\\opt}} \geq S_{\substack{out\\mse}} \tag{4.20}$$

for all choices of N_1 and N_2. Equality holds only when the quantity in Equation 4.12 is 0, namely when $N_2 = 0$. This is the noiseless feedback situation from which the suboptimal system was motivated. In Figure 4-3 the output signal-to-noise ratios are compared for $N_1 = .1$ (i.e., a forward signal-to-noise ratio of $SNR_F = 20$). The output signal-to-noise ratio for the feedback systems discussed in Chapter 3; from the figure the effective

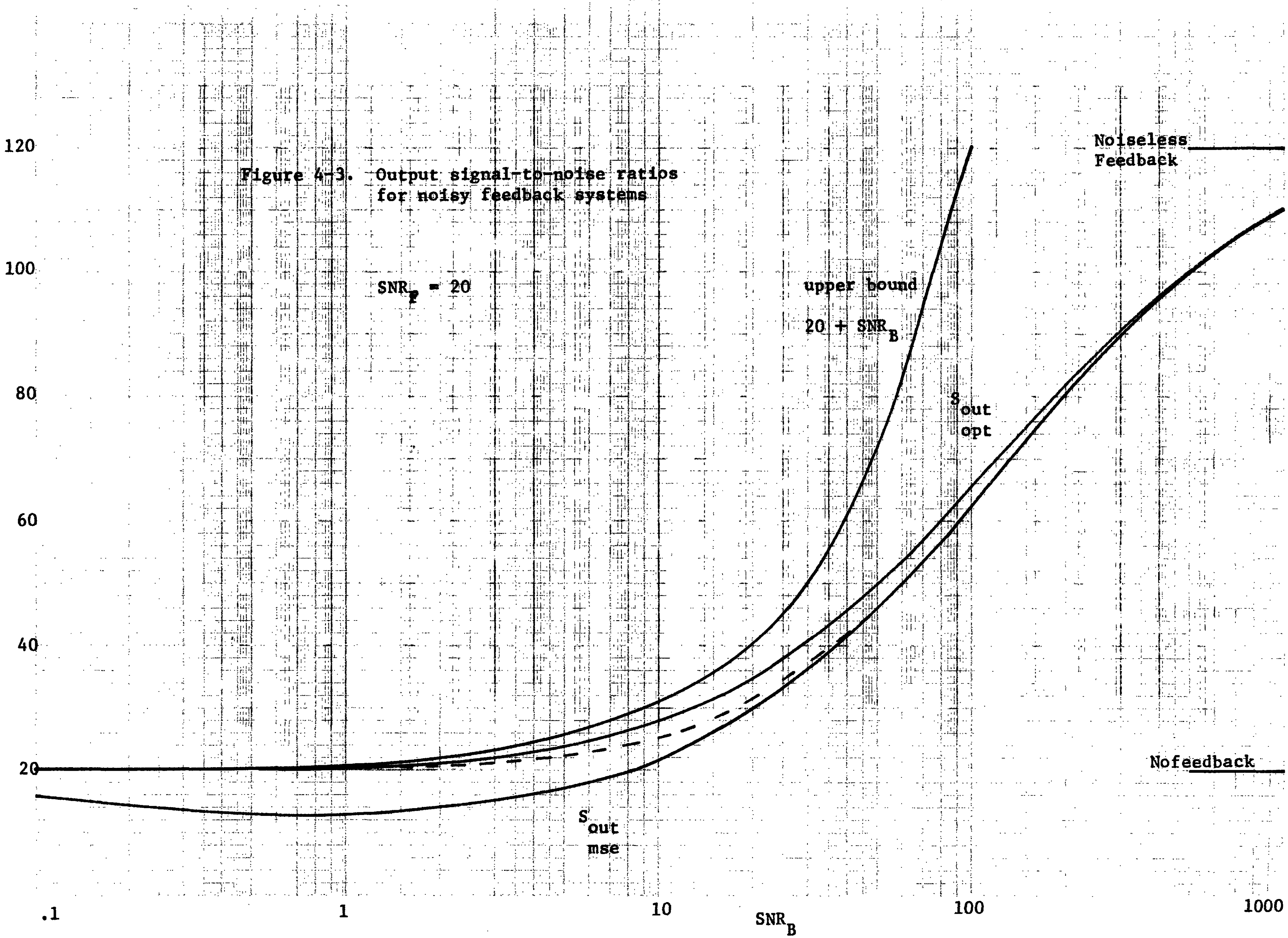


Figure 4-3. Output signal-to-noise ratios for noisy feedback systems

signal-to-noise ratio of the optimal system is always greater than 20 as it must be. The suboptimal system is not always above 20 in output signal-to-noise ratio; for those regions less than 20 the implication is that it is better to disconnect the feedback channel rather than use the suboptimal feedback system.

In the high SNR_B/SNR_F regions of operations the difference in performance of the two systems is quite small. As this ratio approaches infinity, both systems become identical and are the optimal noiseless feedback systems with an output signal-to-noise ratio of 120. For low values of this ratio both systems approach the nofeedback signal-to-noise ratio of 20 although the suboptimal system approaches 20 from below.

One conclusion of this example which is important for the remainder of this chapter is that the suboptimal MSE system is an effective utilization of feedback which is asymptotically (large SNR_B/SNR_F) optimal. Since the optimal solution for arbitrary noisy feedback systems is generally unattainable, the MSE system remains as an easily constructed, potentially effective feedback system whose performance is investigated in the remainder of this chapter.

Clearly the MSE system is probably a poor system whenever the signal-to-noise ratios in the forward and feedback paths are roughly equal. A slight modification of the MSE system can be made to improve the performance.

The above calculations of the MSE system assumed that each use of the forward channel should use equal energy. Allowing N_1 and N_3 to be free (although satisfying Equation 4.2) gives

$$S_{\substack{out\\mse\\mod}} = \frac{1 + (N_1 + N_3)\ (1 + N_2)}{N_1\ [\ N_3(1+N_2) + \frac{N_2}{N_1+N_2+N_1N_2}]} \qquad (4.19a)$$

The maximization of Equation 4.19a subject to constraint Equation 4.2 was carried out numerically and plotted in Figure 4-3 as a dashed line. This same type of modification is used in Section 4.6 in discussing the continuous-time MSE system.

Before proceeding to the investigation of several suboptimum systems, one important observation must be made relative to the previous studies of noisy feedback systems. Many authors (including Kushner [12] on stochastic differential equations and as recently as Omura [31]) state that in the presence of feedback noise the optimal system does in fact estimate the receiver state as outlined in the above MSE system. Elias' result is a direct counterexample to this idea; many more counterexamples are contained in the various MSE systems studied in this chapter.

Some idea as to why the MSE system is really suboptimal can be found by comparing the optimal t_2 in Equation 4.8 with the suboptimal $t_{2_{mse}}$ in Equation 4.11. The major difference is that the optimal choice depends on the relative noise levels of the two forward and feedback channels as well as the absolute levels. Even though both noise levels $\overline{n_1^2}$ and $\overline{n_2^2}$ might be small, $\overline{n_2^2}$ can still be much larger than $\overline{n_1^2}$, implying that the feedback channel is much worse than the forward channel <u>even though both channels are very good</u>. The optimal system recognizes this fact that the feedback channel, even though very good absolutely, is actually relatively poor compared to the forward channel; t_2 in Equation 4.8 will become much smaller in this case than $t_{2_{mse}}$ (smaller in magnitude), thereby tending to rely less on the feedback channel.

4.2 MSE Feedback System Formulation

In this section the MSE noisy feedback system will be formulated in general terms. Succeeding sections apply these general results to specific types of feedback channel noise models. Recall that the MSE system is not the optimal linear noisy feedback system, but it will be shown to be a most useful noisy feedback system.

In Chapter 2 the study of noiseless feedback indicated that the optimal transmitted signal is of the form

$$m(t) = -\frac{g(t)P(t)}{\lambda}(x(t) - \theta) \qquad (4.21)$$

where g(t) is arbitrary and P(t) is known as the solution of a differential equation. x(t) is the receiver state obtained over the noiseless feedback channel. In the presence of feedback noise

Define also the variance or mean square error of the estimate x(t) as

$$V(t) \equiv E[\tilde{x}^2(t)] = -E[x(t)\tilde{x}(t)] \tag{4.26}$$

V(t) is the mean square error the transmitter makes in estimating the receiver state x(t).

Inserting m(t) in Equation 4.23 into the state equation for x(t) (Equation 2.1 with $\phi(t) = 0$) gives

$$\begin{aligned} \frac{d}{dt}\hat{x}(t) &= -\frac{g^2(t)P(t)}{\lambda}(\hat{x}(t) - \theta) + g(t)n(t) \\ &= -\frac{g^2(t)P(t)}{\lambda}x(t) - \frac{g^2(t)P(t)}{\lambda}\tilde{x}(t) \\ &\quad + \frac{g^2(t)P(t)}{\lambda}\theta + g(t)n(t) \end{aligned} \tag{4.27}$$

If $\tilde{x}(t) = 0$ (i.e., no feedback noise), then this equation reduces to the state equation of the noiseless system of Chapter 2. The effect of the feedback channel noise is the additional driving term in Equation 4.27 involving $\tilde{x}(t)$. Now the receiver state is the output of a linear system driven by two noise inputs associated with the forward and feedback channel noises. The actual feedback channel noise enters implicitly in the error waveform $\tilde{x}(t)$.

Since the mean of x(t) in the noisy feedback case is the same as in the noiseless case, define the variance of x(t) as

$$Q(t) \equiv Var[x(t)] \tag{4.28}$$

The differential equation for Q(t) follows easily from the state equation 4.27 as

$$\frac{d}{dt} Q(t) = -\frac{2g^2(t)P(t)}{\lambda} [Q(t) - V(t)] + \frac{N_0}{2} g^2(t) \tag{4.29}$$

$$Q(0) = 0$$

Q(t) is the solution of a linear differential equation and can be broken up (by superposition) into two parts corresponding to the two drives to the equation. Define these two parts as

$$Q(t) = Q_s(t) + Q_v(t) \tag{4.30}$$

where the parts satisfy

$$\frac{d}{dt} Q_s(t) = -\frac{2g^2(t)P(t)}{\lambda} Q_s(t) + \frac{N_0}{2} g^2(t) \tag{4.31}$$

$$Q_s(0) = 0$$

$$\frac{d}{dt} Q_v(t) = -\frac{2g^2(t)P(t)}{\lambda} Q_v(t) + \frac{2g^2(t)P(t)}{\lambda} V(t) \tag{4.32}$$

$$Q_v(0) = 0$$

Of the two components $Q_s(t)$ represents the fluctuations of $x(t)$ due to the forward channel noise assuming noiseless feedback; hence, $Q_s(t)$ is exactly the same as the $\mathrm{Var}[x(t)]$ used in Chapter 2 in the treatment of noiseless feedback systems. $Q_v(t)$ is the fluctuation due to the noisy estimate of the receiver state caused by the noisy feedback channel; it is a linear function of the mean square error $V(t)$ of the receiver state estimate.

Having defined the appropriate functions, the average instantaneous transmitted power of the MSE feedback system is

$$E[m^2(t)] = (\text{noiseless system power}) + \frac{g^2(t)P^2(t)}{\lambda^2}[Q_v(t) - V(t)] \tag{4.33}$$

and the performance (mean square error) of the MSE system is

$$E[(x(T) - \theta)^2] = (\text{noiseless system error}) + Q_v(T) \tag{4.34}$$

The performance (mean square estimation error) of the noisy system is just that of the noiseless plus the effect of the feedback noise on the final state of the receiver, namely $Q_v(T)$. The transmitted energy is slightly more complicated, being the integral of Equation 4.33.

Observe that the characteristics of the MSE system depend only on $V(t)$, variance of estimating the receiver using the feedback channel observations. For whichever type of feedback noise is present, one need only calculate the variance $V(t)$ to determine the overall performance of the noisy feedback system by inserting $V(t)$ into Equation 4.32. For example, if the feedback channel noise is additive white noise, then $V(t)$ is the solution of a Ricatti equation and can be used to determine Equations 4.33 and 4.34.

Suppose that the type of feedback noise has been specified and that $V(t)$ can be calculated. The remainder of the solution is to find the best choice of $g(t)$. Again the functional

$$J = E[(x(T) - \theta)^2] + \lambda \int_0^T dt\, E[m^2(t)] \qquad (4.35)$$

is formed and perturbed to find the optimal $g(t)$. Analytically this is a formidable problem. For any particular $g(t)$ the performance of the noisy feedback system can be determined numerically, but numerical optimization over $g(t)$ appears difficult. Some numerical results for additive noise feedback channels are given in Section 4.6. Lacking the best $g(t)$, the performance of the MSE system is not the best it could be if the optimal $g(t)$ were known.

In the next several sections different types of feedback noise will be studied. By using the results in Equations 4.33 and 4.34 the performance of the MSE system for these noises can be determined.

4.3 Solution for a Constant Variance Estimate

Assume that the feedback channel noise is such that the best estimate x(t) of the receiver state has a constant variance. A constant variance estimate is a good approximation for many noisy feedback systems. For those applications where $SNR_B >> SNR_F$ the feedback channel probably will have a larger bandwidth than the forward channel; hence, perhaps a steady state approximation for the feedback channel is appropriate. Another situation in which a constant variance estimate of the receiver state might arise is a noiseless feedback system in which the receiver state (voltage) x(t) must be quantized due to measurement limitations. For example x(t) might range between -.5 and +.5 volts with a measurement accuracy to the nearest millivolt. Even in situations in which a constant variance approximation is not valid, the results of this section can be used to make a good guess as to how well the noisy feedback system will perform.

Assume that the feedback channel is such that the estimate x(t) has a mean square error of

$$V(t) = V_0 < E[\theta^2] \tag{4.36}$$

which is assumed less than the signal (θ) variance. If the inequality in Equation 4.36 does not hold, the feedback channel cannot be of any value.

Since V(t) is constant, Equations 4.32 and 4.33 can be integrated to evaluate J in Equation 4.35 as

$$J = V_0 + (\overline{\theta^2} - V_0)s_0 - \frac{N_0}{2}\lambda \ln(s_0) \qquad (4.37)$$

which differs slightly from the noiseless expression (Equation 2.28) by the additional terms involving V_0. The quantity s_0 is the same quantity of Chapter 2, namely

$$\frac{1}{s_0} = 1 + \frac{1}{\lambda}\int_0^T dt\, g^2(t) \qquad (4.38)$$

and contains the dependence on g(t). Perturbation of the above equation for J in order to determine the best choice of g(t) leads to an integral square constraint (as before) with a performance of

$$E[(x(T) - \theta)^2] = V_0 + (\overline{\theta^2} - V_0)\exp[-2E_0/N_0] \qquad (4.39)$$

This is the performance of the MSE feedback system subject to a transmitter energy constraint of E_0 and a constant variance receiver estimate.

Examining the performance of the noisy feedback system, it is clear that if V_0 is small enough, then the noisy system performance is essentially the same as the noiseless system. V_0 is small enough if

$$\frac{\overline{\dot{\theta}^2}}{V_0} >> \exp[2E_0/N_0] \tag{4.40}$$

The quantity on the left in Equation 4.40 can be recognized as approximately the signal-to-noise ratio of the feedback channel since it is the reciprocal of the normalized error in estimating the receiver state using the feedback channel observations. The quantity on the right in Equation 4.40 is approximately the effective signal-to-noise ratio of a noiseless feedback system operating over the same forward channel. Thus, the inequality in Equation 4.40 states that, in order for the noisy feedback system to be essentially noiseless, the feedback channel signal-to-noise ratio must be much greater than the effective signal-to-noise of the noiseless feedback system. For $SNR_F = 2E_0/N_0 = 5$ the feedback channel signal-to-noise ratio must be much greater than $e^5 = 150$ for the noisy system to be essentially noiseless.

At the other end of the scale, as $SNR_F = 2E_0/N_0$ tends to infinity the noisy feedback system performance tends to V_0 which is not 0. This result is ample indication that the MSE system is actually suboptimum. As SNR_F tends to infinity, it is easy to construct a nofeedback system which will have a mean square error tending to 0 (which is better than a mean square error of V_0). As noted before, the behavior of the MSE system fails to take into account the <u>relative</u> poorness of the feedback channel.

In Sections 3.5 and 4.1 the usefulness of the feedback systems can be indicated by the effective signal-to-noise ratio of the feedback system (or equivalently the output signal-to-noise ratio). The effective signal-to-noise ratio for the constant variance noisy feedback MSE system is

$$\text{SNR} \xrightarrow[\text{feedback}]{\text{noisy}} \frac{e^{\text{SNR}} - 1}{1 + \dfrac{V_0}{\overline{\theta^2}}\left(1 + e^{\text{SNR}}\right)} \tag{4.41}$$

This quantity is plotted in Figure 4-4 for two values of the feedback signal-to-noise ratio $\overline{\theta^2}/V_0$. Also included in the plot is the effective signal-to-noise ratio for noiseless feedback (Equation 4.41 when $V_0 = 0$) and the effective signal-to-noise ratio of the nofeedback system (which is the actual channel signal-to-noise ratio). As the forward SNR_F increases, the effective signal-to-noise ratio approaches $\text{SNR}_B = \overline{\theta^2}/V_0$ so that the effective signal-to-noise ratio of the MSE system can never be greater than the feedback channel signal-to-noise ratio. Even if there is no noise in the forward channel, the MSE system will not have a mean square error of 0 in estimating θ, but the performance will be a mean square error of V_0 (the feedback channel variance). Again, this is another demonstration that the MSE system is suboptimum.

The effective signal-to-noise ratio in Equation 4.41 can be converted to a probability of error (P_e) for the message coding

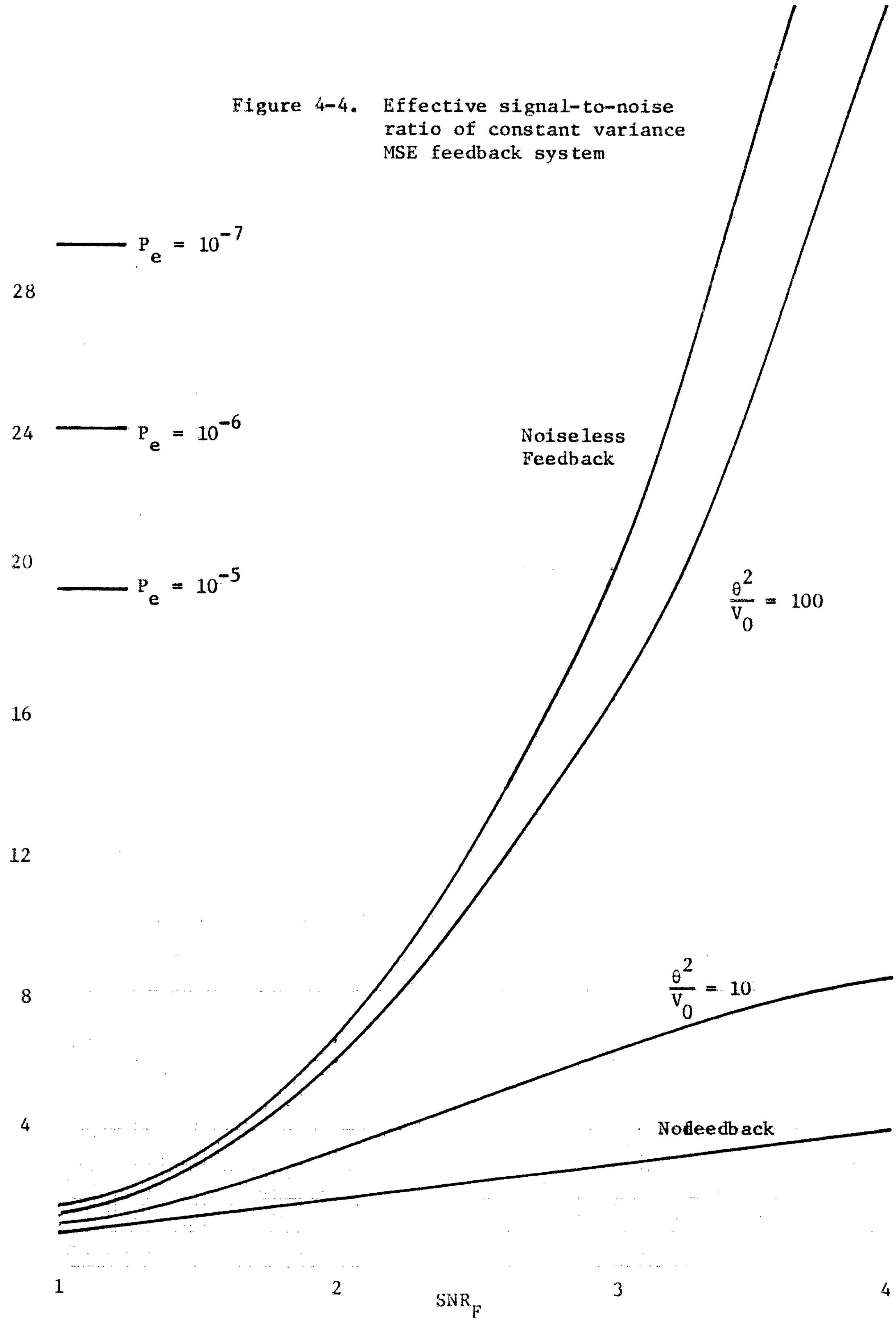


Figure 4-4. Effective signal-to-noise ratio of constant variance MSE feedback system

scheme of Chapter 2 by noting the dependence of P_e on the effective signal-to-noise ratio. If the effective signal-to-noise ratio for the system is SNR, then P_e is

$$P_e \simeq 2\,Q\left[\frac{SNR^{1/2}}{2(M-1)\,\overline{\theta^2}}\right] \tag{4.42}$$

where the Q function is the area in the tail of the normal density (Equation 2.59). For comparison purposes several points can be calculated from Equation 4.42 as (for M=2)

P_e	SNR
10^{-5}	19.4
10^{-6}	24
10^{-7}	29.2

This scale change could be used in Figure 4-4 to convert the vertical scale from effective signal-to-noise ratio to P_e to determine the P_e improvement of the feedback system.

As an example of how these results might be applied to a physical system, suppose that the forward channel has $SNR_F = 3.5$ and operation with a $P_e \leq 10^{-6}$ for a single bit is desired. From the table an effective signal-to-noise ratio of 24 is needed, but without feedback the effective signal-to-noise ratio would only be 3.5. If a noiseless feedback link were available, it would boost the effective signal-to-noise ratio to $e^{3.5} - 1 = 32.1$ which is more than enough for a $P_e \leq 10^{-6}$. For practical reasons there are no voltmeters, say, which could read the receiver state x(t) to an infinite number of decimal places for the noiseless feedback link; hence, truly noiseless feedback is not possible. Suppose instead a binary quantizer is to be purchased for measuring x(t) to some finite number of bits. The exact number of bits of quantization needed to achieve $P_e \leq 10^{-6}$ can be determined so that no extra expense is involved in purchasing the quantizer.

First the required fractional variance of the "noisy" feedback link can be calculated by equating Equation 4.41 to 24 and solving for

$$\frac{V_0}{\overline{\theta^2}} \leq .0099 \qquad (4.43)$$

Next the fractional error of a k-bit quantizer must be determined and then equated in Equation 4.43. The error due to quantization is (approximately) uniformly distributed in an interval of width 2^{-k}. For example, a 3-bit system which measures x(t) as .0625 implies that x(t) actually lies in the interval 0 to .125, an interval of width $2^{-3} = .125$. If the error in estimating x(t) is uniform in the interval, then the fractional variance of the estimate is the variance of a random variable uniform in an interval of width 2^{-k} or

$$\frac{v_0}{\overline{\theta^2}} = \frac{2^{-2k}}{12} \qquad (4.45)$$

Equating the fractional variance of the quantizer to the required feedback error in Equation 4.43 implies that $k \geq 2$, or that at least 2 bits of quantization are needed. Thus, if the receiver informs the transmitter of the quarter of the interval [-.5,.5] the receiver state is in, the system will operate with an effective signal-to-noise ratio of 27.2, more than adequate to achieve the desired $P_e \leq 10^{-6}$. Observe that if only the sign of x(t) (1 bit) is available at the transmitter, this is sufficient to operate at $P_e = 2 \times 10^{-5}$ which is still an improvement over the nofeedback $P_e = 6 \times 10^{-2}$.

Using the result in Equation 4.41 for the effective signal-to-noise ratio of the MSE system, the expected noisy feedback improvement for different types of feedback systems can be approximated. In the next section the effects of delay are evaluated.

4.4 Approximate Performance of Feedback Systems with Delay

Suppose that instead of measurement noise, a constant loop delay of t_0 seconds is present. The transmitter only has available $x(t-t_0)$ at time t and must estimate x(t). The procedure in this section will be to calculate (approximately) the fractional estimation error made by the transmitter in estimating x(t) given $x(t-t_0)$. Then this error will be used in Equation 4.41 to determine the MSE system performance in the presence of delay.

In Chapter 2 in the study of noiseless systems the choice of g(t) in the system was found to be almost arbitrary; however, there was a unique g(t) which seemed to give the best operational characteristics (e.g., constant transmitted power, constant bandwidth). Assume that this choice of g(t) is made. The state equation for x(t) for this choice of g(t) is

$$\frac{d}{dt}\, x(t) \;=\; -\, k\, x(t) \;-\; k\, x(t) \;+\; k\,\theta + \; g(t)n(t) \qquad (4.46)$$

where k is the constant pole location

$$k \;=\; \frac{2E_0}{N_0 T} \;=\; \frac{SNR_F}{T} \qquad (4.47)$$

x(t) is the estimation error in estimating x(t) given $x(t-t_0)$.

In order to approximate the mean squarred error $E[x^2(t)]$, consider a stationary one-pole random process a(t) with a **spectrum**

$$S_a(f) \;=\; \frac{A}{(2\pi f)^2 + k^2} \qquad (4.48)$$

Given $a(t-t_0)$, the minimum variance estimate of $a(t)$ has a fractional steady-state mean square error of

$$\frac{E[(\hat{a}(t) - a(t))^2]}{E[a^2(t)]} = 1 - \exp[-2kt_0] \tag{4.49}$$

where the estimate $\hat{a}(t)$ is

$$\hat{a}(t) = E[a(t) \big|_{a(t-t_0)}] \tag{4.50}$$

The above expression for the fractional estimation error can be applied directly to the feedback system to give the approximate fractional mean square error of the transmitter's estimate of the receiver state. Admittedly the feedback system estimation error is not in the steady state, but Equation 4.49 is a reasonable upper bound to the estimation error. It is an upper bound because at the start of the interval the estimation error is zero (the transmitter knows that the receiver is initially at rest) and increases toward the steady state value. The system is roughly in the steady state at the end of the interval since $kT = SNR_F$ which is normally around 5 or more if the forward channel is not too noisy.

Inserting Equation 4.49 into Equation 4.41 gives the approximate signal-to-noise ratio improvement of the MSE system in the presence of a loop delay of t_0 seconds as

$$\text{SNR} \xrightarrow{\text{delay } t_0} \frac{\exp[\text{SNR}] - 1}{1 + (1 - \exp[-2kt_0])(1 + \exp[\text{SNR}])} \tag{4.51}$$

or by using the value of k in Equation 4.47

$$\text{SNR} \xrightarrow{\text{delay } t_0} \frac{\exp[\text{SNR}] - 1}{1 + (1-\exp[-2\text{SNR}(t_0/T)])(1+\exp[\text{SNR}])} \tag{4.52}$$

as the effective signal-to-noise ratio of the MSE system with delay t_0.

The important quantity which indicates whether or not the delay t_0 is important to the otherwise noiseless feedback system is

$$kt_0 = \text{NSR}_F \quad (\frac{t_0}{T}) \tag{4.53}$$

which must be small compared to $\exp[-\text{SNR}_F]$ in order for the system improvement in Equation 4.52 to essentially the same as the noiseless delayless feedback system.

As an example of the application of Equation 4.52, consider the system used for an example in the previous section. Suppose that there is no quantization error, but that operation at $P_e \leq 10^{-6}$ is desired for $\text{SNR}_F = 3.5$. The maximum loop delay for this operation is desired. In Equation 4.43 the maximum fractional variance allowable in estimating the receiver state was determined; therefore, Equation 4.49 must be less than

$$1 - \exp[-2(3.5)(t_0/T)] \leq .0099 \tag{4.54}$$

or solving

$$t_0 \leq \frac{T}{700} \tag{4.55}$$

in order for $P_e \leq 10^{-6}$ in the MSE feedback system.

Delay is a significant problem for the MSE feedback system because the estimation error increases rapidly with increasing delay. Heuristically the bandwidth of the feedback system is roughly the location of the pole, namely 3.5/T. It would seem that as long as the delay in the loop is much less than the reciprocal bandwidth, the feedback system performance should be almost as good as the noiseless feedback system. A delay of 1/10 the reciprocal bandwidth would appear (from an engineering point of view) to be reasonable for ignoring the delay; this would correspond to a delay T/35. The inequality in Equation 4.55 is much stronger and tends to negate the "engineering point of view" for this system.

Authors who have dealt with discrete-time feedback systems have indicated that loop delay can be handled in discrete-time feedback systems by time multiplexing several feedback systems so that each system effectively has no delay. For every additional unit of time delay there must be an additional message and receiver so that many separate feedback systems (using the single channel at different times) may be necessary to eliminate the effects of the delay. This solution to delay is not directly applicable to continuous-time systems unless one proceeds by discretizing the continuous-time system.

4.5 Additive White Feedback Noise

Consider a feedback system in which the feedback channel is an additive white noise channel like the forward channel. This model is one of the more realistic feedback system models since no system (or channel) is really noiseless. The most general linear system for such a feedback communications problem is shown in Figure 4-1 for the discrete-time case. The four filters LRF1,...,LRF4 are to be chosen optimally subject to the forward and reverse transmitter power constraints. Since the optimal solution is unknown, the performance of the MSE system will be studied. In this section the techniques of the preceding two sections will be used to approximate the performance of the MSE system operating with additive white noise in the feedback channel. Section 4.6 gives more accurate numerical results for several types of MSE systems.

If the feedback channel has additive white noise of spectral density $W_0/2$, then a feedback transmitter energy constraint must be imposed to make the problem meaningful. Assume that the receiver is restricted to an energy E_B in transmitting back to the transmitter in the interval [0,T]. This implies that a reverse or backward signal-to-noise ratio can be defined

$$SNR_B = \frac{2E_B}{W_0} \tag{4.56}$$

in a manner analogous to the definition of SNR_F, the forward signal-to-noise ratio.

Considering the one-state receiver outlined in Section 4.2 and drawn in Figure 4-5, if g(t) is chosen to make x(t) have a constant pole location (bandwidth), then the fractional estimation error in estimating x(t) at the transmitter (based on the feedback channel observations) is approximately the steady state realizable Wiener filtering error if the receiver sends back a multiple of the receiver state x(t). If a(t) is a one-pole (located at -k) stationary process observed in white noise, then if

$$A = \frac{4P_{ave}}{N_0 k} \tag{4.57}$$

and $N_0/2$ is the white noise density and P_{ave} is the average transmitted power, the normalized realizable steady state error is estimating a(t) observed in the white noise is

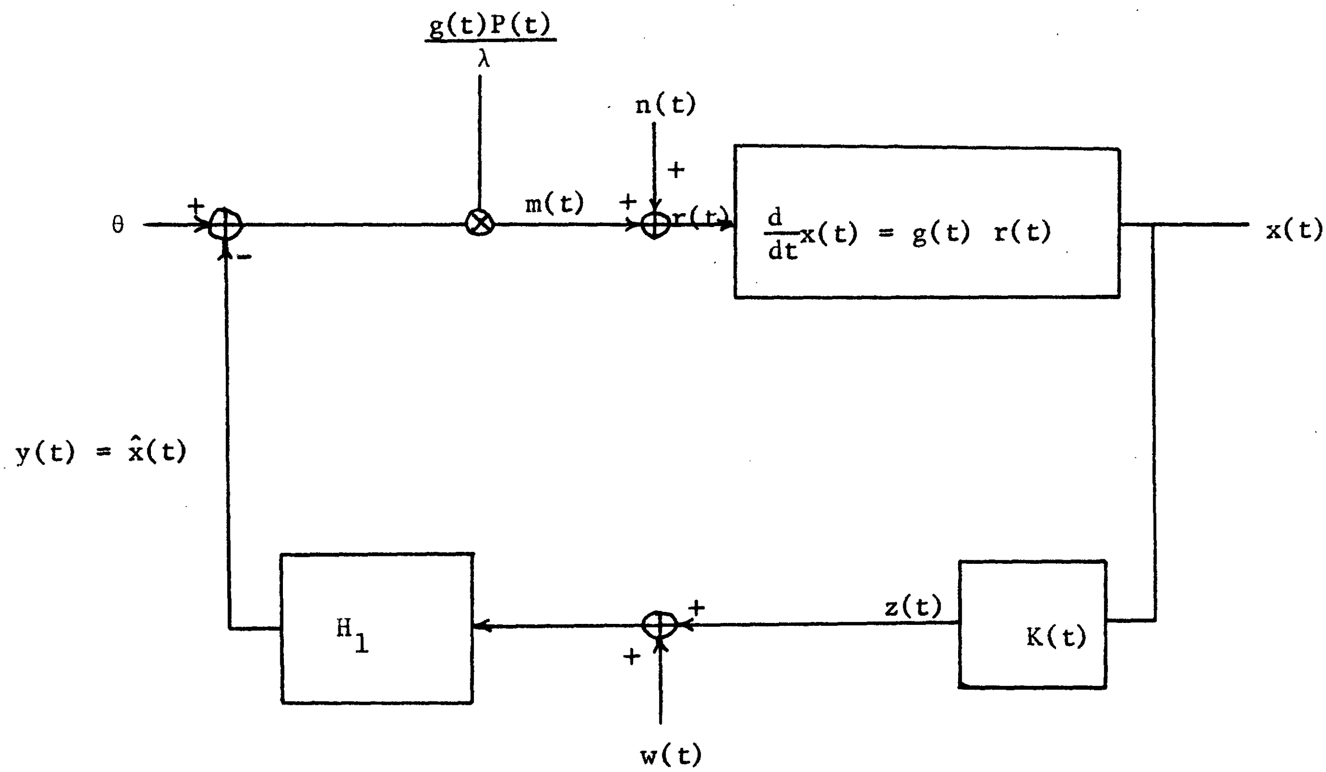


Figure 4-5. Additive noise feedback system

$$\xi = \frac{2}{1 + (1+A)^{1/2}} = \frac{2}{1 + (1 + 2SNR_B/SNR_F)^{1/2}} \qquad (4.58)$$

The second part of Equation 4.58 follows by observing k is given in Equation 4.47 and $P_{ave} = E_B/T$. The above expression for the normalized estimation error is pessimistic due to the fact that the true feedback system starts out at t = 0 with zero error and only approaches the error in Equation 4.58 as steady state is reached; T may not be long enough for steady state to be reached, in which case the fractional error is never as great as that in Equation 4.58.

With these reservations the value of ξ in Equation 4.58 can be inserted in Equation 4.41 to give the approximate performance (signal-to-noise ratio improvement of the MSE system)

$$SNR \xrightarrow[\text{noise}]{\text{white}} \frac{\exp[SNR] - 1}{1 + \dfrac{2(\exp[SNR] + 1)}{1 + (1+2SNR_B/SNR)^{1/2}}} \qquad (4.59)$$

The above expression for the effective signal-to-noise ratio of the MSE system implies fairly poor performance unless $SNR_B >> SNR_F$. For example, recall the system in Section 4.3 and 4.4 which had a $SNR_F = 3.5$ and needed a fractional variance less than .0099 to achieve $P_e \leq 10^{-6}$. In order to achieve this performance with white noise in the feedback path, the signal-to-noise ratio in the backward direction must be approximately $SNR_B = 10000$, a rather large signal-to-noise ratio.

For these cases when SNR_B is much much greater than SNR_F the MSE system is almost optimal; this is the limit in which the MSE system is asymptotically optimal. The implication is that the optimal system (if it were known) would not give significantly better improvement than the MSE system performance calculated in Equation 4.59, 4.52, and 4.41. Whenever SNR_B is not much much greater than SNR_F, the MSE system performance (which is still given by these equations) is much poorer than would be obtained from the optimal linear system.

In the next section the basic MSE system is modified slightly to obtain system improvement even when SNR_B is on the order of SNR_F. The modification is an extra gain parameter which takes into account the effect of the feedback noise on the overall system performance. As noted in Section 4.1 in the study of Elias' system, this appears to be the fault of all MSE systems.

4.6 Numerical Results for Additive Feedback Noise

Considering the same additive white noise feedback problem, suppose $SNR_F = 5$ and $SNR_B = 100$. Using these values in Equation 4.59 implies that the addition of the MSE feedback link changes the effective signal-to-noise ratio from 5 (without feedback) to 3.6 with feedback. The addition of feedback has degraded the performance; hence, this is a situation in which the MSE system for utilizing feedback is quite suboptimal. The modification of the MSE system developed in this section will achieve an effective signal-to-noise ratio of about 19 which is certainly a much better feedback system. In fact the modified system is modified in a manner which prohibits the effective signal-to-noise ratio of the feedback system from being less than SNR_F, the nofeedback signal-to-noise ratio.

The structure of the feedback system is essentially the same as that shown in Figure 4-5. The receiver transmits a multiple of the receiver state x(t) back to the transmitter over the feedback channel. The filter H_1 which generates y(t) = x(t) is, therefore, a Kalman filter to be calculated shortly. The feedback signal is

$$z(t) = K(t)\, x(t) \tag{4.60}$$

and must satisfy the constraint on feedback transmitter energy

$$E_B = \int_0^T dt\, E[z^2(t)] \tag{4.61}$$

The differential equation of the receiver is

$$\frac{d}{dt} x(t) = -\frac{g^2(t)P(t)}{\lambda} y(t) + \frac{g^2(t)P(t)}{\lambda} + g(t)\, n(t) \tag{4.62}$$

where $y(t)$ is the transmitter's estimate of the receiver state. The filter H_1 which generates $y(t)$ observes $z(t) + w(t) = K(t)x(t) + w(t)$. The minimum mean square error estimate $y(t)$ which estimates $x(t)$ is a Kalman filter; the only difference from the ordinary Kalman [17] formulation is that the "message" $x(t)$ depends on the estimate $y(t)$. Nevertheless, the filter H_1 satisfies

$$\frac{d}{dt} y(t) = -\left[\frac{2K^2(t)V(t)}{W_0} + \frac{g^2(t)P(t)}{\lambda}\right] y(t) + \frac{g^2(t)P(t)}{\lambda} \theta + \frac{2K^2(t)V(t)}{W_0} [z(t) + w(t)] \qquad (4.63)$$

$$y(0) = 0$$

where the variance $V(t)$ of the estimate $y(t)$ satisfies the Ricatti equation

$$\frac{d}{dt} V(t) = \frac{N_0}{2} g^2(t) - \frac{2K^2(t)V^2(t)}{W_0} \qquad (4.64)$$

$$V(0) = 0$$

This value of $V(t)$ is what is required to insert in Equation 4.29 to determine the performance of the noisy feedback system.

Unfortunately direct use of $V(t)$ in this manner is not possible without first determining the feedback channel signal-to-noise ratio.

The noisy feedback MSE system is completely specified by Equations 4.62 and 4.63, leaving $g(t)$, $K(t)$, and λ is determined by the forward transmitter energy constraint. The remaining $g(t)$ and $K(t)$ are arbitrary and can be varied to improve the performance.

Recall that for the noiseless case and the constant variance estimate case the shape of $g(t)$ did not affect the system performance; this fact led to other (power, bandwidth) considerations to determine a suitable $g(t)$. Now, however, the estimate variance is not constant; hence, the best choice of the shape of $g(t)$ might be expected to affect the system error.

Consider what the choice of $g(t)$ affects. The transmitted signal for all MSE feedback systems is

$$m(t) = -\frac{g(t)P(t)}{\lambda}(y(t) - \theta) \tag{4.65}$$

where $y(t)$ is the estimate of the receiver state. If the estimate $y(t)$ is not very good during some part of the interval, one would hope that the transmitter weighting $g(t)P(t)$ would be relatively small in that section of the transmission interval. Since in the time-varying system the transmitter knows that the receiver state is 0 initially, the variance of the estimate $y(t)$ is small near $t = 0$, but generally increases toward a steady-state value. Perhaps the best shape of $g(t)P(t)$ would be large initially and tapering off (perhaps to zero) as the estimate becomes worse later in the interval. Such a design philosophy would certainly guarantee a signal-to-noise ratio (effective) no worse than the nofeedback system since all of the energy could be concentrated early in the interval. Toward this end several different shapes of $g(t)$ are tried. By varying the amplitude of $g(t)$ to optimize the feedback system performance, the MSE feedback system performs quite well. The exact effect of the shape of $g(t)$ on the transmitter weighting $g(t)P(t)$ is quite complex.

In order to calculate the performance of the MSE feedback system, several equations in addition to Equation 4.64 are needed. Equation 4.29 specifies the behavior of $Q(t)$, the variance of the receiver state $x(t)$. Define

$$\mu(t) = E[x(t) - \theta] = E[y(t) - \theta] \qquad (4.66)$$

in order to calculate mean values. It follows that the differential equation for $\mu(t)$ is

$$\frac{d}{dt}\mu(t) = -\frac{g^2(t)P(t)}{\lambda}\mu(t) \qquad (4.67)$$

$$\mu(0) = -\theta$$

Some convenience is obtained by writing P(t) directly in terms of μ(t)

$$P(t) = \frac{\mu(T)}{\mu(t)} \tag{4.68}$$

$$\mu(T) = \frac{-\theta}{1 + \frac{1}{\lambda}\int_0^T dt\, g^2(t)}$$

For these definitions the mean square error of the MSE system in estimating θ is

$$E[(x(T) - \theta)^2] = Q(T) + \mu^2(T) \tag{4.69}$$

with a forward signal-to-noise ratio of

$$SNR_F = \frac{2}{N_0}\int_0^T dt\, [Q(t) - V(t) + \mu^2(t)]\, \frac{g^2(t)P^2(t)}{\lambda^2} \tag{4.70}$$

and a feedback channel signal-to-noise ratio

$$SNR_B = \frac{2}{W_0}\int_0^T dt\, [Q(t) + (\theta+\mu(t))^2]\, K^2(t) \tag{4.71}$$

Summarizing, the MSE feedback system in Figure 4-5 (for H_1 being the Kalman filter in Equation 4.63) has been analyzed to yield the performance and constraints of Equations 4.69 to 4.71 in terms of functions specified by differential equations. The functions g(t) and K(t) and the constant λ are free subject to the signal-to-noise ratio constraints.

At this point the analysis stopped and numerical evaluation of these equations started. The optimization over K(t) and g(t) even on a computer is not at all straightforward since the functions are almost completely aribtrary. Instead some modified optimization was carried out numerically to determine the performance of several MSE systems.

The feedback transmitter gain K(t) was assumed to be constant, that constant which made SNR_B be the desired value in Equation 4.71. The value of λ is that which makes SNR_F equal the desired forward signal-to-noise ratio. For this procedure there are no constraints on g(t); g(t) was selected as a parameterized waveform such as

$$g(t) = A \sin(\frac{\pi t}{T}) \tag{4.72}$$

where the optimal choice of the constant A was found numerically. For this class of problems the computational chore is to find the two constants (λ and K) which yield the correct SNR_F and SNR_B and then choosing the best value of A which minimizes

Equation 4.69.

The computation was carried out and plotted in Figure 4-6 for the probability of error P_e as a function of the feedback signal-to-noise ratio (SNR_B) for different values of SNR_F. g(t) was the half sine wave above with T = 1 (T does not affect the performance). Other data in the figure include the performance of the system without feedback and with noiseless feedback. Observe that the modified MSE system always performs better than the nofeedback system which is in contrast to other MSE system studied in this chapter which perform worse than nofeedback in some regions.

The system performance plotted in Figure 4-6 is for g(t) = half sine wave. In Figure 4-7 several different shaped g(t) are compared at a fixed forward SNR_F = 5. The three types of g(t) considered are: 1) constant, 2) half sine wave cycle, 3) full sine wave cycle. From the results plotted in Figure 4-7 each choice of g(t) performs slightly differently although there is no clear "best" g(t) of these three shapes.

The choice of g(t) (as in the noiseless case) affects the power distribution at the transmitter. Roughly, with noisy feedback the instantaneous average transmitted power is proportional to $g^2(t)$; this result is an empirical one based on the numerical results. For g(t) a constant, the transmitted power is approximately constant except for a gradual peaking at the end of the interval. The power distribution in the feedback channel was always ramp-like for the constant K(t).

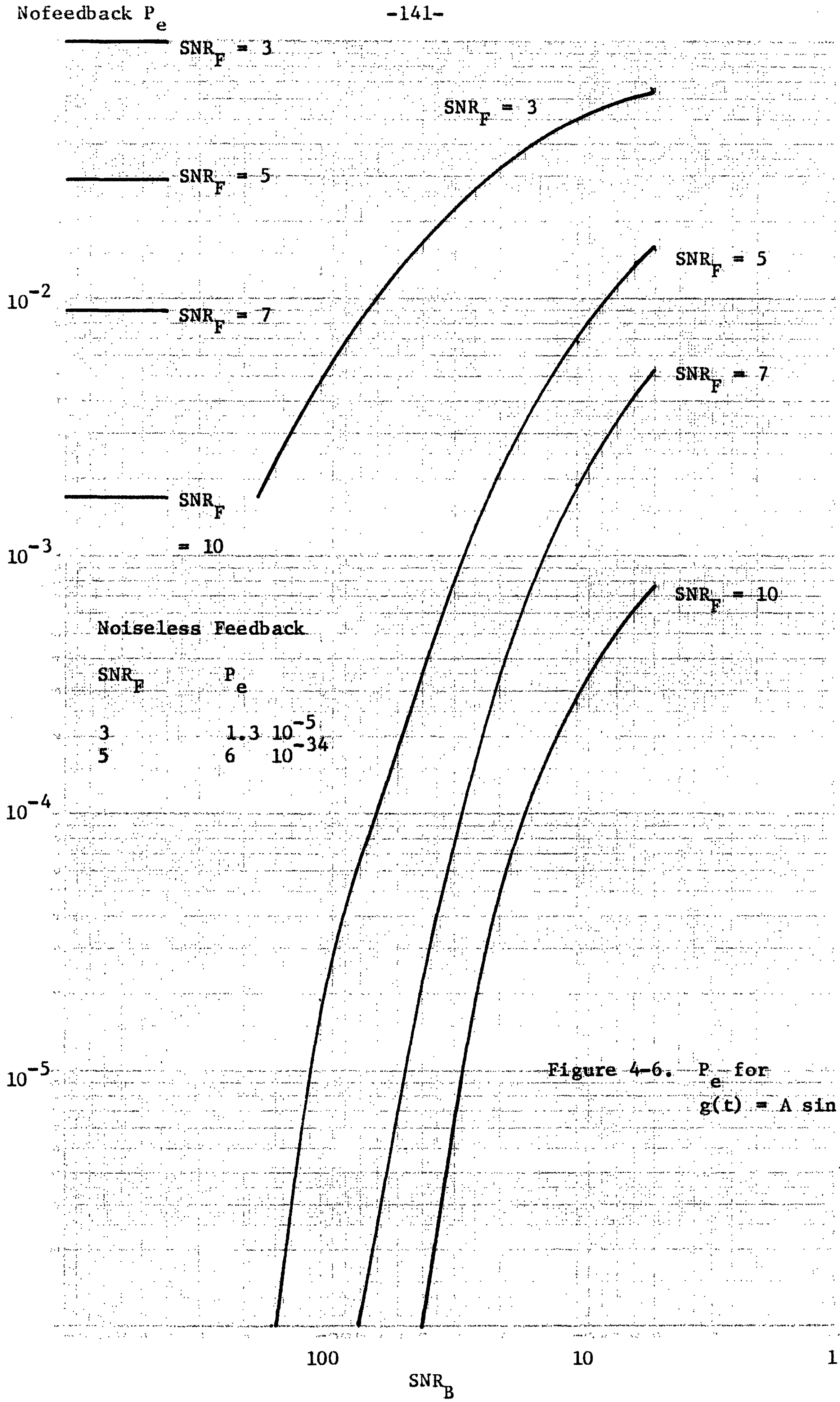


Figure 4-6. P_e for $g(t) = A \sin(\pi t/T)$

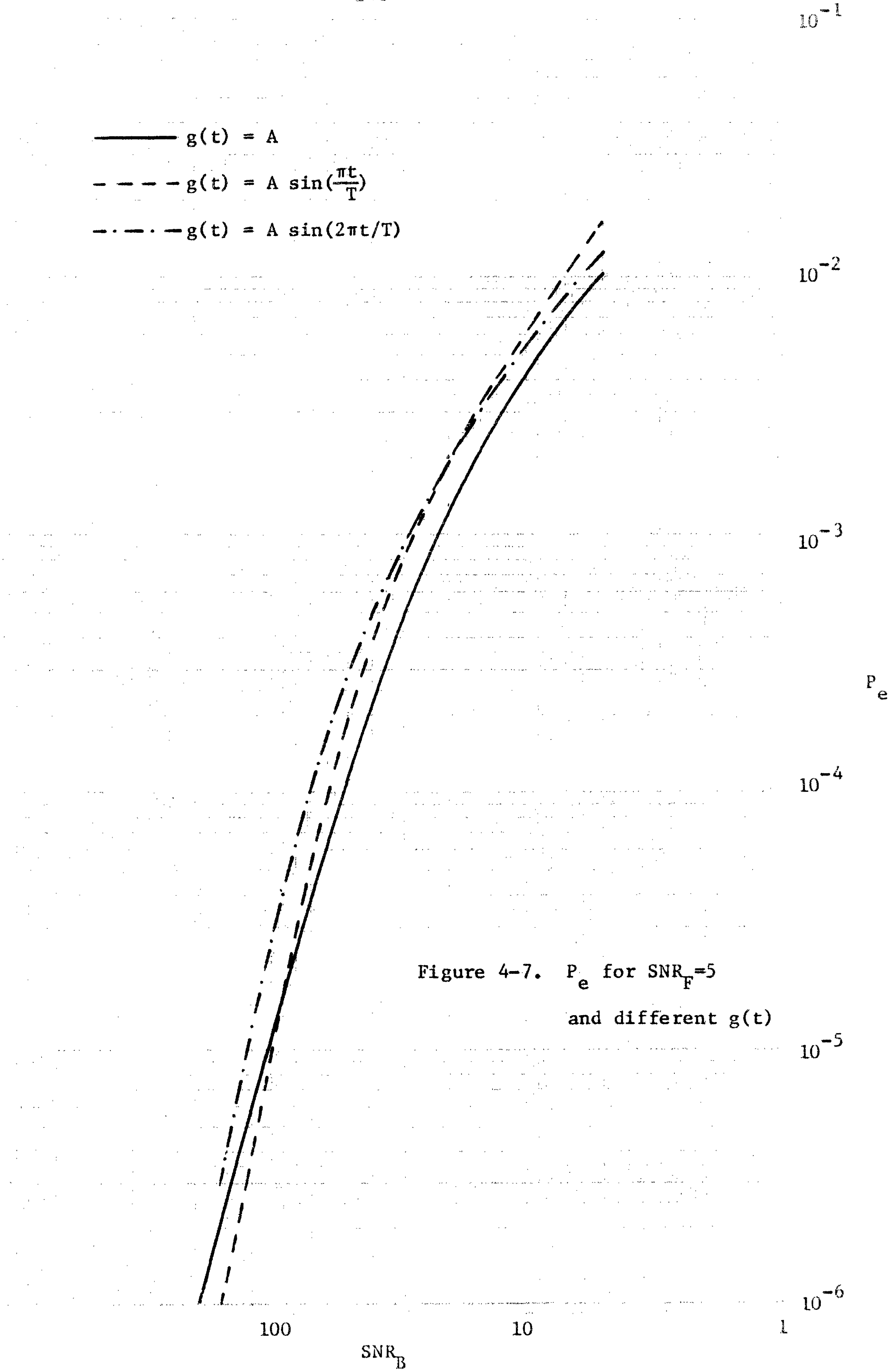


Figure 4-7. P_e for SNR_F=5 and different g(t)

As noted in Chapter 3, feedback systems can be viewed as reducing the effective channel noise or as increasing the effective signal-to-noise ratio. In Figure 4-8 the data from Figure 4-6 are plotted to indicate the relationship between the effective signal-to-noise ratio of the feedback system and the feedback channel signal-to-noise ratio. Figure 4-8 also indicates the increase in effective signal-to-noise ratio to be expected if the feedback channel were noiseless which is substantially larger than the noisy feedback results.

A comparison of the results of Figure 4-8 with the approximate results of Equation 4.59 shows that the latter is quite pessimistic for the SNR_B range plotted in the figure. For example, for $SNR_F = 3$ and $SNR_B = 180$, Equation 4.59 predicts an effective signal-to-noise ratio of 4.3 whereas the figure gives 10. One reason for the difference is that the true estimation error is actually less than half that given by Equation 4.58; another reason is that better use of the feedback channel is obtained by a different g(t).

4.7 Comments on Noisy Feedback Systems

In this chapter digital systems employing noisy feedback channels have been analyzed. Approximate solutions have been obtained for different types of feedback channels. The formulation of the exact optimal feedback system is present, but the problem remains unsolved. Guided by the results of Elias, it appears that

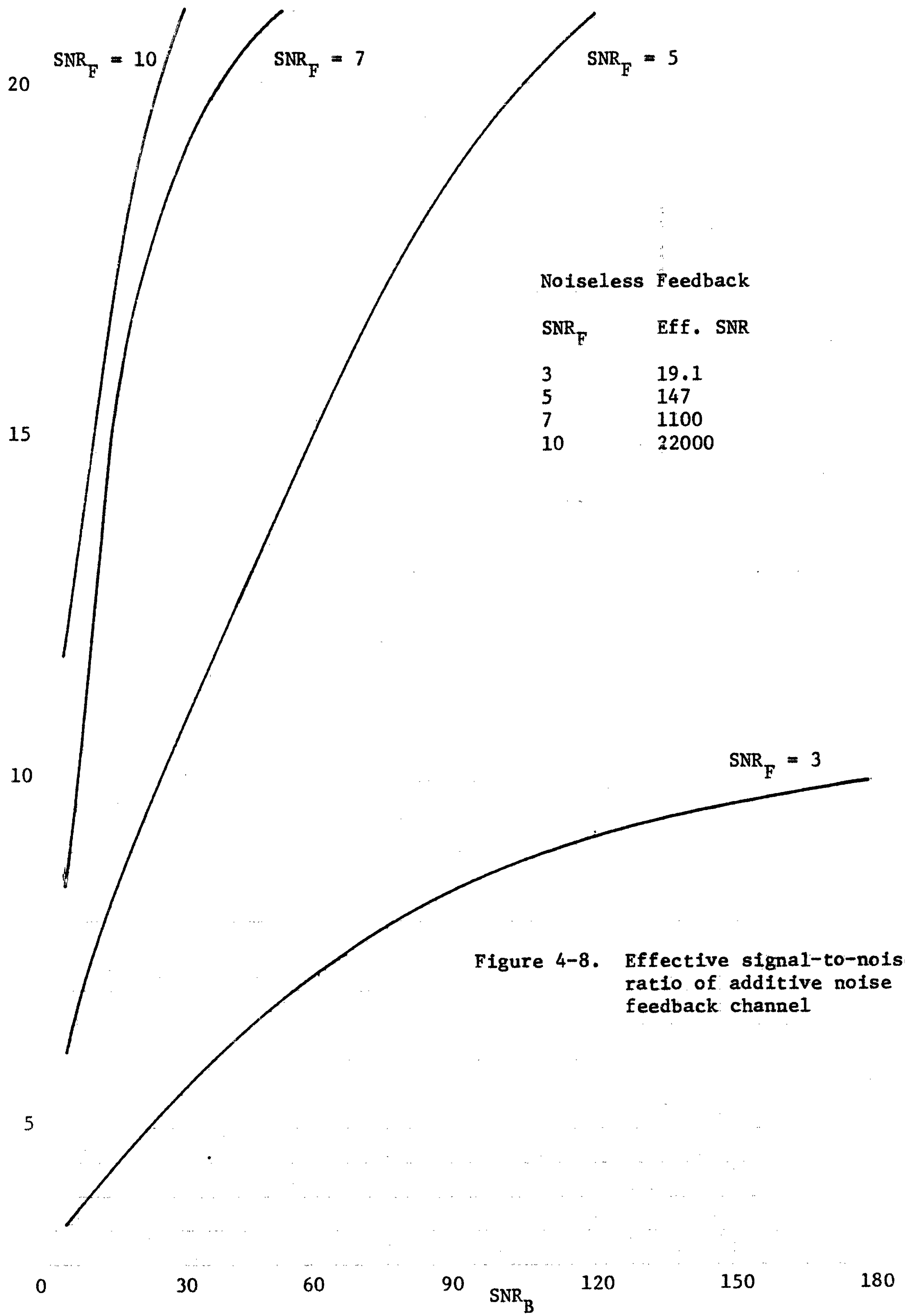


Figure 4-8. Effective signal-to-noise ratio of additive noise feedback channel

the optimal feedback system (if it could be determined) does not perform much better than the suboptimal MSE systems studied for large SNR_B/SNR_F ratios.

By starting from the known solution for noiseless feedback in which the transmitter sends the error between the receiver and the message parameter (θ), the suboptimal MSE system operating in the presence of noisy feedback estimates the receiver state and sends the estimated error of the receiver as its transmitted signal. Many have stated that this technique yields the optimal noisy feedback systems, but Elias' results prove otherwise. Nevertheless, the MSE system is asymptotically optimal and provides a system capable of using the noisy feedback channel. A close examination of Elias' system indicates that the MSE system does not consider the effects of the feedback noise on the overall system performance, but only the effect of the feedback noise as it alters the transmitted signal power. Several examples of this chapter demonstrate this fact.

The performance of the noisy feedback systems discussed depend on the fractional or normalized mean square estimation error of the transmitter's estimate of the receiver state. Knowledge of this normalized error enables one to approximate the performance of the MSE system. For most practical cases this error is far too large for the noisy feedback system to obtain the dramatic improvement of noiseless feedback. The performance improvement of noisy feedback systems computer here may or may not be large enough to warrant the use of a linear feedback system. More complex signalling techniques would be necessary to improve the system performance.

The signal-to-noise improvement calculated and plotted is only for the digital example. Other systems (such as analog estimation problems) have signal-to-noise ratios associated with them which may or may not be related to the digital system signal-to-noise ratio. The characterization of feedback systems by the increase in effective signal-to-noise ratio is a convenience which is limited to the particular system analyzed. The effective signal-to-noise ratio for analog estimation problems would have to be calculated separately.

CHAPTER 5

Summary and Extensions

At this point a summary of the preceding three chapters will be given followed by a series of related problems which are extensions of the feedback systems studied here. Some of these problems do not appear promising in view of the results of this thesis; others are appropriate for further research.

5.1 Summary

As reported in Chapter 1, previous studies of feedback systems have been restricted to discrete-time versions of noiseless feedback systems. Many similar systems have been described which operate at error-free rates up to channel capacity.

In Chapter 2 the noiseless feedback problem is formulated directly in a continuous-time variable, thereby saving the limiting argument of discrete-time versions and providing a direct differential equation structure which is easily implemented. Since the channel noise model is usually a continuous-time white noise model, this solution is more desirable. In other cases a sampled-time system might be appropriate.

Formulation of the problem in continuous-time necessitated the introduction of two relatively new mathematical disciplines: stochastic differential equations and stochastic optimal control.

The two topics are very closely related.

The formulation of the digital communications problem (transmitting one of a finite number of messages) in terms of stochastic optimal control and dynamic programming allowed a solution by judicious guessing of the optimal value function. By imposing a linear receiver the optimal transmitter is also linear. For other linear Gaussian problems with quadratic costs similar solution guessing can be most useful. Also for linear stochastic problems the interpretation of the white noise in the differential equations in noncontroversial since the two interpretations (Ito and Stratonovich) lead to identical mathematics and system performance.

The evaluation of the performance of the differential systems necessitates the derivation of the corresponding differential equations for variances and powers. By using the results of the Ito calculus to develop a simple algebra for treating stochastic differential equations, the step from stochastic equation to the deterministic power or variance equation is obvious. These techniques are most useful for manipulating all stochastic state variable systems, particularly linear systems.

The performance of the continuous-time system evaluated in Chapter 2 is identical to the limiting form of Butman's [23] discrete-time system. By suitably altering the work of the other discrete-time authors to improve their system performance, the performance there can be made identical. The performance of feedback systems is conveniently expressed either in terms of the probability of error or in terms of an increase in effective signal-to-noise ratio.

Although the performance of the systems (both discrete- and continuous-time) can be made identical, one difference has been noted. For continuous-time systems the waveform g(t) (and hence the power distribution) is almost arbitrary whereas the discrete-time system requires a unique g_k and a uniform power distribution in the transmission interval. For many reasons the choice of g(t) such that a uniform power distribution results in the continuous-time feedback system is a most desirable choice of g(t); this choice is not necessary, however.

In Chapter 3 the topic of analog estimation communication using noiseless feedback is discussed and developed. This topic has not been studied before and represents further application of feedback channels. Much of the results of the analog estimation systems were strongly motivated by the use of noiseless feedback in the digital problem studied in Chapter 2.

In Chapter 3 simple linear feedback systems are shown to achieve the rate-distortion bound on mean square estimation error. Achieving the rate-distortion bound in analog systems is the equivalent of achieving channel capacity in digital systems. Noiseless feedback systems can achieve the rate-distortion bound for the transmission over a white noise channel of a Gaussian random variable or a Gaussian random process in a finite interval of time or a stationary process over an infinite (or very long) time interval. The feedback system is linear in all cases.

In order to achieve the rate-distortion bound for processes some delay was necessary. A brief discussion of a one-pole process example showed that removing the delay caused some degradation in performance, but not nearly as much as removing the feedback channel; hence, useful application of feedback does not require delay for processing, but delay does improve the system performance. Kalman filtering is easily adaptable to noiseless feedback.

The overall improvement of analog estimation with noiseless feedback is not nearly as dramatic (when viewed as a signal-to-noise ratio improvement) as is the digital system improvement. This fact is a result of the rate-distortion bound which specifies the ultimate improvement attainable by any system; the feedback system achieves this ultimate and can do no better. The potential advantage of using feedback is much less in analog systems; in fact there exist nofeedback systems (see Van Trees [29]) which operate very close to the rate-distortion bound. The only possible advantage of analog feedback systems in these cases is system simplicity.

In Chapter 4 the digital problem with noisy feedback is studied. Unfortunately analytic solution of noisy feedback is unavailable in general. Elias [3] has solved the simplest discrete-time noisy feedback system, but the more complex extensions of his discrete-time feedback system are not easily solved.

Similarly attempts to solve the noisy feedback problem in continuous time have failed. Instead of the performance of many suboptimum MSE noisy feedback systems have been evaluated. These systems are similar to the noiseless feedback system of Chapter 2 and are asymptotically optimal as the feedback noise density tends to zero. Based on Elias' solution, it appears that the suboptimal MSE system performance for $SNR_B > SNR_F$ is not significantly different from the unknown optimal performance. Approximation techniques are also demonstrated which are useful in estimating how well a feedback system can be expected to perform without actually performing the system calculations.

Without feedback linear modulation systems cannot achieve channel capacity as the length of the transmission interval increases. With noiseless feedback the linear system can achieve channel capacity with an increasing time interval. With noisy feedback the linear system cannot achieve channel capacity with increasing time interval given a fixed average transmitted power in the feedback link.

An extension of a result of Elias [3] implies that the effective signal-to-noise ratio of a linear noisy feedback digital system is bounded by the sum of the forward and reverse signal-to-noise ratios, $SNR_F + SNR_B$. The actual performance (effective signal-to-noise ratio) calculated for several suboptimal MSE systems is much less than the bound. Even for the exact optimal solution of Elias the optimal system is generally much below this bound, implying that the bound is not too tight. Consider the feedback situation when both channels are white noise channels with the same signal-to-noise ratios, $SNR_F = SNR_B$. The bound implies that the effective signal-to-noise ratio is at best doubled. The conclusion is that, for feedback to be effective in offering improvement, the feedback channel must be considerably better than the forward channel. In those cases where the feedback channel is not significantly better than the forward channel, perhaps some signalling scheme other than linear amplitude modulation will utilize the available feedback channel better. Whether or not a more complex modulation system could actually utilize the poor feedback channel is yet to be determined; perhaps no system can achieve much improvement whenever the reverse channel is no better than the forward channel.

The text includes many examples indicating that the optimal noisy feedback system does not follow from the noiseless feedback system by replacing the receiver state in the noiseless system by the estimate of the receiver state to get the optimal noisy feedback system. Such MSE systems are only asymptolically optimal. There is no way to interpret the optimal system of Elias so as to visualize it as estimating the receiver state. In other words it appears that the optimal feedback system cannot be determined except by a direct solution (which seems very difficult).

5.2 Related Topics in Feedback Systems

In this thesis only linear feedback schemes have been considered.

The main reason for this restriction is the simplicity of linear systems. Others have considered more complex signalling (e.g., orthogonal message sets) which perform better than the linear system both with and without feedback. In some applications the performance requirements may necessitate considering nonlinear feedback systems.

Just as more complex systems are possible, more complex channels are also of interest. Additive colored noise in the forward channel is one example. A multiplicative noise or fading channel is another example of a forward channel distúrbance which could be combatted by using a feedback channel. Fading severely inhibits one-way communication systems; a good feedback channel perhaps would offer a means of system simplification and improvement. If the feedback link is essentially no better than the forward channel, it is not clear if feedback can offer much in the way of improved performance.

Two-way communication over two channels between two locations is a feedback system problem. Both channels could be operated independently of one another or used in conjunction with each other to form feedback systems. The results of the preceding chapters are directly applicable with very little additional computation.

Similarly feedback applications to the various analog estimation systems are also possible in the presence of other types of channels.

5.3 Suggestions for Future Research

The results of Chapters 2 and 3 essentially complete the study of noiseless feedback systems over additive white noise channels. Conceivably additional colored noise could be added to the forward channel along with a bandwidth constraint to prevent operation at high frequencies where only the white noise is present, but this problem does not seem to be the most useful. The most desirable result would be an analytic solution for the optimal noisy feedback system for additive white noise in the feedback channel. Such a result for digital and/or analog systems would be an appropriate conclusion to the approximation techniques of Chapter 4. This author along with others has attempted to solve this problem without success; whether or not a solution is possible remains to be seen.

In addition to continued effort to solve the general noisy feedback problem, attention should be focused on fading channels. It appears that feedback could be most useful in lessening the effects of fading. Suppose the noiseless feedback system of Chapter 2 actually operated in a fading channel where the transmitted message waveform was scaled by a random variable. Since the feedback system transmits the error, the fading only affects the time it takes the receiver to approach the correct message point, not the ultimate message point. In effect the only change the slow fading of the channel has on the noiseless feedback system is to make the effective transmitted energy a random variable. The overall performance in a fading environment is reduced, but not enough to make the feedback system ineffective. For fast fading channels the effects are not as obvious and need to be studied. Certainly other than linear modulation systems are needed for effective communication in a fading environment.

Theoretically fading does not alter the white noise channel capacity. Yet there are no known nofeedback systems capable of achieving this capacity; conceivably a noiseless feedback system could be designed which would achieve this capacity and thereby produce a system capable of achieving the theoretical capacity.

CHAPTER 6

Angle Modulation System Performance Relative to the Rate-Distortion Bound

The results of this chapter are essentially unrelated to the preceding chapters. The channel capacity of an rms bandlimited white noise channel is calculated. Then, the results of Van Trees [29] pertaining to angle modulation subject to mean square bandwidth constraints are compared to the ultimate performance implied by the rate-distortion bound.

6.1 rms Bandlimited Channel Capacity

The usual definition of the channel capacity of a band-limited additive noise channel implies that the channel is strictly bandlimited. In some applications a strictly band-limited assumption cannot be realistically imposed on the transmitted signal and/or channel. For example, a transmitted signal of finite duration is obviously not strictly bandlimited. To compare the performance of such an approximately bandlimited system to the theoretical performance implied by the strictly bandlimited channel capacity can lead to contradictions (such as system performance better than the "theoretical" ultimate performance). In this section the strictly bandlimited assumption of channel capacity is replaced by a mean-square bandwidth (rms) constraint and the resulting channel capacity computed.

As is well known the channel capacity of an additive white noise channel (spectral density $N_0/2$) for which the transmitter spectrum is S(f) is

$$C = \frac{1}{2} \int_{-\infty}^{\infty} df \, \ln \left(1 + \frac{2S(f)}{N_0}\right). \qquad (6.1)$$

It is convenient to define a normalized spectrum, $\sigma(f)$

$$S(f) = P \, \sigma(f) \qquad (6.2)$$

where P is the average transmitted power which is assumed finite. Equation 6.1 becomes

$$C = \frac{1}{2} \int_{-\infty}^{\infty} df \, \ln \left(1 + \frac{2P}{N_0} \sigma(f)\right) \quad \text{nats/sec.} \qquad (6.3)$$

The remaining part of the solution for C is to maximize Equation 6.3 subject to any transmitter or channel constraints. Here an infinite bandwidth channel is assumed with power and bandwidth constraints on the transmitter.

For example, if a strictly bandlimited constraint is made at the transmitter

$$\sigma(f) = 0 \qquad |f| > W \qquad (6.4)$$

and the optimal choice of $\sigma(f)$ is

$$\sigma(f) = \frac{1}{2W} \qquad |f| \leq W \qquad (6.5)$$

with the resulting well known capacity formula from Equation 6.3

$$C = W \ln \left(1 + \frac{P}{N_0 W}\right). \tag{6.6}$$

Defining a signal-to-noise ratio Λ in the transmitter bandwidth

$$\Lambda = \frac{P}{N_0 W} \tag{6.7}$$

implies that the channel capacity increases logarithmically with increasing signal-to-noise ratio.

For an rms bandwidth B constraint at the transmitter,

$$B^2 = \int_{-\infty}^{\infty} df\, f^2\, \sigma(f) \tag{6.8}$$

which represents a constraint on $\sigma(f)$. The other implies constraints are

$$\sigma(f) \geq 0 \tag{6.9}$$

$$\int_{-\infty}^{\infty} df\, \sigma(f) = 1. \tag{6.10}$$

In order to maximize Equation 6.3 subject to the three constraints on $\sigma(f)$ (Equations 6.8-6.10), define

$$J = \frac{1}{2} \int_{-\infty}^{\infty} df\ \ell n\ (1 + \frac{2P}{N_0} \sigma(f)) + \alpha \int_{-\infty}^{\infty} \sigma(f)\ df + \gamma \int_{-\infty}^{\infty} f^2\ \sigma(f)\ df \quad (6.11)$$

where α and γ are Lagrange multipliers. Perturbation of J with respect to $\sigma(f)$ yields

$$\sigma(f) = \max(0, -\frac{1}{2} [\frac{1}{\alpha+\gamma f^2} + \frac{N_0}{P}]). \quad (6.12)$$

The maximum operation is necessary to satisfy $\sigma(f) \geq 0$. Clearly if α, γ are positive, $\sigma(f) = 0$ which does not satisfy the constraint Equations 6.8 and 6.10. Similarly if the two multipliers are of different signs, the constraints cannot be satisfied; hence, α and γ are both negative. Define two new positive multipliers Q and f_c such that

$$\sigma(f) = \max(0, \frac{1}{2B\Lambda} (\frac{Q^2 - \left(\frac{Qf}{f_c}\right)^2}{1 + \left(\frac{Qf}{f_c}\right)^2})) \quad (6.13)$$

where the signal-to-noise ratio in the rms bandwidth

$$\Lambda = \frac{P}{N_0 B} \quad (6.14)$$

has been introduced. The transmitter spectrum is that of a one-pole process shifted down to cutoff at $f = f_c$.

For $\sigma(f)$ as given in Equation 6.13, direct evaluation of the constraints Equations 6.8 and 6.10 yield

$$\Lambda B^3 = f_c^{\,3} \left(\frac{2}{3} + \frac{1}{Q^2} - \left(\frac{1}{Q} + \frac{1}{Q^3}\right) \tan^{-1} Q\right) \qquad (6.15)$$

and

$$\Lambda B = f_c \left\{\left(\frac{1}{Q} + Q\right) \tan^{-1} Q - 1\right\}. \qquad (6.16)$$

These two equations determine the unknowns f_c and Q. Given f_c and Q as the solution of these equations, the channel capacity from Equation 6.3 is

$$C = 2 f_c \left\{1 - \frac{1}{Q} \tan^{-1} Q\right\}. \qquad (6.17)$$

It can be shown from the above equations that C can also be written

$$C = B\, g(\Lambda) \qquad (6.18)$$

where $g(\Lambda)$ is a complicated implicit function. The important observation is that channel capacity for rms bandwidth is of the same functional form as the strictly bandlimited form Equation

6.6 providing signal-to-noise ratios in the transmitter bandwidth are defined. Unfortunately $g(\Lambda)$ is implicit and cannot be determined analytically.

The equations can be solved approximately for $\Lambda >> 1$. For large Q the channel capacity in Equation 6.17 is

$$C \approx 2 f_c . \tag{6.19}$$

Similarly for large Q, Equation 6.15 implies

$$\frac{2}{3} f_c^3 \approx \Lambda B^3 \tag{6.20}$$

or combining

$$C \approx B (12)^{1/3} \Lambda^{1/3} \qquad (\Lambda >> 1) \tag{6.21}$$

which implies that channel capacity increases as the cube root of Λ for an rms constraint, but only logarithmically for a strict bandwidth constraint. Thus, using the strict bandwidth capacity formula for channels which are actually rms band-limited yields a capacity much lower than the true capacity.

$g(\Lambda)$ is plotted in Figure 6-1 along with its asymptote (Equation 6.21).

6.2 Rate-Distortion Bound for rms Bandlimited Channels

Goblick [16] states the rate-distortion function for a stationary Gaussian process a(t) with a monotonic spectrum $S_a(f)$ in parametric

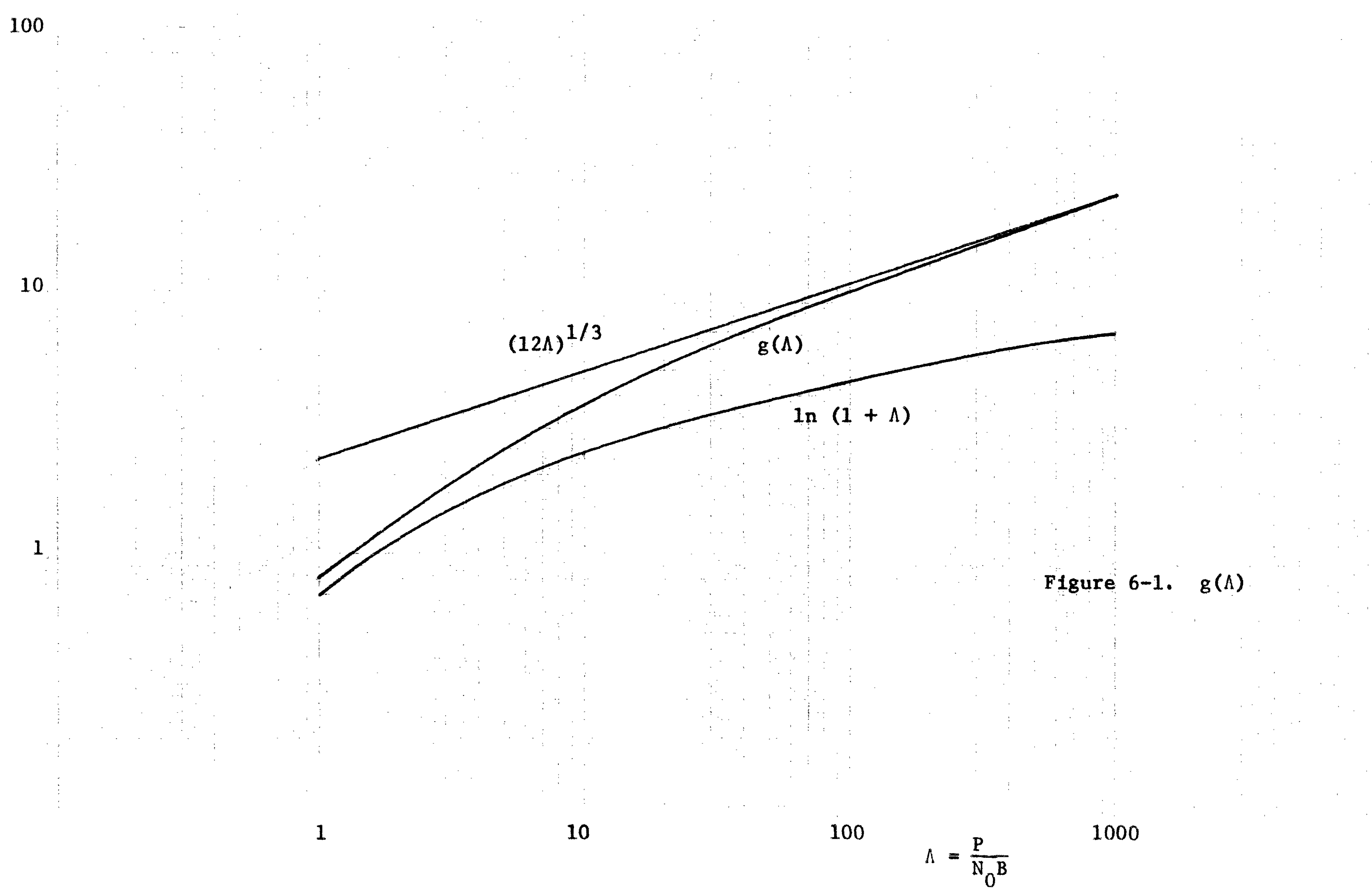


Figure 6-1. $g(\Lambda)$

form as

$$R = \int_0^{\phi} df \quad \ln \left[\frac{S_a(f)}{S_a(\phi)}\right] \qquad \text{(nats/sec)} \tag{6.22}$$

$$\xi = 2\phi\, S_a(\phi) + 2 \int_{\phi}^{\infty} df\; S_a(f) \tag{6.23}$$

where ξ is the minimum mean square error in estimating a(t) for a given information rate R. In this chapter only message spectra of the Butterworth family will be treated. The n^{-th} order unit power Butterworth spectrum is

$$S_a(f) = \frac{n}{\pi W_M} \quad \frac{\sin(\frac{\pi}{2n})}{1 + (\frac{f}{W_M})^{2n}} \tag{6.24}$$

which is monotonic. Other nonmonotonic spectra could be treated with modification of Equations 6.22 and 6.23. Goblick [16] has plotted the rate-distortion function for several different Butterworth orders.

To determine the rate-distortion bound for transmitting a(t) through a white noise channel with an rms transmitter bandwidth constraint, the rms channel capacity is equated to R in Equation 6.21 and the resulting ξ minimum mean square estimation error ξ can be determined for the particular message spectrum $S_a(f)$.

The rms channel capacity given in Equation 6.18 is for a low pass channel. Since angle modulation systems operate at bandpass, the correct rms capacity is the bandpass channel capacity. The important bandwidth for bandpass systems is the rms bandwidth about the carrier. Denoting this bandwidth by B and assuming that the signal-to-noise ratio Λ is defined in this rms bandwidth B, then the transmitted signal spectrum has half its power at positive frequencies and half at the negative frequencies. For the positive frequencies only the signal-to-noise ratio is $\Lambda/2$ and the channel capacity contribution for the positive frequencies only is $Bg(\Lambda/2)$. The negative frequencies contribute the same, giving a total bandpass capacity of

$$C_{\substack{\text{rms} \\ \text{bandpass}}} = 2\,B\,g(\Lambda/2) \qquad \text{(nats/sec)} \qquad (6.25)$$

Given a particular Λ and B (signal-to-noise ratio and rms bandwidth), Equation 6.25 indicates the channel capacity which is then inserted into the left side of Equation 6.22. For the message spectrum $S_a(f)$ the value of ϕ can be determined and used in Equation 6.23 to determine the minimum mean square error ξ which is the rate distortion bound. No system can produce lower mean square error operating under the same constraints. In general the solution must be done numerically.

6.3 Comparison of Angle Modulation Systems for Butterworth Message Spectra

The choice of an rms bandwidth constraint for angle modulation systems is quite judicious since the rms bandwidth of the transmitted signal is easily determined even though the modulated spectrum is rather complex. Van Trees [29] has determined the system performance (reciprocal normalized mean square estimation error) for several types of angle modulation systems assuming a Butterworth message spectrum.

In Figure 6-2 the performance of various modulation systems is compared for a first order Butterworth message spectrum. A bandwidth expansion of $B/W_M = 10$ is used. The horizontal axis is labelled as the signal-to-noise ratio in the message bandwidth. The lowest curve corresponds to realizable FM, the next unrealizable FM, the next optimum (preemphasis) angle modulating and finally the rate-distortion bound. The breaks in the three curves correspond to the approximate threshold region of the angle modulation systems. For the first order Butterworth message the optimum angle modulation system is only 6 db worse than the ultimate rate-distortion performance as the signal-to-noise ratio increases.

The numerical performance comparisons could be carried out for other order spectra in this same fashion. Some analytical results can be obtained for large signal-to-noise ratios for arbitrary order Butterworth spectra. In fact for all Butterworth orders (other than the first) the optimum angle modulation performance

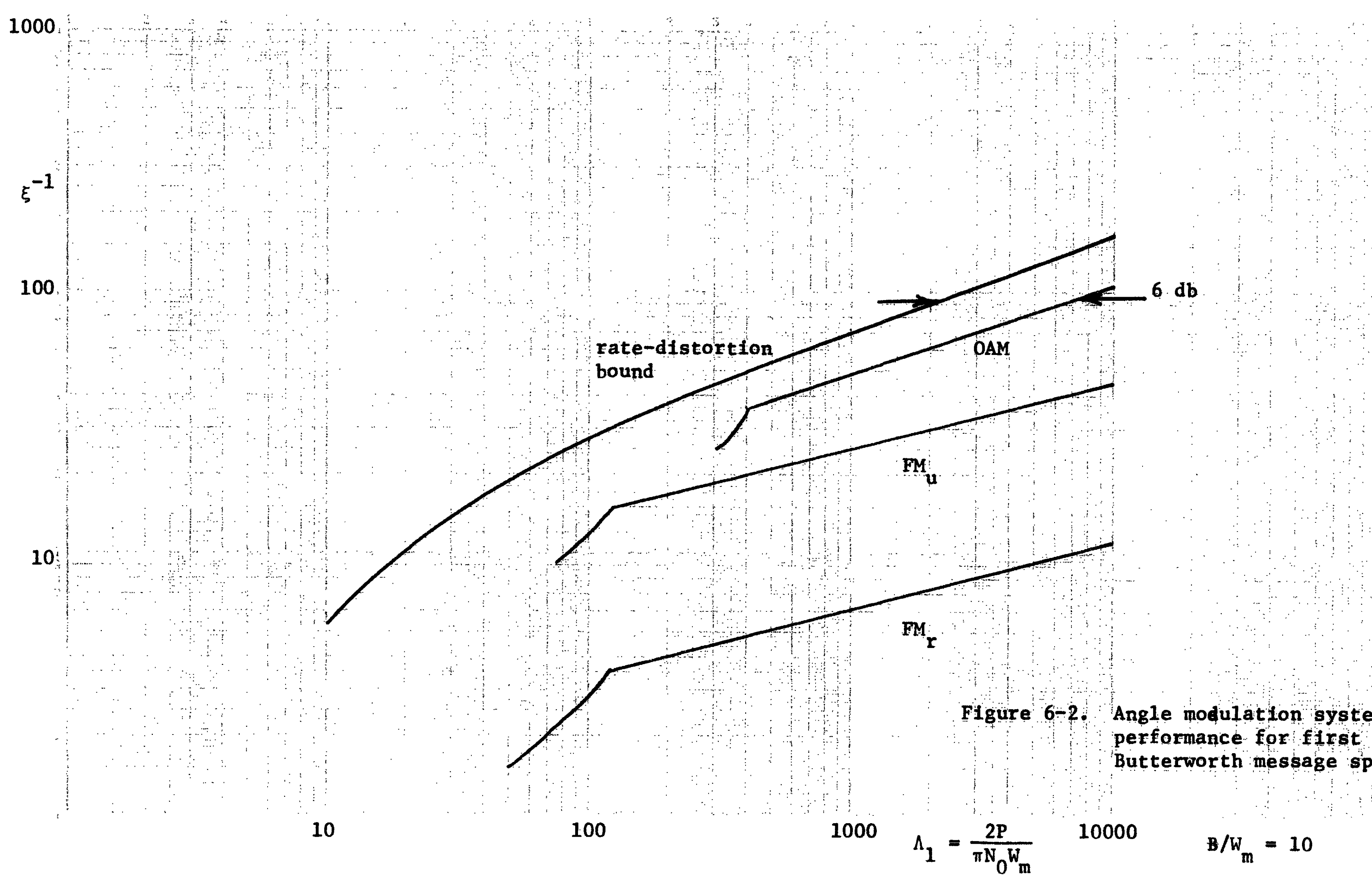


Figure 6-2. Angle modulation system performance for first order Butterworth message spectrum

diverges from the rate distortion bound rather than paralleling it as shown in Figure 6-2.

For large Λ

$$C_{\substack{rms\\bandpass}} \sim \Lambda^{1/3} \tag{6.26}$$

Using this along with the Butterworth spectrum in Equation 6.22 and 6.23 implies that the maximum reciprocal error is

$$\xi^{-1}_{\substack{rate\\distortion}} \sim \Lambda^{\frac{2n-1}{3}} \tag{6.27}$$

where n is the order of the Butterworth spectrum. Using the integral expressions for the angle modulation performance obtained in Van Trees [29] and approximating them for large Λ, it follows that

$$\xi^{-1}_{FM_u} \sim \xi^{-1}_{FM_r} \sim \Lambda^{\frac{2n-1}{2n+2}} \tag{6.28}$$

Of course, $\xi^{-1}_{FM_u} > \xi^{-1}_{FM_r}$ always although they have the same asymptotic slope. Similarly it can be shown that

$$\xi^{-1}_{OAM} \sim \Lambda \qquad (n \geq 3) \tag{6.29}$$

which increases slightly faster than FM, but not nearly as fast as ξ_{RD}^{-1} . As shown in Figure 6-2, for n = 1 $\xi_{RD}^{-1} \sim \xi_{0AM}^{-1} \sim \Lambda^{1/3}$ and the optimum angle modulation system performance does not diverge from the rate-distortion bound.

As the order n tends to infinity, the Butterworth spectrum becomes strictly bandlimited. For this situation

$$\xi_{\substack{rate\\distortion}}^{-1} \sim e^{\Lambda^{1/3}} \tag{6.30}$$

$$\xi_{FM_u}^{-1} \sim \xi_{0AM}^{-1} \sim \Lambda \tag{6.31}$$

$$\xi_{FM_r} \sim \frac{\Lambda}{(\ell n\ \Lambda)^3} \tag{6.32}$$

Thus, for strictly bandlimited message spectra angle modulation system performance diverges rapidly from the rate-distortion bound.

REFERENCES

[1] P. E. Green, Jr., "Feedback Communications Systems," ch. 14, in Lectures in Communication System Theory, Baghdady, Ed., McGraw-Hill, New York, 1961

[2] P. Elias, "Channel capacity without coding," appendix to ch. 14, Lectures in Communications System Theory, Baghdady, Ed., McGraw-Hill, New York, 1961

[3] P. Elias, "Networks of Gaussian channels with applications to feedback systems," IEEE Trans. on Info. Theory, July 1967

[4] J. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback -- Part I: No bandwidth constraint," IEEE Trans. on Info. Theory, April 1966

[5] C. E. Shannon, "The zero error capacity of a noisy channel," IRE Trans. on Info. Theory, Sept. 1956

[6] J. Schalkwijk, "A coding scheme for additive noise channels with feedback -- Part II: Bandlimited signals," IEEE Trans. on Info. Theory, April 1966

[7] J. K. Omura, "Signal optimization for channels with feedback," Tech. Rep. No. 7001-1, Stanford Electronics Laboratories, August 1966

[8] J. K. Omura, "Signal optimization for additive noise channels with feedback," Session 7/4, WESCON, August 1966

[9] G. L. Turin, "Signal design for sequential detection systems with feedback," IEEE Trans. on Info. Theory, July 1965

[10] G. L. Turin, "Comparison of Sequential and nonsequential detection systems with uncertainty feedback," IEEE Trans. on Info. Theory, Jan. 1966

[11] M. Horstein, "On the design of signals for sequential and nonsequential detection systems with feedback," IEEE Trans. on Info. Theory, Oct. 1966

[12] H. J. Kushner, "Optimal stochastic control," IRE Trans. on Auto. Control, Oct. 1962

[13] H. J. Kushner, "Near optimal control in the presence of small stochastic perturbations," Journal of Basic Eng., ASME, March 1965

[14] M. Athans and P. L. Falb, Optimal Control, McGraw-Hill, New York, 1966

[15] T. J. Cruise, "Achievement of rate-distortion bound over additive white noise channel utilizing a noiseless feedback channel," IEEE Proceedings (Letters), April 1967

[16] T. J. Goblick, "Theoretical limitations on the transmission of data from analog sources," IEEE Trans. on Info. Theory, Oct. 1965

[17] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," Journal of Basic Eng., ASME, March 1961

[18] K. Ito, "On stochastic differential equations," Memoirs of Am. Math. Soc., No. 4, 1951

[19] R. L. Stratonovich, "A new form of representation of stochastic integrals and equations," Tech. Rep. No. 7050-9, Stanford Electronics Laboratories, August 1966

[20] J. M. Wozencraft and I. M. Jacobs, Principles of Communication Engineering, John Wiley and Sons, 1965

[21] R. L. Kashyap, "A feedback coding scheme for an additive noise channel with a noisy feedback link," First Annual Princeton Conference on Information Sciences and Systems, 1967

[22] A. J. Kramer, "Analysis of communication schemes using an intermittent feedback link," TR 7050-11, Systems Theory Lab., Stanford Univ., March 1967

[23] S. Butman, "Optimum linear coding for additive noise systems using information feedback," TR No. 1, Comm. Theory Lab., Calif. Inst. of Tech., May 1967

[24] S. Butman and T. L. Grettenberg, "Optimum linear coding for additive noise systems using feedback," Paper 25.5, 1967 IEEE Conf. on Communications

[25] L. Pontryagin et al, The Mathematical Theory of Optimal Processes, Interscience Publishers, Inc., New York, 1962

[26] A. D. Wyner, "On the Schalkwijk-Kailath coding scheme with a peak energy constraint," IEEE Trans. on Info. Theory, Jan. 1968

[27] T. J. Cruise, "Scheme to utilize noiseless feedback to reduce channel noise," RLE QPR No. 88, Mass. Inst. of Tech., Jan. 1968

[28] W. B. Davenport and W. L. Root, Random Signals and Noise, McGraw-Hill, New York, 1958

[29] H. L. Van Trees, Detection, Estimation, and Modulation Theory Part II, John Wiley and Sons, (to be published)

[30] H. L. Van Trees, Detection, Estimation, and Modulation Theory Part I, John Wiley and Sons, New York, 1968

[31] J. K. Omura, "Optimum linear transmission of analog data for channels with feedback," IEEE Trans. on Info. Theory, Jan. 1968

[32] T. J. Cruise, "Channel capacity for an rms bandwidth constraint," RLE QPR No. 90, Mass. Inst. of Tech., July 1968

BIOGRAPHY

Theodore Joseph Cruise was born September 9, 1942 in San Francisco. He graduated from Willow Glen Senior High School in San Jose, California in 1960.

He received the S.B. degree in electrical engineering from Massachusetts Institute of Technology in 1965 and the S.M. and E.E. degrees from M.I.T. in 1966. He was a National Science Foundation Fellow for four years and also taught courses in the Department of Electrical Engineering. He is a member of Tau Beta Pi and Eta Kappa Nu.

He married the former Sherry Ormsby in 1967.