3.

NONLINEAR LEAST-SQUARES FILTERING AND FREQUENCY MODULATION

ARTHUR DICKSON HAUSE

TECHNICAL REPORT 371

AUGUST 25, 1960

LOAN COPY ONLY

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
RESEARCH LABORATORY OF ELECTRONICS
CAMBRIDGE, MASSACHUSETTS

The Research Laboratory of Electronics is an interdepartmental laboratory of the Department of Electrical Engineering and the Department of Physics.

The research reported in this document was made possible in part by support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, jointly by the U. S. Army (Signal Corps), the U. S. Navy (Office of Naval Research), and the U.S. Air Force (Office of Scientific Research, Air Research and Development Command), under Signal Corps Contract DA36-039-sc-78108, Department of the Army Task 3-99-20-001 and Project 3-99-00-000.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

RESEARCH LABORATORY OF ELECTRONICS

Technical Report 371 August 25, 1960

NONLINEAR LEAST-SQUARES FILTERING AND FREQUENCY MODULATION

Arthur Dickson Hause

Submitted to the Department of Electrical Engineering, M. I. T., August 22, 1960, in partial fulfillment of the requirements for the degree of Doctor of Science.

Abstract

This thesis concerns the use of optimum nonlinear filtering for the recovery of messages from modulated signals in the presence of additive Gaussian noise. Signal-to-noise ratio, defined in a manner especially suited to nonlinear filtering, is taken as the measure of performance. This definition also provides a useful connection with the mean-square-error criterion.

Each filter has a linear section and a nonlinear, memoryless section. Attention is focused primarily on the latter. A simplified review of important properties of the orthogonal polynomials employed by Wiener for representation of the nonlinear filter section is included; special emphasis is placed on the properties of Hermite polynomials because this is a cornerstone in the derivation of succeeding results. The low signal-to-noise-ratio performance of communication systems is shown to be related in a simple way to the structure of the optimum filter polynomial.

The use of this nonlinear filter theory confirms the notion that conventional frequency-modulation practice produces somewhat less than optimum reception at low signal-to-noise ratios. This study indicates that the type of frequency modulation that has potentially the best performance in this range is narrow-band FM with a phase-synchronized carrier. Also, the proper use of amplitude information in the noisy FM signal improves performance moderately.

TABLE OF CONTENTS

I. INTRODUCTION

1.1 PURPOSE OF THIS RESEARCH

This report contains a study of nonlinear least mean-square-error filters and their signal-to-noise-ratio performance in communication systems that are subject to additive Gaussian noise. Special emphasis is placed on frequency-modulation (FM) systems, particularly at low signal-to-noise ratios.

The theory of optimum linear least-squares filtering developed by Wiener, Lee, and others (1), is a standard analytical tool in modern statistical communication theory. However, the nonlinear filter theory of Wiener (2) and others (3) has thus far found little application in communication research, partly because of its relative newness, but even more because of its complexity. The aim of the present study is to simplify concepts and notation to the point where the nonlinear theory reveals a number of simple, interesting facts about the optimum performance of noisy communication systems. Expressions for corresponding optimum filters are found to be particularly simple at low signal-to-noise ratios.

This investigation is oriented toward continuous systems and signals. Discrete cases are discussed either for comparative purposes or as approximations to continuous cases.

Special interest in FM systems derives from the fact that most analytical techniques have proved clumsy, if not inadequate, in the treatment of frequency modulation at low signal-to-noise ratios, the region in which the Wiener theory finds its simplest application.

1.2 COMMUNICATION SYSTEM MODEL

Consider the simplified communication system model shown in Fig. 1. The nonlinear operator is an abstraction of a transmitter. The nonlinear filter represents a receiver. Message and Gaussian noise sources are assumed to operate independently.

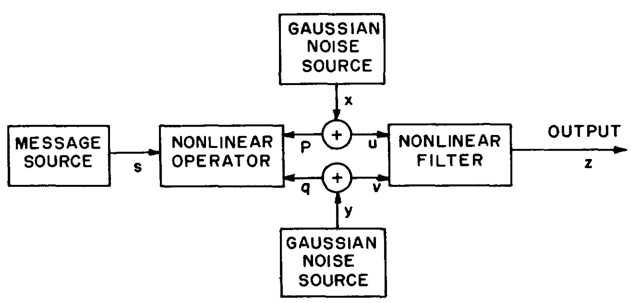

Fig. 1. Communication system model.

The message, s, noise components, x and y, and filter output, z, are assumed to have zero mean values ($\bar{s}=0$, $\bar{x}=0$, $\bar{y}=0$, $\bar{z}=0$, where the bar stands for "average value of"). Regarding Fig. 1, the purpose of this report is to discuss optimization of the nonlinear filter and the resulting signal-to-noise-ratio performance of the system, when we are given the message and noise statistics and a description of the nonlinear operator.

It is important to be aware of several subtle or implicit features of the model of Fig. 1. One feature is suppression of the time variable. Mathematically, the various signals are treated as random variables or random functions. A random variable represents a signal at a single instant of time. We define a random function as the representation of a signal over a continuous range of time values. (Thus, by this definition, a random function is a generalization of a random vector.) For example, the noisy signals u and v in Fig. 1 are treated as random functions, since the nonlinear filter generally has memory. Averages given in this report are usually statistical. In physical situations we must deal with time averages. The two averages are linked by the ergodic hypothesis (4). The Gaussian noise sources are assumed to be ergodic. If the message source is ergodic and the nonlinear operator is time-invariant, all signals in the system are ergodic. If message statistics or the nonlinear operator vary periodically, one possible procedure is to sample at the periodic rate and ascribe the ergodic property to the samples. In any case, the important point is that we convert problems involving time signals into their statistical equivalents, whenever it is possible. Another potential source of confusion is the statement that z is an estimate (or reconstruction) of s. In a physical problem, we would say that the signal z(t) estimates the signal $s(t-\tau)$ for some specified delay, τ (prediction, if τ is negative). We therefore assume that τ is specified implicitly (if not explicitly) in any particular situation, and then omit it, as well as t, from the notation.

Another feature of the Fig. 1 model is the lack of an explicit carrier. In a double-sideband carrier communication system, with sinusoidal carrier

$$\cos(2\pi f_o t+\theta_o) \tag{1}$$

a general expression for the transmitted signal which explicitly includes the time variable is

$$p(t)\cos(2\pi f_o t+\theta_o) + q(t)\sin(2\pi f_o t+\theta_o) \tag{2}$$

where p(t) is the amplitude of the in-phase component and q(t) is the amplitude of the quadrature component. Omission of the carrier is a convenience which, like omission of the time variable, does not affect any essential elements of the problems considered herein. Physically, carrier removal can be accomplished by ideal product demodulation. In phase-synchronous systems, θ_o is a known constant, and demodulation yields the p and q signals depicted. If the system is not phase-synchronous, θ_o must be replaced by $\theta_e(t)$, an appropriate phase-error signal. Then the demodulated signal corresponding to p in Fig. 1 is given by

$$p(t) \cos \theta_e(t) + q(t) \sin \theta_e(t) \tag{3}$$

and that corresponding to q is given by

$$q(t) \cos \theta_e(t) - p(t) \sin \theta_e(t) \tag{4}$$

It will frequently be simpler to redefine p and q to represent expressions 3 and 4 directly, rather than to carry along all the extra notation or define additional variables.

The only important effect of ideal product demodulation on Gaussian noise is to shift its frequency spectrum to the base band. Under the assumption that the original noise band is symmetrical about the carrier frequency, the in-phase and quadrature base-band noise components, x and y in Fig. 1, are independent and have power spectra similar to that of the original noise, but centered about zero frequency instead of the carrier (5).

A base-band system is represented in Fig. 1 by eliminating quadrature signals, so that $q \equiv 0$ and $y \equiv 0$ (hence $v \equiv 0$). Other systems, such as single-sideband systems, may require notational changes for accurate representation, but they can generally be put into a form that is essentially the same as that outlined in Fig. 1.

1.3 FREQUENCY-MODULATION MODEL

Referring again to Fig. 1, the FM signal model is defined by

$$\left.\begin{aligned} p &= \cos \theta \\ q &= \sin \theta \end{aligned}\right\} \tag{5}$$

where the phase, θ, is a zero-mean ($\bar{\theta} = 0$) operator, controlled by the message, s, and expressed in functional notation as

$$\theta = \theta[s] \tag{6}$$

Signals p and q in Eq. 5 may be thought of as derived from $\cos(\theta + 2\pi f_o t + \theta_o)$ by ideal product demodulation. For FM systems that are not phase-synchronous the introduction of a phase error, θ_e, as in Eqs. 3 and 4, leads to the alternative definition

$$\left.\begin{aligned} p &= \cos(\theta - \theta_e) \\ q &= \sin(\theta - \theta_e) \end{aligned}\right\} \tag{7}$$

It should again be stressed that removal of the carrier is a simplification involving no theoretical limitation on the usefulness of the results. However, this does not imply that the optimum physical receiver must contain a product demodulator.

1.4 WIENER FILTER REPRESENTATION

Figure 2 diagrams the Wiener description of the nonlinear filter in the Fig. 1 system. Linear networks with impulse responses $h_1, \ldots, h_L$ provide a finite repre-

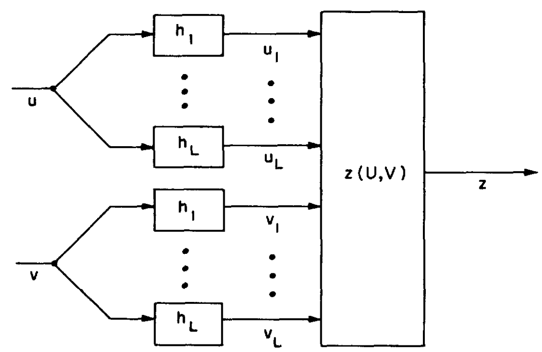

Fig. 2. Nonlinear filter representation.

sentation of the u, v signal components at the filter input. Network outputs may be expressed in the conventional manner,

$$\left.\begin{aligned} u_j(t) &= \int_{-\infty}^{\infty} h_j(\tau)\, u(t-\tau)\, d\tau \\ v_k(t) &= \int_{-\infty}^{\infty} h_k(\tau)\, v(t-\tau)\, d\tau \end{aligned}\right\} \tag{8}$$

These networks are orthonormal in the sense that

$$\int_{-\infty}^{\infty} h_j(t)\, h_k(t)\, dt = \delta_{jk} \tag{9}$$

For example, Laguerre networks (6) and tapped delay lines would be suitable physical networks. If the linear filtering is accomplished before demodulation in a physical carrier system, only a single set of filters is required. This is the main reason for indicating in Fig. 2 identical sets of filters for the separate u and v components. The subscript notation for network outputs is suitable because linearity permits resolution of additive components. That is, $u = p + x$ implies

$$u_j = p_j + x_j \tag{10}$$

and similarly for v.

The nonlinear, memoryless part of the filter (Fig. 2) is shown as a function, $z(U,V)$, where

$$z(U, V) = z(u_1, \ldots, u_L, v_1, \ldots, v_L) \tag{11}$$

That is, the symbols U and V represent the sets of variables

$$
\begin{aligned}
U &= \{u_1, \ldots, u_L\} \\
V &= \{v_1, \ldots, v_L\}
\end{aligned}
\tag{12}
$$

In the Wiener representation, z(U, V) is a polynomial. Because polynomials can approximate continuous functions (7), their use entails no physical restriction on generality of the results.

The importance of the Wiener filter representation should not be underestimated. Representation of filters by abstract nonlinear functionals may be fine theoretically, but it gives little intuitional insight into problems of practical interest. Even the time-amplitude representation suggested by Bose (6) is less than ideal in this respect, considering its difficulty in classifying types of nonlinearity (in particular, in distinguishing between linear and strictly nonlinear characteristics of a filter). An important problem that will receive scant treatment in this report is that of choosing the networks in the linear section of the Wiener filter so that they provide an adequate representation of the filter input. (This could also be thought of as approximating the infinite-dimensional vectors u and v by the finite-dimensional vectors U and V over their ranges of variation, in some average sense.) In order to focus attention on the nonlinear, memoryless part of the filter, standard procedure will be to assume that linear networks have been satisfactorily chosen for any given problem and are fixed, while z(U, V) and desired system parameters are varied.

II. SIGNAL-TO-NOISE-RATIO PERFORMANCE

2.1 PRELIMINARY REMARKS

Signal-to-noise ratio has been chosen as the performance criterion, in preference to mean-square error, because the former is gain-invariant and has greater intuitive meaning in many communication problems than does the latter. However, these two criteria are closely related, as we shall see. Both find greater application in continuous than in discrete systems; this feature fits in with the scope of this investigation. Much of the following discussion is based on an earlier report (8).

Suppose we have a desired signal, S, and an independent noise, N, combined to give the noisy signal, S + N. If $\bar{S} = 0$ and $\bar{N} = 0$, the conventional definition of signal-to-noise power ratio, γ, for the noisy signal S + N, is

$$\gamma = \frac{\overline{S^2}}{\overline{N^2}} \tag{13}$$

The adjective "power" is dropped, as in conventional practice. The quantity $10 \log_{10} \gamma$, expressed in decibels, is also conventionally called signal-to-noise ratio. The distinction between γ and its logarithmic measure will be evident whenever it is important. The purpose of section 2.2 is to generalize the definition of Eq. 13 to situations in which the desired signal and the noise are not additive, independent components of the noisy signal.

The present discussion applies to systems of the form shown in Fig. 3. We shall assume that the signal, s, and the noisy signal, z, have zero means ($\bar{s} = 0$, $\bar{z} = 0$).

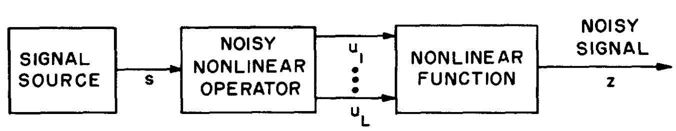

Fig. 3. General noisy system model.

As suggested by the notation, the system shown in Fig. 1 is a special case of the system shown in Fig. 3. The nonlinear operator, Gaussian noise sources, and linear networks (Fig. 2) belonging to the nonlinear filter shown in Fig. 1 constitute one possible form of the noisy nonlinear operator shown in Fig. 3. (If it is desired, the quadrature variables in Fig. 1 may be implicitly included in Fig. 3; it complicates the notation unnecessarily to include $v_1, \ldots, v_L$ explicitly.) The system illustrated in Fig. 3 is used not only for increased generality, but also for the simplification achieved by suppressing the linear part of the Wiener filter, as we have mentioned. The nonlinear function represented by z(U) (see Eqs. 11 and 12) need not be a polynomial for present purposes; any integrable function will do. Time variables are deleted, as before, under the implicit assumption that $z(t)$ estimates $s(t-\tau)$, for some fixed τ.

2.2 GENERALIZED SIGNAL-TO-NOISE-RATIO DEFINITION

It is clear from Fig. 3 that the noisy signal, z, will not generally be the sum of the desired signal, s, and an independent noise. However, it is possible to define a new desired signal, S, by the relation

$$S = cs \tag{14}$$

where c is a constant. Consider the additive noise, N, defined as

$$N = z - S \tag{15}$$

Since $\bar{s} = 0$ and $\bar{z} = 0$, $\bar{S} = 0$ and $\bar{N} = 0$. Therefore, if

$$\overline{SN} = 0 \tag{16}$$

the desired signal and noise components are linearly independent (at least for the fixed time shift, τ). If we ignore the trivial solution $c = 0$, Eqs. 14-16 are satisfied simultaneously if c is given the unique value

$$c = \frac{\overline{sz}}{\overline{s^2}} \tag{17}$$

Hence, it seems reasonable to extend the definition of Eq. 13, with the aid of Eqs. 14 and 15, to obtain

$$\gamma = \frac{c^2\,\overline{s^2}}{\overline{(z-cs)^2}} \tag{18}$$

where c is given by Eq. 17.

It is of incidental interest to observe that the noise power, $\overline{N^2}$, is minimized by this choice for the constant, c. For, the equation

$$\frac{d\overline{N^2}}{dc} = -\,2\,\overline{s(z-cs)} = 0 \tag{19}$$

is satisfied only when c is given by Eq. 17, and

$$\frac{d^2\overline{N^2}}{dc^2} = 2\,\overline{s^2} > 0$$

The form of Eq. 18 is not particularly convenient for a definition (with its dependence on Eq. 17) or for computation. Hence, we introduce the correlation coefficient ρ_{sz},

$$\rho_{sz} = \frac{\overline{sz}}{\left(\overline{s^2}\ \overline{z^2}\right)^{1/2}} \tag{20}$$

which is valid when $\bar{s} = 0$ and $\bar{z} = 0$. Then, Eq. 18 can be rewritten

$$\gamma = \frac{1}{\frac{1}{\rho_{sz}^2} - 1} \tag{21}$$

which is considered to be the signal-to-noise-ratio definition. The gain-invariance of γ is evident because ρ_{sz} is normalized.

If it is desired, Eq. 21 may be expanded in the power series

$$\gamma = \sum_{j=1}^{\infty} \rho_{sz}^{2j} \tag{22}$$

Observe that the series converges unless $\rho_{sz}^2 = 1$, in which case it is still correct because $\gamma = \infty$. This expansion is clearly most useful at low signal-to-noise ratios, when one or only a few terms need be considered.

In the Fig. 1 system, the output signal-to-noise ratio γ_o is defined by Eq. 21. At the input of the nonlinear filter, the input signal-to-noise ratio γ_i is defined as

$$\gamma_i = \frac{\overline{p^2} + \overline{q^2}}{\overline{x^2} + \overline{y^2}} \tag{23}$$

This formula may be justified as follows. From the model of a carrier signal given in Eq. 2, signal power (mean-square value) is found to be

$$\frac{1}{2}\left(\overline{p^2}+\overline{q^2}\right)$$

For example, the FM signal given in Eqs. 5 and 7 has power 1/2. The noise power at carrier frequencies is the same as the power in either base-band noise component, $\overline{x^2} = \overline{y^2}$. If we make use of the definition given in Eq. 13, Eq. 23 follows. Nothing has been assumed about the mean values of the signals p and q. If one of these means does not vanish, it indicates that the respective carrier component is not suppressed by the modulator.

Observe that Eq. 23 is also valid for a base-band system ($q \equiv 0$, $y \equiv 0$) in which $\bar{p} = 0$. Appropriate signal-to-noise ratios can be worked out for single-sideband or other systems. However, Eq. 23 may no longer be correct, unless the notation is chosen specifically to provide this convenience.

2.3 APPLICATION TO LEAST-SQUARES FILTERING

The least-squares nonlinear function illustrated in Fig. 3, chosen to minimize $\overline{(z-s)^2}$, and distinguished from other functions by the subscript ∞, is the conditional mean (9):

$$z_\infty(U) = \int_{-\infty}^{\infty} s\, p(s|U)\, ds \tag{24}$$

where p is the conventional symbol for probability density function. Recall that

$$p(s|U)\, p(U) = p(s, U)$$

We are now equipped to compute the average $\overline{sz}$ ($z(U)$ is an arbitrary integrable function, and $dU = du_1 \dots du_L$),

$$\begin{aligned}\overline{sz} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} s\, z(U)\, p(s, U)\, ds\, dU \\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} s\, p(s|U)\, ds\right] z(U)\, p(U)\, dU \\ &= \int_{-\infty}^{\infty} z_\infty(U)\, z(U)\, p(U)\, dU \\ &= \overline{z_\infty z}\end{aligned} \tag{25}$$

The relation

$$\rho_{sz} = \rho_{sz_\infty}\rho_{z_\infty z}$$

can be derived from Eq. 25 and from the definition of Eq. 20. But $\left|\rho_{z_\infty z}\right|_{max} = 1$, the maximum being achieved whenever $z(U)$ is a (nonzero) scalar multiple of $z_\infty(U)$. Therefore, by Eq. 21, we have

$$\gamma_{max} = \frac{1}{\dfrac{1}{\rho^2_{sz_\infty}} - 1} \tag{26}$$

That is, a maximum signal-to-noise-ratio filter is a least-squares filter followed by any amount of ideal, noiseless gain. In particular, a least-squares filter is a maximum signal-to-noise-ratio filter. (Remember that in these calculations the linear part of the filter has been held invariant, and only the nonlinear function has been changed.)

There is an alternative expression for the correlation coefficient associated with least-squares filters that also happens to be valid for orthogonal approximations to the least-squares filter. Such filters arise in the Wiener representation, as we shall see. An orthogonal approximation, $z_N(U)$, is defined to satisfy the equation

$$\overline{z_N (z_\infty - z_N)} = 0 \tag{27}$$

From Eqs. 25, 27, and the definition, Eq. 20,

$$\rho^2_{sz_N} = \frac{\overline{z_N^2}}{\overline{s^2}} \tag{28}$$

This formula contains one less average than the one in Eq. 20.

2.4 DISCRETE-SYSTEM EXAMPLE

Optimum discrete systems operating at low noise levels tend to have much higher output signal-to-noise ratios than comparable continuous systems. As a simple example, let the noisy nonlinear operator in Fig. 3 be a binary symmetric channel with probability of error p_e. If the signal source output, s, assumes the values +1 or −1 with equal probability, the least-squares function (conditional mean) is

$$z_\infty(\pm 1) = \pm(1-2p_e)$$

Next, define

$$p_e = 1 - \Phi\left(\gamma_i^{1/2}\right) \tag{29}$$

where $\Phi(x)$ is the normal distribution function. (This is a model of a binary transmission system subject to additive Gaussian noise, with signal-to-noise ratio γ_i at the receiver input.) By approximating $1 - \Phi(x)$ (see Davenport and Root (10) or Feller (11)) and using the formulas of section 2.3, we find that

$$\gamma_o \sim \left(\frac{\pi}{8}\gamma_i\right)^{1/2} e^{\gamma_i/2} \tag{30}$$

as $\gamma_i \to \infty$. This demonstrates our contention because the continuous systems that are commonly used have γ_o proportional to γ_i for large γ_i.

III. WIENER FILTERING

3.1 GENERAL REMARKS

This section concerns the construction and performance of optimum least-squares Wiener filters for use in noisy communication systems of the type indicated in Fig. 1. The nonlinear, memoryless part of the Wiener filter, a polynomial, is constructed from basic polynomials having an orthogonality property that will be discussed presently. Two random variables, x and y, are said to be orthogonal if $\overline{xy} = 0$. Whenever $\overline{x} = 0$ or $\overline{y} = 0$, orthogonality is therefore the same as linear independence. (The terminology suggests that x and y be viewed as vectors, with inner product $\overline{xy}$. This convenient viewpoint is not exploited in the present study.) Although we shall not explore the point in detail, it is worth noting that an orthogonal representation for the least-squares polynomial guarantees that a sequence of approximating polynomials converges in the mean-square sense to a function that is independent of the exact nature of the sequence. Other advantages of this orthogonal representation will be made clear in the discussion.

It will often be necessary to indicate a product of U-variables (Eq. 12) in a form such as

$$u_{j_1} \dots u_{j_N}$$

in which the subscripts j_m and j_n may be equal if $m \neq n$. The point of this notation is to distinctly label every variable in the product, whether or not the variables are all distinct.

3.2 ORTHOGONAL BASIC POLYNOMIALS

The material presented in sections 3.2 and 3.3 has been treated in one way or another by Wiener (2) and Barrett (3), not to mention others. The generality of the following discussion is sufficient to include systems of the class illustrated by Fig. 3. The emphasis placed on the nonlinear, memoryless part of the Wiener filter by a Fig. 3 system is advantageous, as we have mentioned.

An N-degree basic polynomial in the set $U = \{u_1, \dots, u_L\}$ is defined as a single N-degree leading term, $u_{j_1} \dots u_{j_N}$, plus a polynomial of less-than-N-degree, the latter being so chosen that the entire basic polynomial is orthogonal to any less-than-N-degree polynomial in variables of the set U. The 0-degree basic polynomial is defined to be the constant, 1. Since all but the 0-degree basic polynomial are orthogonal to constants, they must have zero means, and orthogonality is then directly equivalent to linear independence. Notation for a basic polynomial is

$$F_U\left(u_{j_1} \dots u_{j_N}\right) \tag{31}$$

Notice that the leading term (hence the degree) and the set over which orthogonality holds are explicitly exhibited.

It is important, but nearly trivial, to observe that the basic polynomials form a basis for all polynomials. This fact does not depend on orthogonality. Specifically, an arbitrary N-degree polynomial P(U) in the set U can be expressed as a weighted sum of basic polynomials with degrees not exceeding N. In proving this statement we may inductively assume its truth for all degrees less than N. But the arbitrary N-degree polynomial can be written in the form

$$P(U) = \sum_{\substack{j_1=1 \\ \cdots \\ j_N=1 \\ j_1\leq\cdots\leq j_N}}^{L} a_{j_1\cdots j_N} u_{j_1} \cdots u_{j_N} + Q(U) \tag{32}$$

where Q(U) is a less-than-N-degree polynomial. From the definition of a basic polynomial, it follows that

$$\sum_{\substack{j_1=1 \\ \cdots \\ j_N=1 \\ j_1\leq\cdots\leq j_N}}^{L} a_{j_1\cdots j_N} F_U\left(u_{j_1}\cdots u_{j_N}\right) = \sum_{\substack{j_1=1 \\ \cdots \\ j_N=1 \\ j_1\leq\cdots\leq j_N}}^{L} a_{j_1\cdots j_N} u_{j_1} \cdots u_{j_N} + R(U)$$

where R(U) is a less-than-N-degree polynomial. By using simple arithmetic, we obtain

$$P(U) = \sum_{\substack{j_1=1 \\ \cdots \\ J_N=1 \\ j_1\leq\cdots\leq j_N}}^{L} a_{j_1\cdots j_N} F_U\left(u_{j_1}\cdots u_{j_N}\right) - R(U) + Q(U) \tag{33}$$

Since $-R(U) + Q(U)$ can be expressed, by assumption, as a weighted sum of less-than-N-degree basic polynomials, the argument is complete.

There are several possible procedures for constructing the basic polynomials. One of the simplest is a variation of the Gram-Schmidt orthogonalization procedure (12). From the basis property just described, it is sufficient that a basic polynomial be orthogonal to all lower-degree basic polynomials. Assume that all basic polynomials in the set U of less-than-N-degree have been constructed. The form of the desired N-degree basic polynomial, with leading term $u_{j_1} \ldots u_{j_N}$, can be expressed as

$$F_U\left(u_{j_1}\cdots u_{j_N}\right) = u_{j_1}\cdots u_{j_N} - \sum_{n=0}^{N-1} \sum_{\substack{k_1=1\\ \cdots\\ k_n=1\\ k_1\le\cdots\le k_n}}^{L} a_{k_1\ldots k_n} F_U\left(u_{k_1}\cdots u_{k_n}\right) \tag{34}$$

The required orthogonality property is given by

$$\overline{F_U\left(u_{j_1}\cdots u_{j_N}\right) F_U\left(u_{m_1}\cdots u_{m_M}\right)} = 0 \tag{35}$$

whenever $M < N$. By using Eq. 34 and the assumption that $F_U\left(u_{m_1}\cdots u_{m_M}\right)$ is orthogonal to $F_U\left(u_{k_1}\cdots u_{k_n}\right)$ if $n \neq M$ and $n < N$, Eq. 35 can be written as the condition

$$\overline{u_{j_1}\cdots u_{j_N} F_U\left(u_{m_1}\cdots u_{m_M}\right)} = \sum_{\substack{k_1=1\\ \cdots\\ k_M=1\\ k_1\le\cdots\le k_M}}^{L} a_{k_1\ldots k_M} \overline{F_U\left(u_{k_1}\cdots u_{k_M}\right) F_U\left(u_{m_1}\cdots u_{m_M}\right)} \tag{36}$$

For a fixed value of M, the $C_M^{M+L-1} = \dfrac{(M+L-1)!}{M!\,(L-1)!}$ equations of the form of Eq. 36, corresponding to all possible distinct sets of subscripts $\{m_1, \ldots, m_M\}$, constitute a set of simultaneous linear equations in the coefficients $a_{k_1\ldots k_M}$. This set of equations can be solved by Cramer's rule (12), unless the collection of all C_M^{M+L-1} basic polynomials of M-degree is linearly dependent (that is, a nontrivial weighted sum of these basic polynomials can be made to vanish identically). The last possibility can be eliminated by making the following assumptions:

(a) The random variables in the set U are not linked deterministically; that is, there is no function f such that for some j, $u_j = f(U-u_j)$, where $U - u_j$ is the set $\{u_1, \ldots, u_{j-1}, u_{j+1}, \ldots, u_L\}$.

(b) All variables in U range over an infinite number of values.

The first assumption eliminates redundancy in the variables of U. The second assumption is fulfilled in systems like that of Fig. 1 because the additive Gaussian noise has a continuous amplitude distribution. In a finite-state Fig. 3 system, limitation to a suitable finite set of basic polynomials will satisfy the second assumption. Therefore, the coefficients in Eq. 34 can be uniquely chosen by Eqs. 36 to satisfy Eq. 35, and the orthogonalization procedure is complete.

The basic polynomials in the set $cU = \{cu_1, \ldots, cu_L\}$ and those in U are related by the equation

$$F_{cU}\left(cu_{j_1}\dots cu_{j_N}\right) = c^N F_U\left(u_{j_1}\dots u_{j_N}\right) \tag{37}$$

Equation 37 follows from Eq. 36 by induction. Notice that the coefficients in Eq. 36 must satisfy

$$a_{k_1\dots k_M} \propto c^{N-M}$$

as c is varied, if we use the set cU.

3.3 GENERALIZED HERMITE POLYNOMIALS

Here, U is assumed to be a set of zero-mean random variables with a joint Gaussian probability distribution. The corresponding basic polynomials are the generalized Hermite polynomials, designated $G_U\left(u_{j_1}\dots u_{j_N}\right)$ to distinguish them from basic polynomials in non-Gaussian sets. The Hermite polynomials and Gaussian variables from which they are derived possess several useful averaging properties that are not true in non-Gaussian cases.

The fundamental averaging property of zero-mean Gaussian variables is the pairing formula for averaged products:

$$\overline{u_{j_1}\dots u_{j_N}} = \sum_{\substack{\text{distinct}\\ \text{pairings}}} \overline{u_{j_{T(1)}} u_{j_{T(2)}}} \dots \overline{u_{j_{T(N-1)}} u_{j_{T(N)}}} \tag{38}$$

The method indicated by Eq. 38 is to arrange the variables of the product $u_{j_1}\dots u_{j_N}$ in pairs, average each pair independently, and then sum over all distinct ways of pairing (distinct with respect to the second subscripts). Mathematically, T(1), ..., T(N) is a permutation of 1, ..., N, and the sum is over those permutation operators, T, which produce distinct pairwise combinations of 1, ..., N. Notice that the required pairing is impossible if N is odd. This is to be interpreted by the fact that $\overline{u_{j_1}\dots u_{j_N}} = 0$ for odd N. If N is even, the number of pairings is readily deduced: Pick any variable and pair it with one of the other N−1 variables, pick a third and pair it with one of the remaining N−3, and so on. This number may be expressed as

$$(N-1)\,(N-3)\dots(1) = \frac{N!}{2^N\left(\frac{N}{2}\right)!} \tag{39}$$

The average in Eq. 38 is readily derived as the coefficient of $\theta_1\dots\theta_N$ in the power-series expansion (14) of the joint moment-generating function, $M(\theta_1,\dots,\theta_N)$, of the variables $u_{j_1}, \dots, u_{j_N}$. For zero-mean Gaussian variables, the moment-generating function (15) expansion is given by

$$M(\theta_1, \ldots, \theta_N) = \exp\left(\frac{1}{2} \sum_{\substack{m=1 \\ n=1}}^{N} \overline{u_{j_m} u_{j_n}}\, \theta_m \theta_n\right)$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{1}{2} \sum_{\substack{m=1 \\ n=1}}^{N} \overline{u_{j_m} u_{j_n}}\, \theta_m \theta_n\right)^k$$

$$= \sum_{k=0}^{\infty} \frac{1}{2^k k!} \sum_{\substack{m_1=1 \\ \cdots \\ m_k=1 \\ n_1=1 \\ \cdots \\ n_k=1}}^{N} \overline{u_{j_{m_1}} u_{j_{n_1}}} \cdots \overline{u_{j_{m_k}} u_{j_{n_k}}}\, \theta_{m_1} \theta_{n_1} \cdots \theta_{m_k} \theta_{n_k} \tag{40}$$

Inspection of Eq. 40 shows that the coefficient of $\theta_1 \ldots \theta_N$ must come from terms for which $k = N/2$. This proves Eq. 38 if N is odd. For even N, this coefficient is

$$\sum_T \frac{1}{2^{N/2}\left(\frac{N}{2}\right)!} \overline{u_{j_{T(1)}} u_{j_{T(2)}}} \cdots \overline{u_{j_{T(N-1)}} u_{j_{T(N)}}} \tag{41}$$

where the sum includes all permutation operators, T. There are $2^{N/2}$ identical permutations formed by reversing subscripts in each pair; and $(N/2)!$ ways of ordering the pairs. Adding identical permutations and only summing T over distinct combinations therefore leads to Eq. 38.

Since odd-degree products of zero-mean Gaussian random variables have zero mean values, polynomials having only odd-degree terms must be orthogonal to even-degree polynomials. It follows that the basic Hermite polynomials must have all odd-degree or all even-degree terms. It is worth noticing that these statements are true for any sets of variables, U, in which the joint probability density function (or discrete probabilities) possesses the symmetry property

$$p(u_1, \ldots, u_L) = p(-u_1, \ldots, -u_L)$$

The next important formula simplifies the calculation of averages involving a Hermite polynomial.

$$\overline{u_{j_1} \cdots u_{j_N} G_U\left(u_{k_1} \cdots u_{k_M}\right)}$$

$$= \sum_{\substack{\text{combinations} \\ \text{of M } u_j\text{'s} \\ \text{from } u_{j_1} \ldots u_{j_N}}} \overline{u_{j_{T(M+1)}} \cdots u_{j_{T(N)}}} \sum_{\substack{u_j u_k \\ \text{pairings}}} \overline{u_{j_{T(1)}} u_{k_1}} \cdots \overline{u_{j_{T(M)}} u_{k_M}} \tag{42}$$

That is, pair variables in the Hermite polynomial leading term across to a like number of variables in the other product, average the pairs, and sum over all M! such cross pairings; average the remaining unpaired terms as an independent product, and finally sum over all distinct ways of dividing the product $u_{j_1} \ldots u_{j_N}$ into the two required classes. The averaged product $\overline{u_{j_{T(M+1)}} \ldots u_{j_{T(N)}}}$ can be found with the aid of the pairing formula Eq. 38. If M + N is odd, both sides of Eq. 42 vanish, by the odd-even orthogonality pointed out above. If N < M, the left side vanishes by the Hermite polynomial orthogonality, and the right side disappears because the required pairing is impossible. If N = M, the right side is composed solely of the inner sum. If M = 0, the inner sum is defined equal to unity. The following simple example may help clarify operation of the formula.

$$\overline{u_1u_2u_3u_4\, G_U(u_5u_6)} = \sum_{(abcd)} \overline{u_au_b}\left(\overline{u_cu_5}\;\overline{u_du_6} + \overline{u_cu_6}\;\overline{u_du_5}\right) \tag{43}$$

where the subscript set (abcd) is summed over the six distinct permutations (1234), (1324), (1423), (2314), (2413), and (3412), or any equivalent sets found by interchanging a and b or c and d.

Equation 42 is best proved with the aid of an auxiliary classification scheme. The average on the left side of Eq. 42 can be expanded into a sum of products of averaged pairs with the aid of Eq. 38. We classify each such product by the number of pairs, K, containing both a u_j and a u_k. The following example illustrates this classification scheme.

$$\overline{u_{j_1}u_{j_2}u_{k_1}u_{k_2}} = \left(\overline{u_{j_1}u_{k_1}}\;\overline{u_{j_2}u_{k_2}} + \overline{u_{j_1}u_{k_2}}\;\overline{u_{j_2}u_{k_1}}\right)_{K=2} + \left(\overline{u_{j_1}u_{j_2}}\;\overline{u_{k_1}u_{k_2}}\right)_{K=0}$$

Returning to Eq. 42, an expansion of the left side might be expected to include classes for K = M, M − 2, . . . , 0 (if M is even, or ending with K = 1 if M is odd). Examination of the right side shows that it is precisely the M-class of the left side. Hence, we must show that all K-classes for K < M vanish from the average on the left side.

It is sufficient to consider only those cases for which M > 0, N ≥ M, N + M is even, and K + M is even. If M is even, a little study shows that the 0-class for the left side of Eq. 42 contains those terms that make up the average

$$\overline{u_{j_1} \cdots u_{j_N}}\;\overline{G_U\left(u_{k_1} \cdots u_{k_M}\right)}$$

But $\overline{G_U\left(u_{k_1} \cdots u_{k_M}\right)} = 0$, so the 0-class vanishes. If M is odd, the 1-class on the left side of Eq. 42 is the sum of averages

$$\sum_{n=1}^{N} \overline{u_{j_1} \cdots u_{j_{(n-1)}}u_{j_{(n+1)}} \cdots u_{j_N}}\;\overline{u_{j_n}\, G_U\left(u_{k_1} \cdots u_{k_M}\right)}$$

If M = 1, Eq. 42 gives this result because $G_U(u_k) = u_k$. If M > 1, $\overline{u_{j_n}\, G_U\left(u_{k_1} \cdots u_{k_M}\right)} = 0$ for each n, and the 1-class vanishes.

We now assume inductively that for some K^* such that $K^* < M$ and $K^* + M$ is even, all K-classes of the average on the left side of Eq. 42 vanish when $K \leq K^*$. If $K^* = M - 2$, the formula is proved. If $K^* < M - 2$, consider the quantity

$$\overline{u_{j_{T(K^*+3)}} \cdots u_{j_{T(N)}}}\ \overline{u_{j_{T(1)}} \cdots u_{j_{T(K^*+2)}} G_U\left(u_{k_1} \cdots u_{k_M}\right)} \tag{44}$$

The possible K-classes of this average all have $K \leq K^* + 2$. By induction, the K-classes for which $K < K^* + 2$ vanish. On the other hand, the whole average in Eq. 44 is zero because of the orthogonality of $G_U\left(u_{k_1} \cdots u_{k_M}\right)$. Therefore, the $K^* + 2$ -class also vanishes. If we sum the average in Eq. 44 over all ways of choosing $K^* + 2$ of the u_j-variables, it follows that the whole (K^*+2) -class for the left side of Eq. 42 vanishes. This completes the inductive argument; the K-classes with $K < M$ vanish, and Eq. 42 is proved.

The averaged product of two N-degree Hermite polynomials is now readily derived.

$$\overline{G_U\left(u_{j_1} \cdots u_{j_N}\right) G_U\left(u_{k_1} \cdots u_{k_N}\right)} = \overline{u_{j_1} \cdots u_{j_N} G_U\left(u_{k_1} \cdots u_{k_N}\right)}$$

$$= \sum_{\substack{u_j u_k \\ \text{pairings}}} \overline{u_{j_{T(1)}} u_{k_1}} \cdots \overline{u_{j_{T(N)}} u_{k_N}} \tag{45}$$

That is, pair across and average variables between the two leading terms in all possible ways. The first equality holds because the second polynomial is orthogonal to all lower-degree terms of the first polynomial. The second equality is a direct application of Eq. 42.

The basic Hermite polynomials can be expressed in a form similar to that of Eq. 34, but with explicit expressions for the coefficients.

$$G_U\left(u_{j_1} \cdots u_{j_N}\right) = u_{j_1} \cdots u_{j_N} - \sum_{\substack{n=0, \\ (N+n) \\ \text{even}}}^{N-2} \sum_{\substack{\text{combinations} \\ \text{of } n\ u_j\text{'s} \\ \text{from } u_{j_1} \cdots u_{j_N}}} \overline{u_{j_{T(n+1)}} \cdots u_{j_{T(N)}}}\, G_U\left(u_{j_{T(1)}} \cdots u_{j_{T(n)}}\right) \tag{46}$$

As usual, the coefficients $\overline{u_{j_{T(n+1)}} \cdots u_{j_{T(N)}}}$ may be expanded as in Eq. 38. Proof of formula 46 consists in showing orthogonality of the right-hand side to any $G_U\left(u_{k_1} \cdots u_{k_M}\right)$ for $M < N$. The average $\overline{u_{j_1} \cdots u_{j_N} G_U\left(u_{k_1} \cdots u_{k_M}\right)}$ is given directly by Eq. 42. For the remaining terms, the average $\overline{G_U\left(u_{j_{T(1)}} \cdots u_{j_{T(n)}}\right) G_U\left(u_{k_1} \cdots u_{k_M}\right)}$ vanishes unless n = M, in which case Eq. 45 leads to the result. If all of these averages are assembled, careful observation shows that they cancel. Hence, Eq. 46 is correct.

The representation of Eq. 46 can be built up in a step-by-step manner which is illustrated by the following example. We shall choose for the leading term $u_1^2u_2^3$. Pick three variables in this term, average the remaining pair, sum over ways of choosing distinct combinations of three variables, and subtract the result from the leading term. Then we have

$$u_1^2u_2^3 - \overline{u_1^2}\,u_2^3 - 6\,\overline{u_1u_2}\,u_1u_2^2 - 3\,\overline{u_2^2}\,u_1^2u_2 \tag{47}$$

The third-degree terms are leading terms of the third-degree Hermite polynomials in Eq. 46 and have the coefficients given. It follows that expression 47 is orthogonal to all third-degree Hermite polynomials. Averaging expression 47, with one variable in each term fixed, and summing over all possible ways of picking out this variable completes the procedure by orthogonalizing the expression to first-degree polynomials. Thus

$$\begin{aligned} G_U\left(u_1^2u_2^3\right) &= u_1^2u_2^3 - \overline{u_1^2}\left(u_2^3 - 3\,\overline{u_2^2}\,u_2\right) - 6\,\overline{u_1u_2}\left(u_1u_2^2 - 2\,\overline{u_1u_2}\,u_2 - \overline{u_2^2}\,u_1\right) \\ &\quad - 3\,\overline{u_2^2}\left(u_1^2u_2 - \overline{u_1^2}\,u_2 - 2\,\overline{u_1u_2}\,u_1\right) - 3\,\overline{u_1^2u_2^2}\,u_2 - 2\,\overline{u_1u_2^3}\,u_1 \end{aligned} \tag{48}$$

The last two coefficients can be expanded by using Eq. 38 to obtain

$$\left.\begin{aligned} \overline{u_1^2u_2^2} &= \overline{u_1^2}\,\overline{u_2^2} + 2\,\overline{u_1u_2}^{\,2} \\ \overline{u_1u_2^3} &= 3\,\overline{u_1u_2}\,\overline{u_2^2} \end{aligned}\right\} \tag{49}$$

The expression in Eq. 46 shows that a Hermite polynomial is determined solely by the variables in its leading term. This contrasts with the general situation (Eq. 34) in which any or all variables from the set U can appear in the lower-degree terms. Thus, the Hermite polynomials are said to be completely orthogonal, whereas the general basic polynomials are said to be partially orthogonal. One advantage of complete orthogonality will appear later, when averages of products of basic polynomials belonging to different sets of variables are required. If the two sets are jointly Gaussian, the orthogonality property still holds. Otherwise, polynomials from the two sets do not have this property.

3.4 LEAST-SQUARES POLYNOMIAL FILTERS

We now discuss the construction of Wiener filters, using the basic polynomials just described. These filters are designed to minimize the mean-square error, $\overline{(z-s)^2}$, and therefore to maximize the output signal-to-noise ratio, in Fig. 1 systems. The linear networks of the Wiener filter are assumed to be fixed in advance, as before, and the nonlinear, memoryless part of the filter is then adjusted to give optimum performance. The quadrature variables (q, y, v) are again suppressed, since their explicit inclusion would not be of help.

The least-squares polynomial takes the form

$$z_N(U) = \sum_{n=1}^{N} \sum_{\substack{j_1=1 \\ \cdots \\ j_n=1 \\ j_1 \le \cdots \le j_n}}^{L} b_{j_1 \ldots j_n} F_U\left(u_{j_1} \ldots u_{j_n}\right) \tag{50}$$

There is no constant (n=0) term because $\overline{z_N(U)} = 0$. (For least-squares filtering, $\overline{z} = \overline{s}$, but $\overline{s} = 0$.) The coefficients $b_{j_1 \ldots j_n}$ are determined in the usual Fourier-series manner (16) by equating the averages

$$\overline{s\, F_U\left(u_{k_1} \ldots u_{k_M}\right)} = \overline{z_N(U)\, F_U\left(u_{k_1} \ldots u_{k_M}\right)} \tag{51}$$

From Eq. 50 and the orthogonality of different-degree basic polynomials, Eq. 51 can be written

$$\overline{s\, F_U\left(u_{k_1} \ldots u_{k_M}\right)} = \sum_{\substack{j_1=1 \\ \cdots \\ j_M=1 \\ j_1 \le \cdots \le j_M}}^{L} b_{j_1 \ldots j_M} \overline{F_U\left(u_{j_1} \ldots u_{j_M}\right) F_U\left(u_{k_1} \ldots u_{k_M}\right)} \tag{52}$$

Given any fixed degree M, the C_M^{M+L-1} equations of the form of Eq. 52 must be solved simultaneously for the coefficients. That a unique Cramer's rule solution exists follows from the same considerations discussed in connection with Eq. 36.

It is important to observe that

$$z_N(cU) = z_N(U) \tag{53}$$

This looks surprising until we recall that the two z_N's are different polynomials, with $z_N(cU)$ optimum for input cU, and $z_N(U)$ optimum for input U. We can derive Eq. 53 from Eqs. 37, 52, and 50. Physically, Eq. 53 must be correct because scalar multiplication of signal u should not change the operation of the appropriately revised optimum system. The difference in functional form between $z_N(cU)$ and $z_N(U)$ reflects a practical difficulty with optimum nonlinear filters, namely, the dependence of their performance on input signal level. Incidentally, everything mentioned thus far applies also to the more general system shown in Fig. 3.

Suppose that the noise in the Fig. 1 system is zero ($x \equiv 0$), so that $u \equiv p$. Then the least-square Wiener polynomial, now represented by $\hat{z}_N(P)$ $(P = \{p_1, \ldots, p_L\})$, is an inverse or an approximate inverse of the nonlinear operator. As in Eq. 50,

$$\hat{z}_N(P) = \sum_{n=1}^{N} \sum_{\substack{j_1=1 \\ \cdots \\ j_n=1 \\ j_1 \le \cdots \le j_n}}^{L} \hat{b}_{j_1 \cdots j_n} F_P\left(p_{j_1} \cdots p_{j_n}\right) \tag{54}$$

in which the F_P's are the basic polynomials in P. The coefficients $\hat{b}_{j_1 \cdots j_n}$ are determined as in Eq. 52.

One important property of $\hat{z}_N(P)$ is that it can replace s in Eqs. 51 and 52. For,

$$u_{j_1} \cdots u_{j_n} = \sum_{m=0}^{n} \sum_{\substack{\text{distinct} \\ \text{combinations}}} p_{j_{T(1)}} \cdots p_{j_{T(m)}} x_{j_{T(m+1)}} \cdots x_{j_{T(n)}} \tag{55}$$

where T is a permutation operator. This expansion follows directly from the fact that $u_j = p_j + x_j$. Hence,

$$\overline{s\, u_{j_1} \cdots u_{j_n}} = \sum_{\substack{m=0, \\ n-m \text{ even}}}^{n} \sum_{\substack{\text{distinct} \\ \text{combinations}}} \overline{s\, p_{j_{T(1)}} \cdots p_{j_{T(m)}}}\; \overline{x_{j_{T(m+1)}} \cdots x_{j_{T(n)}}} \tag{56}$$

because x is independent of s and p, and $\overline{x_{j_{T(m+1)}} \cdots x_{j_{T(n)}}} = 0$ when $n - m$ is odd. But, by construction (Eq. 51),

$$\overline{s\, p_{j_{T(1)}} \cdots p_{j_{T(m)}}} = \overline{\hat{z}_N(P)\, p_{j_{T(1)}} \cdots p_{j_{T(m)}}} \tag{57}$$

if $N \ge m$. Substituting $\hat{z}_N(P)$ for s in Eq. 56 therefore changes nothing, and our contention is proved.

We observe that if $\hat{b}_{j_1 \cdots j_n} = 0$ (Eq. 54) whenever $n < N_o$, then $b_{k_1 \cdots k_m} = 0$ (Eq. 50) whenever $m < N_o$. For, Eq. 57 must vanish if $m < N_o$, and thus, because of Eq. 56, the left side of Eq. 52 is zero if $M < N_o$. But this means that Eqs. 52 are homogeneous for any fixed $M < N_o$, and the corresponding coefficients are all zero. If some $\hat{b}_{j_1 \cdots j_n} \ne 0$, Eqs. 52 are not a homogeneous set for $M = n$, and some $b_{k_1 \cdots k_n} \ne 0$. Taken together, these results demonstrate the important fact that the lowest degrees of the nonvanishing basic polynomials comprising $\hat{z}_N(P)$ and $z_N(U)$ are the same (designated N_o).

Incidentally, it can be shown that $\hat{z}_N(P)$ is a mean-square limit of $z_N(U)$ as $\overline{x^2} \to 0$, if $\overline{p^2}$ and $\overline{s^2}$ are held constant. This fact depends on the continuity of the polynomials representing these functions. However, a detailed proof is beyond the scope of this report.

3.5 OPTIMUM LOW SIGNAL-TO-NOISE-RATIO FILTER PERFORMANCE

The output signal-to-noise ratio, γ_o, of a least-squares filter, $z_N(U)$, in a system such as that of Fig. 1 or Fig. 3 is a function solely of the quantity $\overline{Z_N^2}/\overline{s^2}$, from Eqs. 21 and 28. Assuming that $\overline{s^2}$ is fixed, we shall study the behavior of $\overline{z_N^2(U)}$, and hence of γ_o, as a function of the input signal-to-noise ratio, $\gamma_i = \overline{p^2}/\overline{x^2}$ (Eq. 23). For convenience, $\overline{p^2}$ is assumed to be fixed. Then, qualitative results follow from the proportionality

$$\gamma_i \propto \frac{1}{\overline{x^2}} \tag{58}$$

To justify the use of Eq. 58, we must show that multiplying $\overline{p^2}$ and $\overline{x^2}$ by the same positive scalar (which leaves their ratio, γ_i, invariant) has no effect on $\overline{z_N^2(U)}$. But this is the same thing as replacing the signal $u = p + x$ by the signal $cu = cp + cx$ and using the modified $z_N(cU)$. It follows from Eq. 53 that

$$\overline{z_N^2(cU)} = \overline{z_N^2(U)} \tag{59}$$

as required.

From Eq. 50 and the orthogonality of different-degree basic polynomials,

$$\overline{z_N^2(U)} = \sum_{n=N_o}^{N} \sum_{\substack{j_1=1 \\ \cdots \\ j_n=1 \\ k_1=1 \\ \cdots \\ k_n=1 \\ j_1\le\cdots\le j_n \\ k_1\le\cdots\le k_n}}^{L} b_{j_1\ldots j_n} b_{k_1\ldots k_n} \overline{F_U\left(u_{j_1}\cdots u_{j_n}\right)\left(F_U\, u_{k_1}\cdots u_{k_n}\right)} \tag{60}$$

where N_o is the lowest degree of the nonvanishing basic polynomials in $z_N(U)$, from the discussion in section 3.4. The use of Eqs. 52, 34, and 36 demonstrates that $\overline{z_N^2(U)}$ is equivalent to a rational fraction in averaged products of u_j-variables; some averages also include the variable s. By using Eqs. 55-57, $\overline{z_N^2(U)}$ can then be written as a rational fraction in averages of products of p_j- and x_k-variables. Holding $\overline{p^2}$ constant means that all p_j-averages (and $\overline{s^2}$) remain constant. But

$$\overline{x_{j_1}\cdots x_{j_n}} \propto \overline{x^2}^{\,n/2} \tag{61}$$

Since the average on the left is zero for odd n, only integral powers of $\overline{x^2}$ will arise. We are especially interested in the situation in which γ_i is small — which means large

$\overline{x^2}$ relative to $\overline{p^2}$ (proportionality 58). Thus, an expansion of $\overline{z_N^2(U)}$ in the largest powers of $\overline{x^2}$, by Eq. 61, is equivalent to an expansion of $\overline{z_N^2(U)}$ in small powers of γ_i. From Eq. 22,

$$\gamma_o = \sum_{k=1}^{\infty} \left(\frac{\overline{z_N^2}}{\overline{s^2}} \right)^k \tag{62}$$

Finding the M smallest powers of γ_i in a small-γ_i expansion of $\overline{z_N^2(U)}$ and using the first M terms in Eq. 62 provides an M-term expansion of γ_o in the lowest powers of γ_i. This is most useful near the limit $\gamma_i = 0$, where we only need M = 1 or 2.

From a study of Eqs. 55 and 37 we see that the average

$$\overline{F_U\left(u_{j_1} \dots u_{j_n}\right) F_U\left(u_{k_1} \dots u_{k_n}\right)} \tag{63}$$

is of order $\overline{x^2}^{\,n}$, or lower, for large $\overline{x^2}$. In order to find the order of magnitude of the $b_{j_1 \dots j_n}$ coefficients, it is helpful to obtain a number of intermediate results.

Consider the average

$$\overline{u_{j_1} \dots u_{j_n}} = \sum_{\substack{m=0, \\ n-m \text{ even}}}^{n} \sum_{\substack{\text{distinct} \\ \text{combinations}}} \overline{p_{j_{T(1)}} \dots p_{j_{T(m)}}}\; \overline{x_{j_{T(m+1)}} \dots x_{j_{T(n)}}} \tag{64}$$

which follows directly from Eq. 55 and the independence of p and x. This average is dominated by small-m terms when $\overline{x^2}$ is large, and it might be expected to behave somewhat like an averaged product of Gaussian variables. We shall show that the pairing formula taken from Eq. 38,

$$\overline{u_{j_1} \dots u_{j_n}} \sim \sum_{\substack{\text{distinct} \\ \text{pairings}}} \overline{u_{j_{T(1)}} u_{j_{T(2)}}} \dots \overline{u_{j_{T(n-1)}} u_{j_{T(n)}}} \tag{65}$$

is correct for the m = 0, 1, and 2 terms of Eq. 64. That is, formula 65 is correct for terms proportional to $\overline{x^2}^{\,n/2}$ and $\overline{x^2}^{\,(n-2)/2}$ if n is even, or to $\overline{x^2}^{\,(n-1)/2}$ if n is odd. To obtain these results, we must assume that $\overline{p} = 0$. This assumption is not very restrictive because we are ultimately interested in the basic polynomials for which

$$F_U\left[\left(u_{j_1} - \overline{p_{j_1}}\right) \dots \left(u_{j_n} - \overline{p_{j_n}}\right)\right] = F_U\left(u_{j_1} \dots u_{j_n}\right) \tag{66}$$

This follows from the unique representation argument of Eqs. 32-33.

First, suppose that n is even. The right side of Eq. 65 can then be written

$$\sum_{\substack{m=0,\\ m\ \text{even}}}^{n} \sum_{\substack{\text{distinct}\\ \text{pairings}}} \overline{p_{j_{T(1)}} p_{j_{T(2)}}} \cdots \overline{p_{j_{T(m-1)}} p_{j_{T(m)}}}$$

$$\times\ \overline{x_{j_{T(m+1)}} x_{j_{T(m+2)}}} \cdots \overline{x_{j_{T(n-1)}} x_{j_{T(n)}}} \tag{67}$$

because $\overline{u_j u_k} = \overline{p_j p_k} + \overline{x_j x_k}$. The m = 0 terms in expression 67 are just the pairing formula expansion of the m = 0 average in Eq. 64. Similarly, it can be seen that the m = 2 terms in Eqs. 64 and 67 must also agree. Neither expression has m = 1 terms; thus the approximation claimed for Eq. 65 is correct when n is even.

Now, suppose n is odd. The right side of Eq. 65 vanishes because the pairing is impossible. There are no m = 0 or m = 2 terms in Eq. 64 now, and the m = 1 terms are zero because $\overline{p} = 0$ implies $\overline{p_j} = 0$ for all j. Therefore, the approximation claimed for Eq. 65 is also true for odd n, and hence for all n.

Recall that the pairing formula was the basis for deriving all Hermite polynomial formulas exhibited in section 3.3. Therefore, these formulas apply whenever averaged products of u_j-variables are under consideration, and are correct for terms proportional to the highest possible power of $\overline{x^2}$. We shall show that these formulas apply to a limited (but useful) extent when the averages include basic polynomials in P. The structure of the basic polynomials in U was given by Eq. 34:

$$F_U\left(u_{j_1} \ldots u_{j_n}\right) = u_{j_1} \ldots u_{j_n} - \sum_{m=0}^{n-1} \sum_{\substack{k_1=1\\ \vdots\\ k_m=1\\ k_1 \le \ldots \le k_m}}^{L} a_{k_1 \ldots k_m} F_U\left(u_{k_1} \ldots u_{k_m}\right)$$

We would expect Eq. 46 to be a good large-$\overline{x^2}$ approximation to this equation; this means that

$$a_{k_1 \ldots k_m} \sim \overline{u_{j_{T(m+1)}} \cdots u_{j_{T(n)}}} \tag{68}$$

if $\{k_1, \ldots, k_m\} = \{j_{T(1)}, \ldots, j_{T(m)}\}$ (a subset of the leading term subscripts $\{j_1, \ldots, j_n\}$) and n − m is even. Otherwise, we should have $a_{k_1 \ldots k_m} \sim 0$. From the preceding discussion, these approximations must be correct for terms proportional to $\overline{x^2}^{(n-m)/2}$ (n − m even) or to $\overline{x^2}^{(n-m-1)/2}$ (n − m odd and $\overline{p} = 0$) in a large-$\overline{x^2}$ expansion of $a_{k_1 \ldots k_m}$. No higher powers of $\overline{x^2}$ can be present in such an expansion because of Eqs. 64 and the proportionality, $a_{k_1 \ldots k_M} \propto c^{N-M}$ Hence, Eq. 68

is correct for the terms proportional to the highest possible $\overline{x^2}$ power.

Next, we investigate the average

$$\overline{F_P\left(p_{j_1}\cdots p_{j_n}\right) u_{k_1}\cdots u_{k_m}}$$

$$= \sum_{\substack{K=n,\\ m-K \text{ even}}}^{m} \sum_{\substack{\text{distinct}\\ \text{combinations}}} \overline{F_P\left(p_{j_1}\cdots p_{j_n}\right) p_{k_{T(1)}}\cdots p_{k_{T(K)}}}\ \overline{x_{k_{T(K+1)}}\cdots x_{k_{T(m)}}} \tag{69}$$

which follows from Eq. 55. The sum over n can start with K = n because of the orthogonality property of $F_P\left(p_{j_1}\cdots p_{j_n}\right)$. This also shows that the average is zero if $n > m$.

For large $\overline{x^2}$, we claim that an approximate formula that is closely related to Eq. 42 holds.

$$\overline{F_P\left(p_{j_1}\cdots p_{j_n}\right) u_{k_1}\cdots u_{k_m}}$$

$$\sim \sum_{\substack{\text{distinct}\\ \text{combinations}}} \overline{u_{k_{T(n+1)}}\cdots u_{k_{T(m)}}}\ \overline{F_P\left(p_{j_1}\cdots p_{j_n}\right) u_{k_{T(1)}}\cdots u_{k_{T(n)}}} \tag{70}$$

Since $\overline{F_P\left(p_{j_1}\cdots p_{j_n}\right) u_{k_{T(1)}}\cdots u_{k_{T(n)}}} = \overline{F_P\left(p_{j_1}\cdots p_{j_n}\right) p_{k_{T(1)}}\cdots p_{k_{T(n)}}}$, replacing $\overline{u_{k_{T(n+1)}}\cdots u_{k_{T(m)}}}$ by $\overline{x_{k_{T(n+1)}}\cdots x_{k_{T(m)}}}$ in Eq. 70 yields the K = n terms in Eq. 69. Hence, Eq. 70 is correct for terms proportional to $\overline{x^2}^{(m-n)/2}$, if $m - n$ is even. If $m - n$ is odd and $\bar{p} = 0$, the terms proportional to $\overline{x^2}^{(m-n-1)/2}$ are zero in Eqs. 69 and 70. Therefore, Eq. 70 is correct for terms proportional to the highest possible $\overline{x^2}$ power, just as in the other large-$\overline{x^2}$ approximations.

It is now possible to conclude that the basic polynomials in P and those in U exhibit an approximate joint orthogonality property (similar to complete orthogonality) for large $\overline{x^2}$. Specifically, we claim that the average

$$\overline{F_P\left(p_{j_1}\cdots p_{j_n}\right) F_U\left(u_{k_1}\cdots u_{k_m}\right)} \tag{71}$$

has these properties for large $\overline{x^2}$:

(a) If $n > m$, it is zero.

(b) If $n = m$ or $n = m - 1$, it is of the order of $\overline{x^2}^{0}$ (constant).

(c) If $n < m - 1$, it is of the order of $\overline{x^2}^{(m-n-2)/2}$, or less.

The first statement and the $n = m$ case of the second statement follow directly from Eq. 69. Statements (b) and (c) are true because the basic polynomials in U have been shown to behave like Hermite polynomials in averages, to an approximation that is valid for terms proportional to the highest possible power of $\overline{x^2}$, and this behavior includes approximate complete orthogonality when the set P is joined with the set U. The whole line of argument heretofore has been somewhat sketchy because clarity depends on a real understanding of the properties of Hermite polynomials as outlined in section 3.3, and not on a more detailed discussion. It is possible, but comparatively quite complicated, to prove the statements concerning Eq. 71 without reference to Hermite polynomials.

We are now able to find the order of magnitude of the coefficients $b_{j_1 \dots j_n}$ (Eq. 60) in powers of $\overline{x^2}$ for large $\overline{x^2}$. Recall that these coefficients are determined by the C_n^{n+L-1} equations of the form of Eq. 52, with $M = n$. From expression 63 and the associated discussion, the determinant of these simultaneous equations cannot be of order greater than $\overline{x^2}^{n}$ raised to the C_n^{n+L-1} power. That the order of the determinant is at least this large for large $\overline{x^2}$ can be seen by temporarily letting $p \equiv 0$. The resulting determinant has terms of only the order in question and cannot vanish, for the reasons mentioned in connection with Eq. 36. But returning to the case for $p \not\equiv 0$ does not add any terms of the order in question, and so our contention is proved.

The numerator determinant in the Cramer's rule solution of Eqs. 52 has one column consisting of the averages

$$\overline{s\ F_U\left(u_{k_1} \dots u_{k_n}\right)} = \overline{\hat{z}_N(P)\ F_U\left(u_{k_1} \dots u_{k_n}\right)} \tag{72}$$

if we assume that $N \geq n$. Thus, $\overline{x^2}^{n}$ raised to the $C_n^{n+L-1}-1$ power and multiplied by the largest of these averages gives a bound on the order of the numerator determinant. Let N_o be again the lowest degree of the nonvanishing basic polynomials in $\hat{z}_N(P)$. Then the average in Eq. 72 has these properties for large $\overline{x^2}$:

(a) If $n < N_o$, it is zero.

(b) If $n = N_o$ or $n = N_o + 1$, it is of the order of $\overline{x^2}^{0}$.

(c) If $n > N_o + 1$, it is of the order of $\overline{x^2}^{(n-N_o-2)/2}$, or less.

These properties are a consequence of the similar properties of Eq. 71. If we make use of Cramer's rule, these statements will apply to $\overline{x^2}^{n}\ b_{j_1 \dots j_n}$.

By using Eqs. 58, 60, 62, 63, and 72, we can assemble the low signal-to-noise-ratio qualitative results in compact form. The expansion of γ_o in powers of γ_i for small γ_i can then be characterized by the following statements:

(a) The N_o-degree basic polynomials in $z_N(U)$ supply $\gamma_i^{N_o}$ and higher powers of γ_i to the γ_o expansion.

(b) Any $(N_o + 1)$-degree basic polynomials in $z_N(U)$ supply $\gamma_i^{N_o+1}$ and higher powers of γ_i to γ_o.

(c) Any greater-than-$(N_o + 1)$-degree basic polynomials in $z_N(U)$ supply no lower-power terms than $\gamma_i^{N_o+2}$ to γ_o.

Notice the dependence of these statements on N_o, the lowest degree of the basic polynomials in $\hat{z}_N(P)$ and $z_N(U)$. These statements together show that γ_o behaves as $\gamma_i^{N_o}$ (except for a constant multiplier) when $\gamma_i \to 0$. Thus, the system performance at low signal-to-noise ratios is inversely related to the size of N_o, and systems with a large linear $(N_o = 1)$ component will give the best performance in this range. These statements also show that $z_N(U)$, for any $N \geq N_o$, converges in performance (and in the mean-square sense) to $z_{N_o}(U)$ as $\gamma_i \to 0$. Since $z_{N_o}(U)$ consists solely of N_o-degree basic polynomials, the optimum filter takes on a relatively simple form at low signal-to-noise ratios. Finally, these statements show that $z_{N_o+1}(U)$ is optimum when the expansion of γ_o includes just $\gamma_i^{N_o}$ and $\gamma_i^{N_o+1}$ terms. However, it is not possible to say, in general, that $z_N(U)$ is optimum when γ_o includes terms up to γ_i^N.

Suppose that we have a system like that in Fig. 1 in which the additive noise is not Gaussian. The fact that $\hat{z}_N(P)$ and $z_N(U)$ have the same lowest degree, N_o, for their nonvanishing basic polynomials is unchanged. It is also true that γ_o behaves as $\gamma_i^{N_o}$ when $\gamma_i \to 0$. The difference is that basic polynomials of degree greater than N_o can supply $\gamma_i^{N_o}$ terms to the expansion of γ_o. Hence, $z_{N_o}(U)$ may no longer be optimum near the low signal-to-noise-ratio limit; this necessitates a polynomial filter with higher-degree terms in this range. It would be interesting to explore the low signal-to-noise-ratio case in greater detail for non-Gaussian noises and attempt to relate noise characteristics to the performance and structure of the optimum filter.

3.6 EXAMPLES OF SIMPLE SYSTEMS

Consider the class of systems in which p (Fig. 1) is a zero-mean Gaussian signal. Then $u = p + x$ is Gaussian, and variables in the set $U + P = \{u_1, \ldots, u_L, p_1, \ldots, p_L\}$ have a joint Gaussian distribution. Therefore, the basic polynomials G_U in U and G_P in P are Hermite polynomials, and exhibit complete orthogonality with respect to the set U + P. As might be anticipated, several results of section 3.5 can be strengthened for the Gaussian case.

The major difference between the Gaussian and non-Gaussian cases shows up clearly in Eq. 71. If p is a Gaussian signal, the average shown there is zero unless $m = n$, by the complete orthogonality property. If $m = n$,

$$\overline{G_P\left(p_{j_1} \cdots p_{j_n}\right) G_U\left(u_{k_1} \cdots u_{k_n}\right)} = \overline{G_P\left(p_{j_1} \cdots p_{j_n}\right) u_{k_1} \cdots u_{k_n}}$$

$$= \overline{G_P\left(p_{j_1} \cdots p_{j_n}\right) p_{k_1} \cdots p_{k_n}} \tag{73}$$

which is also correct in the non-Gaussian case. Notice that this average involves none of the x_k-variables. Following essentially the same reasoning as in section 3.5, we are led to these conclusions:

(a) The M-degree basic polynomials contained in $\hat{z}_N(P)$ $(N \geq M)$, and none having other degrees, determine the coefficients of the M-degree basic polynomials in $z_N(U)$.

(b) The M-degree basic polynomials contained in $z_N(U)$ supply only γ_i^M and higher-power terms to a small-γ_i expansion of γ_o.

The first statement shows that the structures of $\hat{z}_N(P)$ and $z_N(U)$ are more closely related here than in the general case. In particular, the degrees of the nonvanishing basic polynomials in $\hat{z}_N(P)$ and in $z_N(U)$ are the same, whereas, in general, this is only true for the lowest (N_o) degree. The strongest version of the first statement, which is generally valid for non-Gaussian signals, is that no basic polynomials in $\hat{z}_N(P)$ of degree greater than M can influence the M-degree coefficients in $z_N(U)$. It has already been emphasized that the second statement is only true in general for $M = N_o$, $N_o + 1$, and $N_o + 2$ (it is trivially true for $M < N_o$).

A simple Gaussian, memoryless example illustrates one possible type of system behavior. Assuming that p is a Gaussian signal, we let

$$s = \hat{z}_N(P) = G_P(p^n)$$

where $N \geq n$, of course. From the preceding discussion, we must have

$$z_N(U) = b\, G_U(u^n) \tag{74}$$

for $N \geq n$. Let $\overline{p^2} = 1$, for simplicity, and notice that this makes proportionality 58 an equality. The formulas of section 3.3 make it easy to compute the averages

$$\overline{s^2} = n!$$

$$\overline{G_U^2(u^n)} = n!\ \overline{u^2}^{\,n}$$

$$= n!\left(1 + \overline{x^2}\right)^n$$

$$\overline{s\, G_U(u^n)} = n!$$

With the aid of Eq. 52, we can now calculate the coefficient b in Eq. 74:

$$b = \frac{1}{\left(1+\overline{x^2}\right)^n}$$

Therefore,

$$\frac{\overline{z_N^2(U)}}{\overline{s^2}} = \frac{1}{\left(1+\overline{x^2}\right)^n} = \frac{\gamma_i^n}{(1+\gamma_i)^n} \tag{75}$$

in which we have introduced $\gamma_i = 1/\overline{x^2}$. Finally, from Eqs. 21 and 28, and the binomial expansion, we have

$$\gamma_o = \frac{\gamma_i^n}{\sum\limits_{m=0}^{n-1} C_m^n \gamma_i^m} \tag{76}$$

where $C_m^n = \dfrac{n!}{m!\,(n-m)!}$. Notice that as $\gamma_i \to 0$, we obtain $\gamma_o \sim \gamma_i^n$, which agrees with the general low signal-to-noise-ratio results. A logarithmic sketch of the behavior of Eq. 76 is shown in Fig. 4. (The particular curve shown is for the case n = 2.)

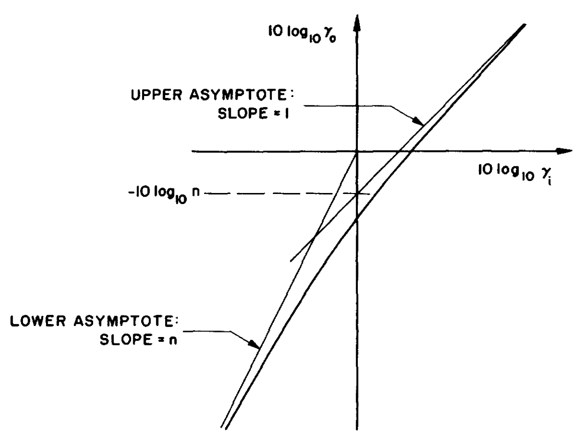

Fig. 4. Signal-to-noise-ratio curve.

It may be possible to improve the performance of a linear system by using a nonlinear filter, if the signal p is not Gaussian. The following example is interesting because the optimum filter at both high and low signal-to-noise-ratio limits is linear, but it is possible to improve midrange performance by nonlinear filtering. We again choose a memoryless case, for simplicity. Let $s = p = w^2 - 1$, where w is an auxiliary variable, having a Gaussian distribution with $\overline{w} = 0$, $\overline{w^2} = 1$. We now compute

$$z_1(U) = b_1 F_U(u)$$

$$z_2(U) = b_1 F_U(u) + b_2 F_U(u^2)$$

and compare their performance.

The calculations are routine, so we shall omit the details. The basic polynomials are

$$F_U(u) = u$$

$$F_U(u^2) = u^2 - \frac{8}{2+\overline{x^2}} u - \left(2+\overline{x^2}\right)$$

The coefficients can then be found:

$$b_1 = \frac{2}{2+\overline{x^2}}$$

$$b_2 = \frac{4\overline{x^2}}{24 + 36\overline{x^2} + 6\overline{x^2}^2 + \overline{x^2}^3}$$

Observe that $\overline{s^2} = \overline{p^2} = 2$, and $\gamma_i = 2/\overline{x^2}$. It follows that the linear filter performance is described by

$$\gamma_o = \gamma_i \tag{77}$$

as expected, since the linear filter merely changes the signal amplitude. The second-degree filter performance is given by

$$\gamma_o = \frac{1 + 7\gamma_i + 9\gamma_i^2 + 3\gamma_i^3}{1 + 3\gamma_i + 5\gamma_i^2 + 3\gamma_i^3} \gamma_i \tag{78}$$

It is clear that Eqs. 77 and 78 are practically the same for large or small values of γ_i, and that the latter is moderately larger than the former when $\gamma_i \sim 1$. The following tabulation indicates the improvement in output signal-to-noise ratio in going from the linear to the second-degree filter, for several values of input signal-to-noise ratio. The results are expressed in decibels.

$10 \log_{10} \gamma_i$	Improvement
−15 db	0.48 db
−10 db	1.22 db
− 5 db	2.18 db
0 db	2.22 db
5 db	1.26 db
10 db	0.51 db

Whether it would be possible to obtain this much improvement in a corresponding physical system is debatable because of the need for careful adjustment of the filter to match the signal and noise levels. On the other hand, there may be cases in which the potential improvement resulting from nonlinear filtering would be great enough to merit a painstaking trial.

IV. FREQUENCY-MODULATION SYSTEMS

4.1 INTRODUCTION

One of the first detailed studies of interference suppression in FM communication systems was carried out by Armstrong (17). The notion that wideband FM could be used to reduce interference at the receiver to a level below that of an AM system with the same carrier power was experimentally confirmed. Carson and Fry (18) put some of the FM theory on a more mathematical basis. In much of the early work emphasis was almost exclusively on operation at relatively high signal-to-noise ratios, partly because many important commercial applications required high-quality performance. Also, there was some indication that frequency modulation did not retain its superiority over amplitude modulation when the signal-to-noise ratio at the receiver input was low.

Almost one decade later, Rice (19, 20), Middleton (21, 22), and Stumpers (23) were leaders in analyzing the FM noise problem from a statistical viewpoint. The input noise was assumed to be additive and Gaussian. The detector in each case was an idealized discriminator sensitive only to rate-of-change of signal phase (instantaneous frequency), although Middleton also showed how to take into account amplitude variations. These studies tended to confirm the conjecture (which checks with experimental evidence) that FM only outperforms AM at relatively high input signal-to-noise ratios. Furthermore, it appeared that at low input signal-to-noise ratios the FM bandwidth should be as narrow as possible, in order to obtain optimum performance.

There have been many attempts to improve the noise performance of FM systems, some of which are mentioned in the studies already discussed. Black (24) lists a number of other important papers. More recently, Baghdady (25) has employed a novel feedback scheme in which the limiter and associated filter are included. However, the optimum performance potentiality of FM systems will remain somewhat uncertain until a more general analysis is undertaken.

In this section, the Wiener polynomial filtering previously described is used to examine optimum FM noise performance at low input signal-to-noise ratios. The results indicate that present detection methods are not optimum, but not too far from it either. It appears to be safe to affirm the advantages of AM over FM at low input signal-to-noise ratios, at least in the transmission of continuous messages.

To emphasize the importance of taking a general nonlinear approach to the FM noise filtering problem (filtering in this sense includes the detector), it should be observed that the optimum filter must make use of amplitude fluctuations in the received signal. This contrasts with the usual approach, in which some form of limiting is employed to smooth out amplitude changes. In the vector FM signal model, amplitude is proportional to vector length, and phase angle is equal to the vector angle. In Fig. 5 we have shown a transmitted signal vector, $\vec{F}$, together with two possible additive noise vectors, $\vec{N}$ and its negative. Because $\vec{N}$ and $-\vec{N}$ have equal probability of

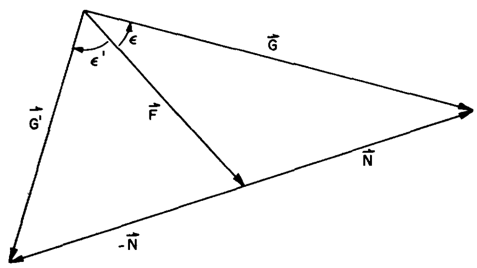

Fig. 5. FM signal model.

occurrence, so do the two resultant vectors, $\vec{G}$ and $\vec{G}'$ (for a given $\vec{F}$). Notice that the length of $\vec{G}$ is greater than the length of $\vec{G}'$, and that for the magnitudes of the phase errors, $|\epsilon| < |\epsilon'|$. We can similarly construct almost all possible resultant vectors in equally likely pairs, and in each pair the longer vector will have a smaller phase error. Hence, the amplitude of the received signal is useful in providing information about the probable size of the phase error. However, we do not claim that this additional information will usually provide more than a moderate improvement in output signal-to-noise ratio. From a practical viewpoint, the complexity of a general nonlinear FM filter may often outweigh the advantage of improved performance attained by its use.

4.2 FREQUENCY-MODULATION SPECTRUM

We shall now study briefly the spectrum of FM signals when the modulation is a Gaussian wave. Knowledge of the spectrum is useful in choosing the linear networks of Wiener filters with economy, but we shall not pursue this relationship here. Much of the material presented in this report has evolved from Wiener's (26) work; a more detailed account has appeared elsewhere (27).

We shall use the complex FM signal model,

$$g = e^{i\underline{m}\theta}$$

instead of the model with two real components, largely as a matter of convenience. The phase, θ, is assumed to have a Gaussian distribution, with $\overline{\theta} = 0$, $\overline{\theta^2} = 1$. The quantity $\underline{m}$ is then the rms phase deviation, as well as being the modulation index. Let us represent the complex FM signal model formally by the series

$$g = \sum_{n=0}^{\infty} \frac{(i\underline{m}\theta)^n}{n!} \tag{79}$$

Wiener first changed this into a series of orthogonal basic polynomials (Hermite, since θ is Gaussian). He then multiplied by the same series shifted in time, averaged, and thus found the autocorrelation function

$$R_{gg}(\tau) = \exp\left\{-\underline{m}^2\left[1-R_{\theta\theta}(\tau)\right]\right\} \tag{80}$$

where $R_{\theta\theta}(\tau)$ is the autocorrelation function of θ. It was not difficult to show that Eq. 80 splits into two parts. One part is the autocorrelation function for $\cos\theta$,

$$R_{cc}(\tau) = e^{-\underline{m}^2} \cosh \underline{m}^2 R_{\theta\theta}(\tau) \tag{81}$$

the other part is the autocorrelation function for sin θ,

$$R_{ss}(\tau) = e^{-\underline{m}^2} \sinh \underline{m}^2 R_{\theta\theta}(\tau) \tag{82}$$

Equations 81 and 82 are interesting because in practice the FM signal can be thought of as containing the two corresponding real components, but not as the complex variable of the FM signal model.

The power density spectrum can be written as a term-by-term Fourier transform of the power series expression for Eq. 80,

$$S_{gg}(f) = e^{-\underline{m}^2} \sum_{n=0}^{\infty} \frac{\underline{m}^{2n}}{n!} H_n(f) \tag{83}$$

where

$$\left.\begin{aligned} H_o(f) &= \delta(f) \\ H_1(f) &= \int_{-\infty}^{\infty} R_{\theta\theta}(\tau)\, e^{i2\pi f\tau}\, d\tau \\ H_n(f) &= \int_{-\infty}^{\infty} H_{n-1}(x)\, H_1(f-x)\, dx \end{aligned}\right\} \tag{84}$$

The even terms in Eq. 83 come from Eq. 81; the odd terms, from Eq. 82. Notice that since $H_1(f) \geq 0$ and $\int_{-\infty}^{\infty} H_1(f)\, df = 1$, $H_1(f)$ could be considered abstractly as the probability density function of some random variable, f. We shall make use of this interpretation later.

Now suppose we desire to represent $S_{gg}(f)$ by a partial sum of the series in Eq. 83. Let Δ represent the total power left out in such a sum. Since $\int_{-\infty}^{\infty} S_{gg}(f)\, df = 1$, Δ is also the fraction of the power left out. The length of the partial sum is determined by the size of Δ. For example, if $\underline{m}^2 \geq 2$, summing over those terms for which

$$(1-a)(\underline{m}^2-1) < n \leq (1+b)\underline{m}^2 \tag{85}$$

gives the bound

$$\Delta < \frac{1}{a^2\underline{m}^2} + \frac{3}{b\left(b - \frac{1}{\underline{m}^2 - 1}\right)\underline{m}^2} \tag{86}$$

as shown elsewhere (27). To hold the bound on Δ constant for large deviation $\underline{m}$, we must have, approximately

$$\left.\begin{aligned} a &\propto \frac{1}{\underline{m}} \\ b &\propto \frac{1}{\underline{m}} \end{aligned}\right\} \tag{87}$$

Therefore, the important terms in Eq. 86 are those for which $\frac{n}{\underline{m}^2} \sim 1$, if $\underline{m}$ is large. However, from Eq. 87 the number of terms with subscripts satisfying Eq. 85 is proportional to $\underline{m}$. Hence, the actual number of terms that must be included in Eq. 83 is of the order of $\underline{m}$.

We can be more definite about the form of $S_{gg}(f)$ for large values of $\underline{m}$. Recall that $H_1(f)$ can be likened to a probability density function. From Eqs. 84, $H_n(f)$ can then be interpreted as the probability density function for the sum of n independent f-variables. The central limit theorem (28) shows that $H_n(f)$ assumes a normal form as $n \to \infty$. The shape of $H_n(f)$ does not change greatly for small percentage changes in n, so

$$S_{gg}(f) \sim \frac{\exp\left(-\dfrac{f^2}{2\underline{m}^2\overline{s^2}}\right)}{\left(2\pi\underline{m}^2\overline{s^2}\right)^{1/2}} \tag{88}$$

as $\underline{m} \to \infty$, where

$$\overline{s^2} = \int_{-\infty}^{\infty} f^2\, H_1(f)\, df \tag{89}$$

and we assume that $H_1(f)$ is a continuous function. Let

$$s = \frac{1}{2\pi}\frac{d\theta}{dt} \tag{90}$$

Then it follows that $f^2\, H_1(f)$ is the power density spectrum of the message, s, which fact gives more significance to Eq. 89. Notice that $\underline{m}s$ is the instantaneous frequency of the FM signal, from Eq. 90, and that replacing f by $\underline{m}s$ in the right side of Eq. 88 gives the probability distribution of this instantaneous frequency. Therefore, Eq. 88 verifies the instantaneous-frequency concept for large $\underline{m}$ and Gaussian phase – that the power in any part of the signal spectrum is proportional to the probability (or average amount of time) that the instantaneous frequency occupies the given part of the spectrum.

Returning to the series in Eq. 79, recall that Wiener changed this into an orthogonal series before proceeding. The orthogonal series is guaranteed to converge in the mean-square sense. It is interesting to see under what conditions Eq. 79 converges in the mean-square sense. (This obviously happens if θ is a bounded variable, for then the series converges uniformly.) If we replace $\underline{m}\theta$ by w and write

$$g_N = \sum_{n=0}^{N} \frac{(iw)^n}{n!} \tag{91}$$

then g_N approaches g in the mean-square sense if and only if

$$\overline{|g - g_N|^2} \to 0 \tag{92}$$

as $N \to \infty$. An obvious necessary condition for convergence is

$$\overline{|g_N - g_{N-1}|^2} = \frac{\overline{w^{2N}}}{(N!)^2} \to 0 \tag{93}$$

as $N \to \infty$.

Next, we calculate formally

$$\overline{\left| \sum_{n=N+1}^{\infty} \frac{(iw)^n}{n!} \right|^2} = \sum_{n=N+1}^{\infty} \left[\sum_{m=0}^{n-N-1} \frac{F(m)}{(n-m)!\,(n+m)!} \right] \overline{w^{2n}} \tag{94}$$

where $F(m) = 2(-1)^m$, $m \neq 0$, and $F(0) = 1$. The bracketed sum is an alternating series bounded in magnitude by $2/(n!)^2$. Hence,

$$\overline{|g - g_N|^2} < 2 \sum_{n=N+1}^{\infty} \frac{\overline{w^{2n}}}{(n!)^2} \tag{95}$$

Inequality 95, together with Eq. 92, yields the sufficient condition for convergence

$$\sum_{n=0}^{\infty} \frac{\overline{w^{2n}}}{(n!)^2} < \infty \tag{96}$$

This is stronger than the necessary condition, Eq. 93.

For a simple example of the preceding results, let $w = x^C$, where x is a zero-mean Gaussian variable. It can be shown that the sufficient condition for convergence, Eq. 96, is satisfied if $C < 2$, or if $C = 2$ and $\overline{x^2} < 1/4$. On the other hand, the necessary condition, Eq. 93, is violated if $C > 2$, or if $C = 2$ and $\overline{x^2} > 1/4$. Hence, we did not need the orthogonal series for the preceding spectrum computation with a Gaussian phase signal ($C = 1$), and this fact has been verified by direct calculation.

4.3 POLYNOMIAL FILTERING

The FM signal model of Eq. 7 will be employed henceforth. In this case, θ_e is the phase-error angle, which might be introduced by lack of perfect phase synchronism in carrier demodulation. We shall begin with "random-phase FM," in which θ_e is assumed to be uniformly distributed over an interval of length 2π. If $\theta_e(t)$ is treated as a time signal, then we are assuming that $\theta_e(t)$ and $\theta_e(t) + c$ are equally likely, where c is a constant. From the standpoint of Eq. 7, this constant c could be limited to an interval of length 2π. Random-phase FM is a suitable model for practical systems that are phase-insensitive such as those employing a conventional discriminator. For this purpose, it is often satisfactory to let θ_e be a randomly chosen constant, rather than a random signal.

For random-phase FM, the probability density function relating the message and the received signal variables has the partially odd symmetry property

$$p(s,u,v) = p(s,-u,-v) \tag{97}$$

For, $u = \cos(\theta-\theta_e) + x$, so that keeping s (hence θ) fixed, shifting θ_e by π, and replacing x by $-x$ changes u into $-u$ without affecting any joint probabilities. The same comments apply to $v = \sin(\theta-\theta_e) + y$. This proof also applies to all higher-order joint probabilities relating s, u, and v, so that

$$p(s,U,V) = p(s,-U,-V) \tag{98}$$

where U and V are the set of outputs of the linear networks in a Wiener filter (Fig. 2) as defined in Eqs. 12. If we average out s in Eq. 98, $p(U,V) = p(-U,-V)$. From the discussion of the symmetry property for any set of variables, U, it follows that the basic polynomials in U and V have all odd-degree or all even-degree terms, just as the Hermite polynomials. With the aid of Eq. 98, we find that the average of s times an odd-degree basic polynomial is always zero. It follows that $z_N(U,V)$, the least-squares polynomial part of the filter, contains only even-degree basic polynomials.

We are now in a position to obtain some qualitative low signal-to-noise-ratio results for random-phase FM. Since $z_N(U,V)$ contains only even-degree basic polynomials, it has none of the first degree. If we assume that not all second-degree polynomials vanish, it follows (see section 3.5) that

$$\gamma_o \sim C\gamma_i^2 \tag{99}$$

as $\gamma_i \to 0$. On a logarithmic plot, Eq. 99 would have an asymptotic slope of 2 as $\gamma_i \to 0$; this agrees with known theoretical results (19-23), as well as many experimental determinations. Hence, we may infer that random-phase FM is a realistic model for conventional FM systems, and that such conventional systems are not grossly inferior to the optimum system at low input signal-to-noise ratios, if we assume that the system is not phase-synchronous.

For comparison purposes, observe that if θ_e is not translation-invariant (the frequency modulation is not random-phase), $z_N(U,V)$ may contain first-degree polynomials. If this is so, we have shown in section 3.5 that

$$\gamma_o \sim C'\gamma_i \tag{100}$$

for optimum filtering, as $\gamma_i \to 0$. Such performance is akin to that of an AM system, although C' could be quite small. No matter how small C' is, Eq. 100 can exceed

Eq. 99 by an arbitrarily large ratio, if γ_i is sufficiently small. Of course, if C′ is too small in relation to C, the improvement attained by fully utilizing phase information may not be significant at signal-to-noise ratios that are large enough to be useful for communication purposes.

It is helpful to introduce some simplifying assumptions, which will be in effect henceforth. These simplifications are not needed to produce tractable problems. First, we require symmetry of the message distribution,

$$p(s) = p(-s) \tag{101}$$

and this is assumed also to hold for s as a random function (replacing s by −s does not change any higher-order probabilities). We assume odd symmetry for the phase functional

$$\theta[-s] = -\theta[s] \tag{102}$$

Notice that Eq. 102 is automatically true if θ is a linear functional. Let the phase error, θ_e, have a symmetrical distribution

$$p(-\theta_e) = p(\theta_e) \tag{103}$$

and this holds also for θ_e as a random function. The linear part of the Wiener filter is assumed to be a tapped delay line, so that U and V are sets of samples of u and v, respectively. We choose sampling because it implies that

$$\left.\begin{aligned} p_n &= \cos(\theta_n - \theta_{en}) \\ q_n &= \sin(\theta_n - \theta_{en}) \end{aligned}\right\} \tag{104}$$

where $\theta_n - \theta_{en}$ is the corresponding sample of the phase angle. Finally, we let the noise samples be independent, which is equivalent to linear independence for the Gaussian noise,

$$\left.\begin{aligned} \overline{x_m x_n} &= 0 \\ \overline{y_m y_n} &= 0 \end{aligned}\right\} \tag{105}$$

if $m \neq n$. Such independence would be true, for example, in uniform sampling of a rectangular noise band at the Nyquist rate, so this assumption is not completely artificial.

In order to study the effect of the preceding assumptions on the class of problems that can be handled, we observe that

$$p(s, u, v) = p(-s, u, -v) \tag{106}$$

For, changing s (hence θ), θ_e, and y into their negatives does not change any joint probabilities (see Eqs. 101-103); it leaves u unchanged, and converts v into −v. This proof also shows that

$$p(s, U, V) = p(-s, U, -V) \tag{107}$$

If we average out s in Eq. 107, $p(U, V) = p(U, -V)$. Thus, the average of a product of U- and V-variables will be zero if an odd number of V-variables is present. This means that the basic polynomials in U and V can be divided into two classes, in one of which all terms in each polynomial have an odd number of V-variables, and in the other, only terms with even numbers of V-variables appear. Furthermore, the basic polynomials from opposite classes are automatically orthogonal, even if they are of the same degree. With the aid of Eq. 107, the average of s multiplied by a basic polynomial from the even-number-of-V-variables class is zero. It follows that only basic polynomials from the odd-number-of-V-variables class can appear in $z_N(U, V)$.

Let us return to random-phase FM. Combining the simplifying assumptions with the earlier results, we can make the following statements.

(a) The nonvanishing basic polynomials in $z_N(U, V)$ (for least-squares filtering) are of even degree, and each term has an odd number of V-variables (hence, also an odd number of U-variables).

Next, suppose that $\pi/2$ is added to θ_e, s is replaced by $-s$ (hence θ by $-\theta$), x by y, and y by $-x$. No probabilities are affected, but U-variables are changed into the corresponding V-variables, and V-variables into $-$U-variables. From Eq. 107 and the fact that all basic polynomials under consideration have odd numbers of U- and V-variables, the averaged product of such a transformed polynomial with s is unchanged in magnitude and sign. This leads to a second observation about random-phase FM filters.

(b) The terms making up the nonvanishing basic polynomials contained in $z_N(U, V)$ can be written in pairs, the second term of each pair being formed from the first by interchanging U- and V-variables and adding a minus sign.

Property (b) is useful because it cuts in half the number of independent coefficients that must be found in constructing $z_N(U, V)$. For example, let us consider some possible second-degree basic polynomials that might appear in $z_N(U, V)$. From the first property, we cannot have $F_{U,V}(u_m u_n)$ because the leading term has an even number of U-variables. From the second property, we can use $F_{U,V}(u_m v_n - u_n v_m)$ in place of $F_{U,V}(u_m v_n)$ and $F_{U,V}(u_n v_m)$. Also, $F_{U,V}(u_m v_m - u_m v_m) = 0$; this shows that $F_{U,V}(u_m v_m)$ does not appear. Notice that

$$F_{U,V}(u_m v_n - u_n v_m) = u_m v_n - u_n v_m \tag{108}$$

because $\overline{u_m v_n} = 0$, and orthogonality to first-degree terms is automatic.

It is relatively easy to calculate the low signal-to-noise-ratio least-squares filter for random-phase FM, under the simplifying assumptions made previously. The important average is

$$\overline{s(u_m v_n - u_n v_m)} = \overline{s \sin(\theta_n - \theta_m)}\ \overline{\cos(\theta_{en} - \theta_{em})} \tag{109}$$

in which we have used all of the given symmetry properties. Under the further simplification that θ_e is a randomly chosen constant, $\overline{\cos(\theta_{en}-\theta_{em})} = 1$. Then,

$$z_N(U,V) \sim 2\gamma_i^2 \sum_{\substack{m=1 \\ n>m}}^{L} \overline{s \sin(\theta_n-\theta_m)}\,(u_m v_n - u_n v_m) \tag{110}$$

as $\gamma_i \to 0$. The coefficient $2\gamma_i^2$ can be discarded for maximum signal-to-noise-ratio filtering, and thus the filter will be independent of γ_i (except that it must be small). This independence of the filter from input signal-to-noise ratio does not usually occur in nonlinear filtering, but it is certainly a great practical help when it does. Another feature of Eq. 110 is that the sum roughly models a lowpass filter such as the one following the discriminator in a conventional FM receiver. The precise nature of the linear filtering represented by this sum depends, of course, on the coefficients $\overline{s \sin(\theta_n-\theta_m)}$.

For further study of Eq. 110, we introduce the polar-coordinate variables r and ϕ:

$$\left.\begin{aligned} u &= r \cos\phi \\ v &= r \sin\phi \end{aligned}\right\} \tag{111}$$

which are especially appropriate to a vector representation, as in Fig. 5. Because we have employed sampling,

$$\left.\begin{aligned} u_n &= r_n \cos\phi_n \\ v_n &= r_n \sin\phi_n \end{aligned}\right\} \tag{112}$$

Polar-coordinate samples are designated by the usual set notation:

$$\left.\begin{aligned} R &= \{r_1, \dots, r_L\} \\ \Phi &= \{\phi_1, \dots, \phi_L\} \end{aligned}\right\} \tag{113}$$

After transformation by Eq. 112, Eq. 110 becomes

$$z_N(R,\Phi) \sim 2\gamma_i^2 \sum_{\substack{m=1 \\ n>m}}^{L} \overline{s \sin(\theta_n-\theta_m)}\, r_m r_n \sin(\phi_n-\phi_m) \tag{114}$$

Notice that only differences between phase variables appear, for both the transmitted phase, θ, and the received phase, ϕ. This suggests differentiation, at least for samples spaced close to each other in time. Hence, it can be argued that the natural modulation for random-phase frequency modulation is to make the instantaneous frequency, $\frac{d\theta}{dt}$, proportional to the message, s. This is not a conclusive argument, by any means, but it is in fairly good agreement with standard practice in FM systems. Another interesting feature of Eq. 114 is the appearance of $r_m r_n$ as a weighting factor that gives more emphasis to the large-r samples. This agrees qualitatively with the discussion of Fig. 5. We found then that larger r can imply smaller probable phase error, $\phi - \theta$.

For random-phase FM, the output signal-to-noise ratio has already been found to be of the form $\gamma_o \sim C\gamma_i^2$ as $\gamma_i \to 0$ with optimum nonlinear filtering. From Eqs. 110, 28, and 22, we find that

$$C \sim \frac{2}{\overline{s^2}} \sum_{\substack{m=1 \\ n>m}}^{L} \overline{s \sin(\theta_n - \theta_m)}^2 \tag{115}$$

as $\gamma_i \to 0$. We can bound Eq. 115 by assuming that $\sin(\theta_n - \theta_m) = \pm 1$ when $s = \pm s_o$ (a constant). The result is

$$C \leq L(L+1) \tag{116}$$

Because the phase differences in Eq. 115 are not all independent, a closer bound for C is around $L^2/2$, at least for large L. When L = 2, the bound in Eq. 116 is clearly attainable.

Next, let us suppose that θ_e is not uniformly distributed. Then the optimum filter at the low signal-to-noise-ratio limit is usually linear. Recall that because of the symmetry conditions (Eqs. 101-103), all of the U-variables drop out (terms must have an odd number of V-variables). The important average is

$$\overline{s\, v_n} = \overline{s \sin \theta_n}\; \overline{\cos \theta_{en}} \tag{117}$$

If we pick the most favorable case, in which there is no phase-error signal $(\theta_e \equiv 0)$, then $\overline{\cos \theta_{en}} = 1$. In this case,

$$z_N(U, V) \sim 2\gamma_i \sum_{n=1}^{L} \overline{s \sin \theta_n}\, v_n \tag{118}$$

as $\gamma_i \to 0$. The coefficient $2\gamma_i$ can be discarded without changing the output signal-to-noise ratio; this makes the filter independent of γ_i, as usual for a linear filter. It might be expected that the coefficients $\overline{s \sin \theta_n}$ usually turn out so as to make the filter essentially lowpass. It is interesting to note that for narrow-band frequency modulation, a linear filter like that of Eq. 118, operating on the quadrature signal, v, would be close to optimum for all signal-to-noise ratios, not just for the lowest.

Transforming Eq. 118 by means of Eq. 112 gives

$$z_N(R, \Phi) \sim 2\gamma_i \sum_{n=1}^{L} \overline{s \sin \theta_n}\, r_n \sin \phi_n \tag{119}$$

Here, the transmitted and received phase variables appear singly; this fact suggests that the appropriate frequency modulation in this situation is phase modulation, with θ proportional to s. Notice the appearance of r_n as a weighting factor. Again, we have agreement with the probability relationship between the magnitudes of r and $\phi - \theta$.

The output signal-to-noise ratio for optimum nonrandom-phase FM performance is of the form $\gamma_o \sim C'\gamma_i$ as $\gamma_i \to 0$, from Eq. 100. With the use of Eq. 118, we find that

$$C' \sim \frac{2}{s^2} \sum_{n=1}^{L} \overline{s \sin \theta_n}^2 \tag{120}$$

as $\gamma_i \to 0$. We can obtain a bound, $C' \leq 2L$, by letting $\sin \theta_n = \pm 1$ when $s = \pm s_o$. This bound is clearly attainable for any L by a binary phase-reversal system, which is really equivalent to a phase-locked binary PCM system with suppressed carrier.

For wideband frequency modulation, the required transmission bandwidth, W, is proportional to the modulation index, $\underline{m}$; that is, $W \propto \underline{m}$. If the interfering noise is white, varying the bandwidth without changing transmitter power means that $\gamma_i \propto \frac{1}{W}$. Under these conditions, we have the well-known relation between output signal-to-noise ratio and modulation index

$$\gamma_o \propto \underline{m}^2 \tag{121}$$

as long as γ_i remains large enough so that noise peaks are almost always less than the peak signal amplitude (29).

Now, suppose that γ_i is small. Observe that the necessary number of samples, L, for achieving a good signal representation is proportional to the bandwidth; that is, $L \propto W$. Also, consider the average

$$\overline{s \sin \underline{m}\theta_n} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s \sin \underline{m}\theta_n \, p(s, \theta_n) \, ds \, d\theta_n \tag{122}$$

If the integrand on the right side of Eq. 122 is an integral (a smoothness condition that will almost always be satisfied in practice for continuous systems), a special form of the Riemann-Lebesgue theorem (30) shows that

$$\overline{s \sin \underline{m}\theta_n} \propto \frac{1}{\underline{m}}$$

or smaller, for large $\underline{m}$. With the use of the results derived after Eq. 120, we can calculate that

$$\gamma_o \propto \frac{1}{\underline{m}^2} \tag{123}$$

or smaller, if $\underline{m}$ is large and $\gamma_i \to 0$. Practically the same proof shows that Eq. 123 holds for random-phase FM, or for any sort of phase-error signal, θ_e. Notice that this is just the inverse of the relation in Eq. 121. Hence, narrow-band frequency modulation generally outperforms wideband frequency modulation at low input signal-to-noise ratios; this is just the reverse of the high-input signal-to-noise ratio situation. Comparison with AM performance can be made by recalling the similarity between phase-synchronous narrow-band FM and AM.

It does not appear profitable to use Wiener polynomial filtering for FM signals at high input signal-to-noise ratios, since a high-degree polynomial must be computed in order to achieve a good filter. Moreover, conventional techniques can be used to produce nearly optimum results in this range. The complexity is also great at intermediate input signal-to-noise ratios, but here it seems certain that conventional techniques are not optimum (review the argument about using signal amplitude and refer to Fig. 5). We leave for future speculation the question of whether or not a simple nonlinear correction filter might be designed to make a worth-while improvement in the performance of a conventional FM discriminator.

4.4 EXAMPLES OF FM SYSTEM PERFORMANCE

In order to develop some feeling for the averages used in obtaining low signal-to-noise-ratio performance of FM systems with optimum filtering, we shall compute the quantity

$$\overline{C} = \frac{\overline{s \sin \underline{m}s}^{\,2}}{\overline{s^2}} \tag{124}$$

for several distributions of s. Similar averages appear in Eq. 115 if $\theta_n - \theta_m \propto s$, or in Eq. 120 if $\theta_n \propto s$. Notice that the bound $\overline{C} \leq 1$ is attainable for a binary distribution (for example, let $s = \pm 1$ and $\underline{m} = \pi/2$).

First, let s have the rectangular distribution

$$p(s) = \begin{cases} \frac{1}{2}, & |s| \leq 1 \\ 0, & |s| > 1 \end{cases} \tag{125}$$

From Eq. 124, we can compute $\overline{C}$. The results are:

$\underline{m}$	$\overline{C}$
0.500π	0.49
0.662π	0.57 (max)
1.000π	0.30

The middle value is the maximum for $\overline{C}$. Since system performance increases with $\overline{C}$, these results give some indication of the importance of properly adjusting the modulation index, $\underline{m}$. Also, note that the maximum value of $\overline{C}$ for this distribution is not much over half that attainable for a binary message. This reflects the fact that a discrete signal usually has better signal-to-noise-ratio performance than a continuous signal.

Next, we try the triangular distribution

$$p(s) = \begin{cases} 1 - |s|, & |s| \leq 1 \\ 0, & |s| > 1 \end{cases} \tag{126}$$

We compute again and the results are:

$\underline{m}$	$\overline{C}$
0.500π	0.29
0.833π	0.44 (max)
1.000π	0.40

The fact that this maximum for $\overline{C}$ is lower than before may be explained roughly by observing that a rectangular distribution is closer to being binary (the optimum) than is a triangular distribution. Also, remember that we are mapping the message, s, into a function, sin $\underline{m}$s, that has a multivalued inverse if the range of $\underline{m}$s exceeds π. Hence, distributions that are more spread out for a given variance might be expected to provide lower performance.

In the light of the preceding argument, it is interesting to compute $\overline{C}$ for an unbounded message. Therefore, we assume that s is normally distributed.

$$p(s) = \frac{e^{-s^2/2}}{(2\pi)^{1/2}}$$

With some effort we can compute the average

$$\overline{s \sin \underline{m}s} = \underline{m}\, e^{-\underline{m}^2/2}$$

Therefore, when $\underline{m} = 1$,

$$\overline{C}_{max} = \frac{1}{e} = 0.37$$

As expected, this maximum is lower than those already computed, largely because the extensive tails of the distribution are mapped many-to-one in the modulation process.

It is worth while to consider a simple example of lowpass behavior for the linear filter of Eq. 118 (used for optimum low signal-to-noise-ratio filtering of quadrature signal v when $\theta_e \equiv 0$). If we discard the coefficient $2\gamma_i$, and define

$$a_n = \overline{s \sin \theta_n}$$

the filter impulse response can be written in the form

$$h(t) = \sum_{n=1}^{L} a_n \delta(t - \tau_n)$$

where the τ_n are the sampling times. The power density spectrum for this filter is readily found.

$$H(f) = \sum_{\substack{m=1 \\ n=1}}^{L} a_m a_n \cos 2\pi(\tau_m - \tau_n) f$$

Assume the spectrum of signal v has no important frequency components above W, so that we need only investigate H(f) for $f < W$. Such a restriction is usually needed for a filter constructed from a finite number of all-pass elements, of course.

For our particular example,

$$\tau_{n+1} - \tau_n = \frac{1}{2W}$$

for all values of n, which implies uniform sampling at the Nyquist rate. Then, we assume that all filter coefficients are equal. Admittedly, this is a somewhat artificial situation, but the results are not unrealistic. We let $a_n = 1/L$ for all values of n, so that $H(0) = 1$. The resulting filter spectrum is

$$H(f) = \frac{1}{L} + \frac{2}{L} \sum_{n=1}^{L-1} \left(1 - \frac{n}{L}\right) \cos n\pi \frac{f}{W} \tag{127}$$

Although we shall not go into details here, it can be shown that if $L > 1$,

$$H\left(\frac{2W}{L}\right) = 0$$

is the lowest-frequency zero of H(f), and most of the area under H(f) is within the band marked by this zero. Since it is reasonable to assume that the message bandwidth is proportional to W/L, this result appears to be qualitatively correct. To help visualize these spectra, we have sketched Eq. 127 for L = 1, 2, and 4 in Fig. 6.

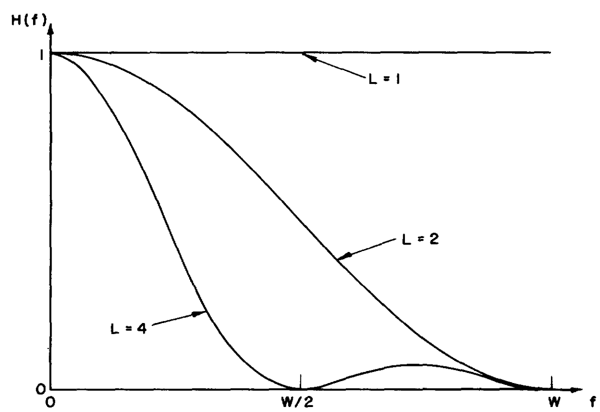

Fig. 6. Filter spectra.

4.5 DIRECT STATISTICAL APPROACH

In this section, the least-squares FM filter is of the form shown in Fig. 2. However, the nonlinear function, z(U, V), is represented directly as a conditional mean, not constructed of orthogonal basic polynomials. We make all the simplifying assumptions of Eqs. 101-104. The polar coordinates defined in Eqs. 112 are also used.

First, let us consider random-phase FM, with the simplification that θ_e is a randomly chosen constant. Straightforward (but somewhat lengthy) computation gives the conditional mean (Eq. 24):

$$z_{\infty}(R,\Phi) = \frac{\overline{s\, I_o\left\{2\gamma_i\left[\sum_{\substack{m=1\\n=1}}^{L} r_m r_n \cos(\phi_m-\phi_n-\theta_m+\theta_n)\right]^{1/2}\right\}}^{s,\theta}}{\overline{I_o\left\{2\gamma_i\left[\sum_{\substack{m=1\\n=1}}^{L} r_m r_n \cos(\phi_m-\phi_n-\theta_m+\theta_n)\right]^{1/2}\right\}}^{\theta}} \tag{128}$$

in which $\overline{X}^{s,\theta}$ means "average X over s and the θ_n-variables"; $\overline{X}^{\theta}$ means "average X over the θ_n-variables"; and $I_o(x) = J_o(ix)$ is the zero-order modified Bessel function of the first kind. Note that only differences between pairs of phase angles appear, as before, thereby suggesting differentiation and the importance of instantaneous frequency. Samples of the vector length (signal amplitude), r, again show up as weighting factors, in accordance with their relation to phase-angle errors.

With the aid of the power series (Jahnke and Emde (31)),

$$I_o(x) = \sum_{n=0}^{\infty}\left[\frac{\left(\frac{x}{2}\right)^n}{n!}\right]^2$$

Eq. 128 may be written as the ratio of two series. Notice, by the way, that only even powers of γ_i appear in these series. As $\gamma_i \to 0$, the denominator series converges in the mean-square sense to 1 (its first term), and the numerator to a first term that is identical to the second-degree polynomial of Eq. 114. This exact agreement in form with the orthogonal polynomial representation ceases, however, after the first term because the two expansions are built up in different ways. Nevertheless, it is encouraging to find this readily demonstrable similarity of the filters near the low signal-to-noise-ratio limit.

Next, we exhibit the conditional mean filter for the case in which there is no phase-error signal ($\theta_e \equiv 0$).

$$z_{\infty}(R,\Phi) = \frac{\overline{s\, \exp\left[2\gamma_i \sum_{n=1}^{L} r_n \cos(\phi_n-\theta_n)\right]}^{s,\theta}}{\overline{\exp\left[2\gamma_i \sum_{n=1}^{L} r_n \cos(\phi_n-\theta_n)\right]}^{\theta}} \tag{129}$$

The phase-angle variables appear singly this time; this suggests the importance of phase modulation, as in the polynomial approach. The weighting by r-samples is still present. We can again expand the numerator and denominator in power series, and as $\gamma_i \to 0$ the result converges in the mean-square sense to the first-degree polynomial of Eq. 119. Actually, the low signal-to-noise-ratio filters for polynomial and conditional-mean approaches agree in form—no matter what distribution is chosen for θ_e, and even if other simplifying conditions are removed.

To increase our familiarity with the use of conditional means, we shall compute two simple low signal-to-noise-ratio examples. Assume that the filter cannot make use of amplitude information, as in a ratio detector. We begin with the conditional probability density function

$$p(R, \Phi | \theta, \theta_e) = \prod_{n=1}^{L} \frac{\gamma_i}{\pi} r_n \exp\left\{-\gamma_i \left[\sin^2 (\phi_n - \theta_n - \theta_{en}) + \left(r_n - \cos (\phi_n - \theta_n - \theta_{en})\right)^2\right]\right\} \tag{130}$$

obtained by using the Gaussian noise distribution. If we form the series expansion for the exponential and integrate out the R-variables, Eq. 130 becomes

$$p(\Phi | \theta, \theta_e) = \left(\frac{1}{2\pi}\right)^L + \left(\frac{1}{2\pi}\right)^{L-1} \frac{1}{2}\left(\frac{\gamma_i}{\pi}\right)^{1/2} \sum_{n=1}^{L} \cos (\phi_n - \theta_n - \theta_{en})$$

$$+ \left(\frac{1}{2\pi}\right)^{L-2} \frac{\gamma_i}{4\pi} \sum_{\substack{m=1 \\ n>m}}^{L} \cos (\phi_m - \theta_m - \theta_{em}) \cos (\phi_n - \theta_n - \theta_{en}) + \ldots \tag{131}$$

for the smallest three terms in powers of γ_i.

First, we consider a random-phase case with $L = 2$, $\underline{m}s = \theta_2 - \theta_1$, and $\theta_{e1} = \theta_{e2} = \theta_e$. Integrating over θ_e in Eq. 131, we obtain

$$p(\phi_1, \phi_2 | s) \sim \frac{1}{4\pi^2} + \frac{\gamma_i}{8\pi} \cos (\phi_2 - \phi_1 - \underline{m}s)$$

as $\gamma_i \to 0$. Let s have the rectangular distribution of Eq. 125. In the simple case with $\underline{m} = \pi$, the conditional-mean filter can be computed. The result is

$$z(\phi_1, \phi_2) \sim \frac{\gamma_i}{2} \sin (\phi_2 - \phi_1)$$

as $\gamma_i \to 0$. The standard low signal-to-noise-ratio computation for this filter then gives

$$\gamma_o \sim 0.38\, \gamma_i^2 \tag{132}$$

With the aid of Eqs. 115 and 124 and the comparable result for $\overline{C}$, the coefficient of γ_i^2 in Eq. 132 must be compared with the value 0.60 that is obtained when the optimum filter can make use of amplitude r. Hence, the performance is improved noticeably by using r, as we have claimed from the study of Fig. 5.

We can make a similar comparison when $\theta_e \equiv 0$. If $L = 1$ and $\underline{m}s = \theta_1$ (let $\phi_1 = \phi$), the first two terms of Eq. 131 become

$$p(\phi|s) \sim \frac{1}{2\pi} + \frac{1}{2}\left(\frac{\gamma_i}{\pi}\right)^{1/2} \cos(\phi-\underline{m}s)$$

as $\gamma_i \to 0$. Again, if we let s have the rectangular distribution of Eq. 125 and take $\underline{m} = \pi$, the conditional-mean filter is represented by

$$z(\phi) \sim \left(\frac{\gamma_i}{\pi}\right)^{1/2} \sin\phi$$

as $\gamma_i \to 0$. Finally, for this filter,

$$\gamma_o \sim 0.48\ \gamma_i \tag{133}$$

Compare the coefficient of γ_i in Eq. 133 with the value 0.60 that is obtained when r is available for use in the optimum filter. We might speculate that even greater differences in performance may occur between optimum filters that do and do not use r, if the input signal-to-noise ratio is intermediate (say, around 0 db). However, it would probably be just about as easy (or difficult) to use the polynomial approach for such a comparison, and we leave the question open for future exploration.

Acknowledgment

The author is indebted to Professor Y. W. Lee for guidance and encouragement during his thesis research and the preparation of this report. Thanks are also due to his fellow-students, particularly to K. L. Jordan, Jr., M. Schetzen, D. W. Tufts, and C. E. Wernlein, Jr., for helpful discussions and advice.

References

1. Y. W. Lee, Statistical Theory of Communication (John Wiley and Sons, Inc., New York, 1960).
2. N. Wiener, Nonlinear Problems in Random Theory (The Technology Press of the Massachusetts Institute of Technology and John Wiley and Sons, Inc., New York, 1958).
3. J. F. Barrett, The use of functionals in the analysis of non-linear physical systems, Statistical Advisory Unit Report 1/57, Ministry of Supply, Great Britain, 1957.
4. W. B. Davenport, Jr. and W. L. Root, An Introduction to the Theory of Random Signals and Noise (McGraw-Hill Book Company, Inc., New York, 1958); see especially pp. 67-68.
5. Ibid., pp. 158-165.
6. A. G. Bose, A Theory of Nonlinear Systems, Technical Report 309, Research Laboratory of Electronics, M. I. T., May 15, 1956; see especially pp. 2-4.
7. E. C. Titchmarsh, The Theory of Functions (Oxford University Press, London, 2d edition, 1939); see p. 414.
8. A. D. Hause, Maximum signal-to-noise ratio estimation, Quarterly Progress Report No. 54, Research Laboratory of Electronics, M. I. T., July 15, 1959, pp. 108-113.
9. A. G. Bose, op. cit., pp. 31-33.
10. W. B. Davenport, Jr. and W. L. Root, op. cit., p. 145.
11. W. Feller, An Introduction to Probability Theory and Its Applications, Vol. 1 (John Wiley and Sons, Inc., New York, 2d edition, 1957); see p. 166.
12. F. B. Hildebrand, Methods of Applied Mathematics (Prentice-Hall, Inc., Englewood Cliffs, N. J., 1952); see pp. 34-35.
13. Ibid., pp. 12, 25.
14. W. B. Davenport, Jr. and W. L. Root, op. cit., pp. 50-53.
15. Ibid., p. 153
16. A. G. Bose, op. cit., pp. 25-26.
17. E. H. Armstrong, A method of reducing disturbances in radio signaling by a system of frequency modulation, Proc. IRE 24, 689-740 (1936).
18. J. R. Carson and T. C. Fry, Variable frequency electric circuit theory with application to the theory of frequency modulation, Bell System Tech. J. 16, 513-540 (1937).
19. S. O. Rice, Mathematical analysis of random noise, Bell System Tech. J. 23, 282-332 (1944); 24, 46-156 (1945).
20. S. O. Rice, Statistical properties of a sine wave plus random noise, Bell System Tech. J. 27, 109-157 (1948).
21. D. Middleton, On theoretical signal-to-noise ratios in FM receivers: A comparison with amplitude modulation, J. Appl. Phys. 20, 334-351 (1949).
22. D. Middleton, The spectrum of frequency-modulated waves after reception in random noise, Quart. Appl. Math. 7, 129-174 (1949).
23. F. L. H. M. Stumpers, Theory of FM noise, Proc. IRE 36, 1081-1092 (1948).
24. H. S. Black, Modulation Theory (D. Van Nostrand Company, Inc., Princeton, N. J., 1953); see especially pp. 215-216; 233-234.
25. E. J. Baghdady, Theory of feedback around the limiter, IRE National Convention

Record, Part 8, 1957, pp. 176-202.

26. N. Wiener, op. cit., pp. 49-62.

27. A. D. Hause, FM spectra for Gaussian messages, Quarterly Progress Report No. 55, Research Laboratory of Electronics, M.I.T., Oct. 15, 1959, pp. 86-93.

28. W. Feller, op. cit., p. 229.

29. H. S. Black, op. cit., p. 224.

30. E. C. Titchmarsh, op. cit., p. 403.

31. E. Jahnke and F. Emde, Tables of Functions with Formulae and Curves (Dover Publications, New York, 4th edition, 1945), p. 128.