Three-Frequency Nonlinear Heterodyne Detection. 2: Digital Communications and Pulsed Radar

Malvin Carl Teich and Rainfield Y. Yen

Part 1 of this paper [Appl. Opt. 14, 666 (1975)] dealt with the cw radar and analog communications uses of three-frequency nonlinear heterodyne detection. In this paper, we evaluate the technique for a number of specific pulsed radar and digital communications applications. Both the vacuum channel and the lognormal turbulent atmospheric channel are considered. It is found that the advantages of the technique in the pulsed/digital system are similar to those obtained in the cw/analog system. Computer generated error probability curves as a function of the input signal-to-noise ratio are presented for a variety of binary receiver parameters and configurations and for various levels of atmospheric turbulence. Orthogonal and nonorthogonal signaling schemes, as well as dependent and independent fading, are considered. When Doppler information is poor, performance is generally superior to that of the conventional heterodyne system.

I. Introduction

Part 1 of this paper[1] was primarily concerned with the behavior of the three-frequency nonlinear heterodyne system for applications in cw radar and analog communications. As such, a determination of the output signal-to-noise ratio $(\mathrm{SNR})_o$ was sufficient to characterize the system. In Part 2 we investigate applications in synchronous digital communications and pulsed radar and therefore examine system performance in terms of the error probability P_e. Evaluation of the probability of error under various conditions requires a decision criterion as well as a knowledge of the signal statistics; we now investigate operation of the three-frequency nonlinear heterodyne scheme in the time domain rather than in the frequency domain.

Because of the added complexity of dealing in the time domain, we limit our investigation to sinewave signals, Gaussian local oscillator (LO) noise, and envelope detection. The configuration of such a receiver is therefore similar to that considered previously in Part 1, with the addition of an envelope detector (see Figs. 1 and 13 of Ref. 1). We therefore examine the case of a particular square-law envelope detector, consisting of a square-law device, a narrowband filter, and an envelope detector.[2] Although envelope detection is generally suboptimum because it is insensitive to phase, it is easy to implement practically and is therefore widely used.[3]

We begin with an investigation of binary communications and pulsed radar for both nonorthogonal and orthogonal signaling formats in the vacuum channel. We then examine envelope probability distributions and binary signaling for sinewave signals in the lognormal channel (clear air turbulent atmosphere). The advantages of the three-frequency nonlinear heterodyne scheme in the digital communications/pulsed radar configuration are similar to those cited in Part 1 for cw radar/analog communications.

II. Envelope Probability Distributions for Sinewave Signals Plus Gaussian Noise (Vacuum Channel)

We assume here, as in Part 1, that the fields incident on the mixer are parallel, plane-polarized, and spatially first-order coherent over the detector aperture. In general, therefore, the input to the square-law device, as previously [see Eq. (19), Part 1], will be two narrowband signals plus white Gaussian noise with zero mean resulting from the LO, over the band $(0, f_n)$. Thus

$$s(t) = A_a \cos(\omega_a t + \phi_a) + A_b \cos(\omega_b t + \phi_b), \quad (1)$$

with A_a, A_b, ϕ_a, and ϕ_b stochastic processes. The amplitudes are assumed to be independent of the phases. In this section, we treat the specific case of sinusoidal signals, i.e., A_a and A_b constant and ϕ_a, ϕ_b independent random variables uniformly distributed over $(0,2\pi)$.

M. C. Teich is with the Department of Electrical Engineering & Computer Science, Columbia University, New York, New York 10027 and is a Fellow of the John Simon Guggenheim Memorial Foundation. R. Y. Yen was with Columbia University; he is now with Columbia Research Corporation, Gaithersburg, Maryland 20760.

Received 17 May 1974.

In the time domain, the white Gaussian noise, which arises from the LO, can be expressed as[4]

$$n(t) = \sum_{k=1}^{\infty} u_k \cos\omega_k t + \sum_{k=1}^{\infty} v_k \sin\omega_k t. \quad (2)$$

Here, $\omega_k = k\omega_0$ with $\omega_0 = 2\pi/2T$. If the input signal is a pulse, the pulse duration is the time interval $(-T,T)$. The coefficients u_k and v_k may therefore be written as

$$u_k = \frac{1}{T}\int_{-T}^{T} n(t) \cos\omega_k t\, dt, \quad (3a)$$

and

$$v_k = \frac{1}{T}\int_{-T}^{T} n(t) \sin\omega_k t\, dt. \quad (3b)$$

Since u_k and v_k are linear transformations of the Gaussian random variable $n(t)$, they are also Gaussian random variables;[5] furthermore it can be shown that for T large, all u_k's and v_k's are uncorrelated and independent of one another.[6] Since the mean of $n(t)$ is taken to be zero, we find

$$\langle u_k \rangle = \left\langle \frac{1}{T}\int_{-T}^{T} n(t) \cos\omega_k t\, dt \right\rangle = 0, \quad (4a)$$

and similarly

$$\langle v_k \rangle = 0, \quad (4b)$$

while the variance $\langle u_k{}^2 \rangle$ is given by

$$\langle u_k{}^2 \rangle = \frac{1}{T^2}\int_{-T}^{T}\int_{-T}^{T} \langle n(t)n(t') \rangle \cos\omega_k t \cos\omega_k t'\, dt dt' = \frac{N}{T}. \quad (5a)$$

Similarly,

$$\langle v_k{}^2 \rangle = N/T. \quad (5b)$$

In calculating these quantities, we have assumed that the Gaussian noise $n(t)$ is stationary, and that the band $[f_i,f_n]$ is sufficiently large so the noise can be approximated to be completely white (over an infinite band) leading to an autocorrelation function $R_n(t - t') \simeq N\delta(t - t')$. Here N is the height of the white noise spectrum.

The input $x(t)$ to the square-law device can now be written as

$$x(t) = s(t) + n(t) = A_a \cos(\omega_a t + \phi_a) + A_b \cos(\omega_b t + \phi_b) + \sum_k u_k \cos\omega_k t + \sum_k v_k \sin\omega_k t. \quad (6)$$

We note that for ω_0 small, it will always be possible to find integers m and n such that $m\omega_0$ and $n\omega_0$ are very close to ω_a and ω_b, respectively. This implies that T is much larger than $2\pi/\omega_a$ and $2\pi/\omega_b$.

By direct substitution, we find the output of the square-law device $y(t)$ to be

$$\begin{aligned} y(t) &= \alpha x^2(t) \\ &= \alpha\Big(\tfrac{1}{2}\sum_k u_k{}^2(1 + \cos2\omega_k t) + \tfrac{1}{2}\sum_k v_k{}^2(1 - \cos2\omega_k t) \\ &\quad + \sum_k u_k v_k \sin2\omega_k t + \sum\sum_{i>j} u_i u_j[\cos(\omega_i - \omega_j)t + \cos(\omega_i + \omega_j)t] + \sum\sum_{i>j} v_i v_j[\cos(\omega_i - \omega_j)t - \cos(\omega_i + \omega_j)t] \\ &\quad + \sum\sum_{i>j} u_i v_j[\sin(\omega_i + \omega_j)t - \sin(\omega_i - \omega_j)t] + \sum\sum_{i<j} u_i v_j[\sin(\omega_i + \omega_j)t + \sin(\omega_j - \omega_i)t] \\ &\quad + \tfrac{1}{2}A_a{}^2[1 + \cos(2\omega_a t + 2\phi_a)] + \tfrac{1}{2}A_b{}^2[1 + \cos(2\omega_b t + 2\phi_b)] \\ &\quad + A_a A_b[\cos\{(\omega_a + \omega_b)t + \phi_a + \phi_b\} + \cos\{(\omega_a - \omega_b)t + \phi_a - \phi_b\}] \\ &\quad + A_a\sum_k u_k[\cos\{(\omega_k + \omega_a)t + \phi_a\} + \cos\{(\omega_k - \omega_a)t - \phi_a\}] \\ &\quad + A_b\sum_k u_k[\cos\{(\omega_k + \omega_b)t + \phi_b\} + \cos\{(\omega_k - \omega_b)t - \phi_b\}] \\ &\quad + A_a\sum_k v_k[\sin\{(\omega_k + \omega_a)t + \phi_a\} + \sin\{(\omega_k - \omega_a)t - \phi_a\}] \\ &\quad + A_b\sum_k v_k[\sin\{(\omega_k + \omega_b)t + \phi_b\} + \sin\{(\omega_k - \omega_b)t - \phi_b\}]\Big), \end{aligned} \quad (7)$$

where we have used the following symmetrical relations:

$$\tfrac{1}{2}\sum\sum_{i>j} u_i u_j[\cos(\omega_i + \omega_j)t + \cos(\omega_i - \omega_j)t] = \tfrac{1}{2}\sum\sum_{i<j} u_i u_j[\cos(\omega_i + \omega_j)t + \cos(\omega_i - \omega_j)t], \quad (8a)$$

and

$$\tfrac{1}{2}\sum\sum_{i>j} v_i v_j[\cos(\omega_i + \omega_j)t + \cos(\omega_i - \omega_j)t] = \tfrac{1}{2}\sum\sum_{i<j} v_i v_j[\cos(\omega_i + \omega_j)t + \cos(\omega_i - \omega_j)t]. \quad (8b)$$

Since it is the effective bandwidth rather than the shape of the final narrowband filter that is important,[1] we choose a realizable impulse response for this filter given by

$$h(t) = 2B\cos2\pi f_c t \qquad 0 < t < 1/B. \quad (9)$$

This choice facilitates the computation in the time domain and provides accord with signal-to-noise ratios calculated previously.[1] Assuming B is very small, the time output from the bandpass filter $z(t)$ is given by

$$\begin{aligned} z(t) &= \int_0^{1/B} h(t - t')y(t')dt' \\ &= A\cos(\omega_c t + \phi) + u\cos\omega_c t + v\sin\omega_c t. \end{aligned} \quad (10)$$

Here

$$A = \alpha A_a A_b, \tag{11a}$$

$$\phi = \phi_a - \phi_b, \tag{11b}$$

and, after a great deal of calculation,[7] u and v turn out to be the sum of an infinite number of random variables, and therefore Gaussian. The means and variances of u and v are found to be[7]

$$\langle u \rangle = \langle v \rangle = 0 \tag{12}$$

and

$$\langle u^2 \rangle = \langle v^2 \rangle \simeq 4\alpha^2 (N/T)[2f_n N + (\langle A_a^2 \rangle + \langle A_b^2 \rangle)], \tag{13}$$

assuming $f_0, f_c \ll f_n$. It is also found that

$$\langle uv \rangle = \langle u \rangle \langle v \rangle = 0, \tag{14}$$

indicating that u and v are uncorrelated and independent processes. Equations (12), (13), and (14) indicate that the last two terms in Eq. (10), $u \cos \omega_c t + v \sin \omega_c t$, constitute a narrowband Gaussian random process with zero mean and center frequency ω_c. In fact, Eq. (13) represents the output noise power N_o.

We can corroborate this rather broad result[7] for a specific case by generalizing the results obtained by Kac and Siegert[8] and Emerson,[9] who have treated a related problem. We assume the output of the heterodyne mixer to consist of two sinewave signals plus uncorrelated (white) Gaussian noise. The system, in this case, consists of a realizable IF Gaussian bandpass filter, with arbitrary width Δf and a center frequency around f_a or f_b (which is large in comparison with f_c), the usual square-law device, and a realizable final narrowband filter with bandwidth B. Under the restrictions $f_c \ll f_a, f_b$ and $B \ll f_c \ll \Delta f$, it may be shown that the output of the final narrowband filter will be a sinewave signal plus a Gaussian random process of nonzero mean. For noise alone, the output will simply be Gaussian. Thus the envelope distribution for noise will be Rayleigh, and that for signal-plus-noise will be Rician. This is, we might add, the same result obtained for conventional two-frequency heterodyne detection, although the means and variances will not have the same relationship in both cases.

For $f_c \ll f_n$, as prescribed previously, it is not difficult to verify that the above description in the time domain is in accord with the frequency-domain results presented in Part 1. Since the relationship between the pulse width T and the minimum bandwidth of the final filter is governed by the Fourier transform property $TB \sim 1$ (see Ref. 10), Eq. (13) for the noise power in this regime may be written as

$$\langle u^2 \rangle = \langle v^2 \rangle = N_o \simeq 4\alpha^2 NB[2f_n N + (\langle A_a^2 \rangle + \langle A_b^2 \rangle)]. \tag{15}$$

Using Eq. (11a), we therefore obtain for the output signal-to-noise ratio,

$$(\mathrm{SNR})_o = \frac{S_o}{N_o} = \frac{\langle A^2 \rangle}{2\langle u^2 \rangle} = \frac{\langle A_a^2 A_b^2 \rangle}{8NB[(\langle A_a^2 \rangle + \langle A_b^2 \rangle)2f_n N]}. \tag{16}$$

Using an input signal-to-noise ratio given by

$$(\mathrm{SNR})_i = (\langle A_a^2 \rangle + \langle A_b^2 \rangle)/4f_n N, \tag{17}$$

we finally obtain

$$(\mathrm{SNR})_o = k_Q (\mathrm{SNR})_i^2 / [1 + 2(\mathrm{SNR})_i], \tag{18}$$

with

$$k_Q = f_n \langle A_a^2 A_b^2 \rangle / [B(\langle A_a^2 \rangle + \langle A_b^2 \rangle)^2]. \tag{19}$$

These expressions are valid in the regime $f_c \ll f_n$ and are analogous to Eqs. (40) and (41) of Part 1. Our treatment is therefore consistent with that presented previously.[1]

According to Eq. (10) and the discussion following, in the presence of signal plus noise, the output of the narrowband final filter $z(t)$, after being passed through the envelope detector, is given by the Rician distribution[11]

$$f_1(r) = (r/\sigma^2) \mathrm{I}_0(Ar/\sigma^2) \exp[-(r^2 + A^2)/2\sigma^2]. \tag{20}$$

Here, r represents the envelope of $z(t)$, $\sigma^2 = \langle u^2 \rangle = 4\alpha^2 NB[2f_n N + (\langle A_a^2 \rangle + \langle A_b^2 \rangle)]$, and $\mathrm{I}_0(x)$ is the modified Bessel function of the first kind and zero order, also expressible as

$$\mathrm{I}_0(x) = \frac{1}{2\pi} \int_0^{2\pi} \exp(x \cos\theta) d\theta. \tag{21a}$$

We may use the asymptotic expansion for $x \ll 1$ (see Ref. 12),

$$\mathrm{I}_0(x) = 1 + x^2/4 + \ldots \simeq \exp(x^2/4), \tag{21b}$$

and for $x \gg 1$,

$$\mathrm{I}_0(x) \simeq e^x/(2\pi x)^{1/2}. \tag{21c}$$

In the presence of noise alone, i.e., for $A_a = A_b = 0$, the probability density function for the envelope $f_0(r)$ is the Rayleigh distribution

$$f_0(r) = (r/\sigma_0^2) \exp[-r^2/2\sigma_0^2]. \tag{22}$$

Here σ_0^2 is the noise power in the absence of signal, i.e.,

$$\sigma_0^2 = \langle u^2 \rangle \big|_{A_a = A_b = 0} = 8\alpha^2 B f_n N^2. \tag{23}$$

We note that in this nonlinear problem $\sigma^2 \neq \sigma_0^2$ because of the presence of $s \times n$ terms in σ^2. In the usual linear systems problem these terms do not appear, and $\sigma^2 = \sigma_0^2$.

III. Binary Communications and Pulsed Radar (Vacuum Channel)

Given the probability distributions for the output signals, we can proceed to investigate binary communications and pulsed radar systems performance on choosing a decision rule. In the following, we consider both orthogonal and nonorthogonal formats for digital signaling.

A. Nonorthogonal Signaling Formats

We first consider pulse-code modulation where it is the intensity that is modulated. This simple nonorthogonal scheme is frequently referred to as PCM/IM.[13] The signal is considered to be present when a 1 is transmitted and absent when a zero is transmitted. To evaluate system performance, we choose the likelihood ratio criterion.[12,13] If Q represents the *a priori* probability that a 1 is transmitted, the signal is judged to be present if

$$Qf_1(r) \geq (1 - Q)f_0(r). \quad (24)$$

For simplicity, we assume throughout that the different types of errors are equally costly. Since the signals are pulse coded, the value of r chosen is the average value over the pulse width. The decision threshold r_D is the value of r for which the equality in Eq. (24) holds. Using Eqs. (20), (21a), and (22) for sinewave signals and Gaussian noise, r_D is therefore the solution to the transcendental equation

$$\frac{1}{2\pi}\int_0^{2\pi} \exp\left(\frac{Ar}{\sigma^2}\cos\theta\right)d\theta = \left(\frac{1-Q}{Q}\right)\frac{\sigma^2}{\sigma_0^2}\, e^{A^2/2\sigma^2} \exp\left(-\frac{\sigma^2 - \sigma_0^2}{2\sigma_0^2\sigma^2}r^2\right). \quad (25)$$

Using Eqs. (15) and (19), it is clear that

$$\sigma^2 = \sigma_0^2[1 + 2(\xi_0/k_Q)^{1/2}], \quad (26)$$

where

$$\xi_0 \equiv \langle A^2\rangle/2\sigma_0^2. \quad (27)$$

For sinewave inputs, A, A_a, and A_b are constant, and the quantity k_Q is identical with the quantity k_P introduced in Part 1 [see Eq. (40), Ref. 1]. Defining $r_0 \equiv r/\sigma_0$, Eq. (25) can be rewritten as

$$\frac{1}{2\pi}\int_0^{2\pi} \exp\left[\frac{(2\xi_0)^{1/2} r_0 \cos\theta}{1 + 2(\xi_0/k_P)^{1/2}}\right]d\theta = \left(\frac{1-Q}{Q}\right)\left[1 + 2\left(\frac{\xi_0}{k_P}\right)^{1/2}\right]\exp\left[\frac{\xi_0}{1 + 2(\xi_0/k_P)^{1/2}}\right] \times \exp\left[-\frac{r_0^2(\xi_0/k_P)^{1/2}}{1 + 2(\xi_0/k_P)^{1/2}}\right]. \quad (28)$$

Therefore, with k_P and Q fixed, the solution to Eq. (28) for r_0, which we call $\hat{r}_0$, is a function only of ξ_0. If we further define

$$\xi \equiv (\mathrm{SNR})_o = \langle A^2\rangle/2\sigma^2, \quad (29)$$

the quantity $\xi = (\sigma_0^2/\sigma^2)\xi_0 = \xi_0/[1 + 2(\xi_0/k_P)^{1/2}]$ is also a function only of ξ_0. Thus, $\hat{r}_0$ is a function only of ξ. The decision threshold $r_D \equiv \sigma_0\hat{r}_0$ is therefore a function of both ξ and σ_0.

The probability of a decoding error P_e is given by

$$P_e = Q\int_0^{r_D} f_1(r)dr + (1 - Q)\int_{r_D}^{\infty} f_0(r)dr, \quad (30)$$

which in the present case may be written as

$$P_e = Q\int_0^{r_D} \frac{r}{\sigma^2} I_0\left(\frac{Ar}{\sigma^2}\right)\exp[-(r^2 + A^2)/2\sigma^2]dr + (1 - Q)\int_{r_D}^{\infty}\frac{r}{\sigma_0^2}\exp(-r^2/2\sigma_0^2)dr. \quad (31)$$

Replacing r/σ by r', we can rewrite the first integral I_1 in Eq. (31) as follows:

$$I_1 = \int_0^{\hat{r}_0\sigma_0/\sigma} r' I_0[(2\xi)^{1/2}r']e^{-\left(\frac{r'^2}{2}+\xi\right)}dr'. \quad (32)$$

Since $\hat{r}_0\sigma_0/\sigma = \hat{r}_0/[1 + 2(\xi_0/k_P)^{1/2}]^{1/2}$ is a function only of ξ_0, which, in turn, is a function only of the output signal-to-noise ratio ξ, this integral is a function only of ξ. The second integral in Eq. (31) can be easily evaluated as follows:

$$\int_{r_D}^{\infty}\frac{r}{\sigma_0^2}e^{-r^2/2\sigma_0^2}dr = \int_{r_D/\sigma_0}^{\infty} xe^{-x^2/2}dx = -e^{-x^2/2}\Big]_{r_D/\sigma_0}^{\infty} = e^{-r_D^2/2\sigma_0^2} = e^{-\hat{r}_0^2/2}, \quad (33)$$

which is also a function only of ξ.

Therefore, with fixed $k_P(\propto f_n/B)$ and fixed Q, the probability of error P_e is a function only of the output signal-to-noise ratio $\xi \equiv (\mathrm{SNR})_o$. By use of Eq. (18), in turn, P_e can be written in terms of $(\mathrm{SNR})_i$. Computer results for the probability of error are presented in Fig. 1 [Eq. (21) has been used for the computer calculation], in which P_e is plotted against $(\mathrm{SNR})_i$ for several values of f_n/B, with the usual choice $Q = 0.5$ and $A_a = A_b$. The solid curves represent this PCM/IM scheme. For fixed f_n, the advantage of using small B is obvious.

Also shown in Fig. 1 is the P_e vs SNR curve for the conventional two-frequency heterodyne system in which no square-law device is used and f_n must be narrowed to Δf to provide a detectable signal. The output for this case is again a sinewave signal plus a narrowband (Δf) Gaussian noise.[13] Thus the computation is the same as for the three-frequency heterodyne case with $\sigma^2 = \sigma_0^2$ and $\xi = \xi_0 = (\mathrm{SNR})_i$. The probability of error at a given signal-to-noise ratio for the ordinary heterodyne system is seen to be higher than for the three-frequency system. This results from the exclusion of noise demanded by the final bandpass filter where $f_n/B > 1$, thus providing higher $(\mathrm{SNR})_o$ and lower P_e for the three-frequency system. Since the Rician and Rayleigh distributions have been calculated only for $B \ll f_c$ (hence $B \ll f_n$) and for white noise, the optimum three-frequency case considered in Part 1 is not shown in Fig. 1.

B. Pulsed Radar Application

The three-frequency nonlinear heterodyne system can also be used for pulsed radar applications. The configuration is similar to that considered previously. Pulses are sent to the target and the maximum likelihood test is used to determine whether the target is or is not present (reflected or scattered signal deemed present or absent). For a detailed treatment of conventional range-gated pulsed radar applications, the reader is referred to the book by Davenport and Root.[14]

C. Orthogonal Signaling Formats

We consider a number of orthogonal signaling formats—we begin with frequency-shift keying (FSK), which is also referred to as PCM/FM. In such a scheme, the frequency of one of the transmitted beams is fixed at the value f_1, whereas the frequency of the other is caused to shift between two values, f_2 and f_2' (not to be confused with the Doppler shifted f_2' considered in Part 1 of this paper). When a 1(0) is transmitted, the second carrier will be at frequency f_2 (f_2'). The difference frequency will therefore shift between $f_c = f_1 - f_2$ and $f_c' = f_1 - f_2'$ (assuming $f_1 > f_2, f_2'$). The frequencies $|f_1 - f_L|$, $|f_2 - f_L|$, and $|f_2' -$

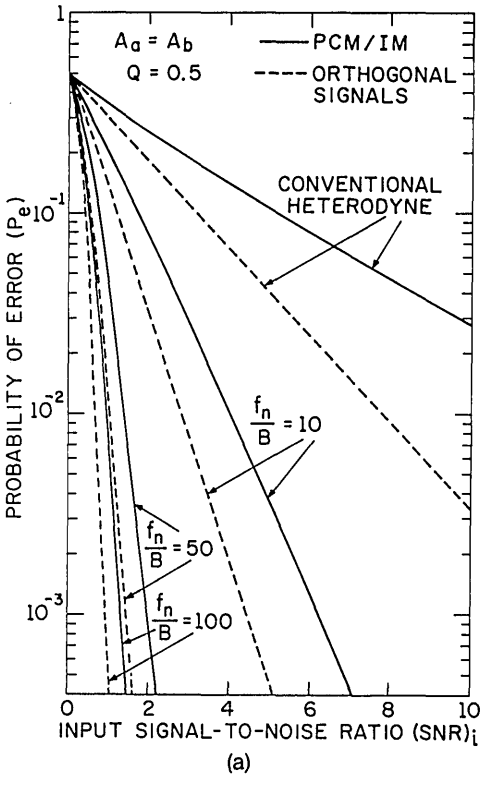
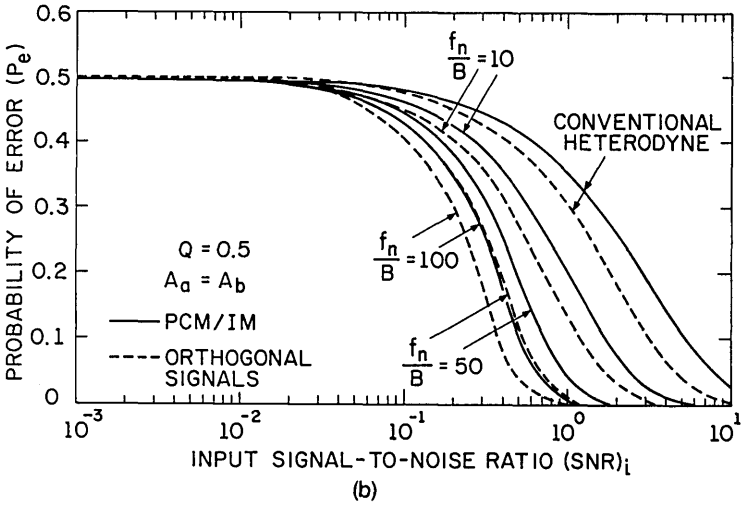


Fig. 1. (a) Probability of error vs $(SNR)_i$ for the three-frequency binary communication system in the vacuum channel. The input signals are assumed to be sinusoidal, and the noise is Gaussian. The result for the conventional heterodyne system is shown for comparison. (Log vs linear plot.) (b) Same curves on a linear vs log plot.



f_2| will all lie within the band f_n. A block diagram for such a system is shown in Fig. 2. Two narrow bandpass final filters with center frequencies at f_c and f_c' (not to be confused with the Doppler shifted f_c' considered in Part 1) are used. Following each bandpass filter is an envelope detector. If a 1(0) is transmitted, the signal will ideally pass through the top (bottom) narrow bandpass filter along with the noise; only noise will be present at the other filter.

For such an orthogonal format, the optimum single detector receiver chooses the largest signal as the correct one. Let the outputs of the first and second envelope detectors be represented by r_1 and r_2, respectively; the probability density functions for r_1 and r_2 are $h_1(r_1)$ and $h_2(r_2)$, respectively. If we assume that a 1 is transmitted, we have

$$h_1(r_1) = f_1(r_1), \tag{34a}$$

$$h_2(r_2) = f_0(r_2), \tag{34b}$$

where $f_1(\cdot)$ and $f_0(\cdot)$ are given by Eqs. (20) and (22), respectively. Using the decision rule of the largest, errors occur when $r_2 > r_1$. The error probability P_{e1} is, therefore,

$$P_{e1} = \int_0^\infty dr_1[f_1(r_1)\int_{r_1}^\infty dr_2 f_0(r_2)]$$

$$= [\sigma_0^2/(\sigma_0^2 + \sigma^2)]e^{-A^2/2(\sigma_0^2+\sigma^2)}. \tag{35}$$

This can be readily shown to be a function only of ξ. In exactly the same manner, the error probability P_{e0} when 0 is transmitted, is given by the same expression; thus $P_{e0} = P_{e1}$. The over-all probability of error P_e is therefore given by

$$P_e = QP_{e1} + (1 - Q)P_{e0} = P_{e1}, \tag{36}$$

which is presented in Fig. 1 in dashed form with the same parameters as for the PCM/IM case. The conventional heterodyne case is also shown.[13] The improvement obtained by using the orthogonal PCM/FM signaling format is seen to be substantial. In this case, however, transmitter power is required for sending both a 0 and a 1.

Another binary orthogonal pulse-code modulation scheme is polarization modulation (PCM/PL). Thus the bit 1(0) is represented by right (left) circular or vertical (horizontal) linear polarization. At the transmitter, a polarization modulator converts the laser beam into one of two polarization states. At the receiver (see Fig. 3), the circularly polarized beam

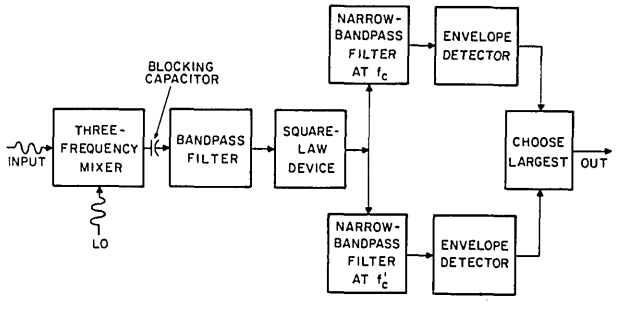


Fig. 2. Block diagram for the PCM/FM three-frequency nonlinear heterodyne receiver.

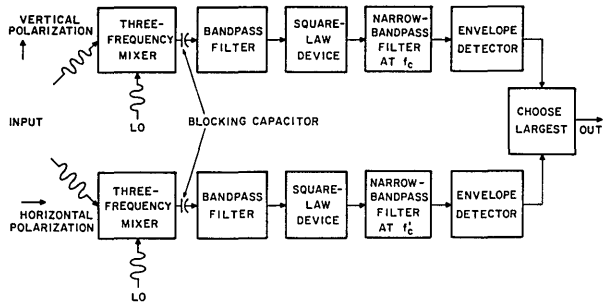


Fig. 3. Block diagram for the PCM/PL three-frequency nonlinear heterodyne receiver.

may be passed through an optical filter and then be converted to horizontal or vertical linear polarization by a quarter-wave plate. The linear polarization components are spatially separated (e.g., by a Wollaston prism) so the vertically polarized component will strike the upper photodetector and the horizontally polarized component will strike the lower photodetector. With 100% modulation, when the bit 1 is transmitted, only vertical polarization will appear at the receiver and the radiation will ideally strike only the upper detector. When a 0 is transmitted, only horizontal polarization will appear and a signal will ideally strike only the lower detector. The choice of largest decision rule is used for decoding. It is not difficult to see that the results for P_e in this case are identical to those for the PCM/FM system. Depolarization effects of the atmosphere, which are not generally large, will result in a decrease of $(\mathrm{SNR})_i$ and thus $(\mathrm{SNR})_o$ (see Refs. 15–18).

The final orthogonal format that we consider is binary pulse-position modulation (PPM/IM). In this scheme, each bit period is divided into two equal subintervals. If a 1(0) is transmitted, the pulse is caused to occur in the first (second) subinterval. A block diagram for one implementation of such a system is presented in Fig. 4. The upper (lower) gate is open for every initial (final) subinterval, and closed for every final (initial) subinterval. A time delay equal to the subinterval length is provided for the signal in the upper gate so the outputs for both intervals can be compared at the same time. The rule of largest decision is used for decoding. The results for the probability of error are again the same as those for the PCM/FM system.

The input signals for the PCM/FM, PCM/PL, and PPM/IM systems possess the orthogonality property

$$\int_{-T}^{T} S_1(t)S_0(t)dt = 0, \tag{37}$$

where $S_1(t)$ is the signal waveform representing a 1 state, and $S_0(t)$ is the signal waveform representing a 0 state. Such orthogonal modulation schemes are generally superior to nonorthogonal schemes in terms of error probability performance[16–18] and have the further advantage of requiring no more than a simple comparison for optimum reception. The M-ary signaling case is a straightforward generalization of the binary case.[19]

IV. Envelope Probability Distributions for Sinewave Signals Plus Gaussian Noise (Lognormal Channel)

The first portion of this paper was concerned with the calculation of system performance for the vacuum channel; we now turn to the error probabilities for three-frequency nonlinear heterodyne detection for the atmospheric channel. The behavior of the clear-air turbulent atmosphere as a lognormal channel for optical radiation has been well documented both theoretically and experimentally.[16–18,20–22] We therefore choose the amplitudes A_1 and A_2 to be lognormally distributed, and the phases ϕ_1 and ϕ_2 to be uniformly distributed over $(0,2\pi)$. Since $A_a \propto A_1$ and $A_b \propto A_2$, while $\phi_a = \phi_1 - \phi_L$ and $\phi_b = \phi_2 - \phi_L$, we can write

$$A_a = u_a B_a, \tag{38a}$$

$$A_b = u_b B_b, \tag{38b}$$

where B_a and B_b are constants and u_a and u_b have the same lognormal distribution:

$$P_N(u_i) = \frac{1}{\sigma_\chi u(2\pi)^{1/2}} \exp\left[-\frac{1}{2\sigma_\chi^2}(\ln u_i - m)^2\right], \quad i = a,b. \tag{39}$$

Here σ_χ is the logarithmic-amplitude standard deviation which is related to the logarithmic-irradiance standard deviation σ_I by the formula $4\sigma_\chi^2 = \sigma_I^2$ (see Ref. 22). Assuming energy is conserved and that there is no scattering of radiation out of the beam, we choose

$$\langle u_i^2 \rangle = 1, \tag{40}$$

which is equivalent to setting $m = -\sigma_\chi^2$.

Using Eq. (11a), the output amplitude A is given by

$$A = \alpha A_a A_b = \alpha B_a B_b u_a u_b. \tag{41}$$

If u_a and u_b are independent, we obtain

$$\langle A^2 \rangle = \alpha^2 B_a^2 B_b^2 \langle u_a^2 \rangle \langle u_b^2 \rangle = \alpha^2 B_a^2 B_b^2, \tag{42a}$$

or

$$\alpha B_a B_b = \langle A^2 \rangle^{1/2} = (\alpha^2 \langle A_a^2 A_b^2 \rangle)^{1/2}. \tag{42b}$$

Furthermore,

$$\ln A = \ln u_a + \ln u_b + \ln \alpha B_a B_b. \tag{43}$$

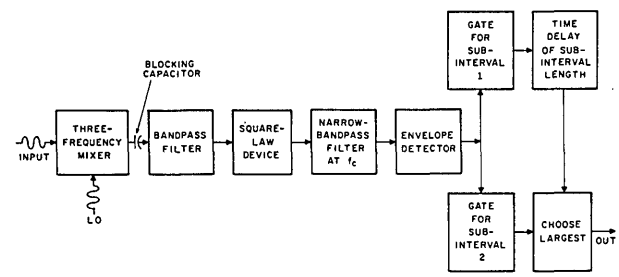


Fig. 4. Block diagram for the PPM/IM three-frequency nonlinear heterodyne receiver.

Since the quantities $y_a \equiv \ln u_a$ and $y_b \equiv \ln u_b$ will both be normally distributed as

$$f_N(y_i) = \frac{1}{(2\pi)^{1/2}\sigma_x} \exp\left[-\frac{1}{2\sigma_x^2}(y_i + \sigma_x^2)^2\right], \quad (44)$$

if u_a and u_b are independent, the variable $y_L = \ln A$ will have the normal distribution

$$f_L(y_L) = \frac{1}{2\sigma_x(\pi)^{1/2}} \exp\left[-\frac{1}{4\sigma_x^2}(y_L + 2\sigma_x^2 - \ln \alpha B_a B_b)^2\right], \quad (45)$$

from which we obtain the probability density for A:

$$f_A(A) = \frac{1}{2\sigma_x(\pi)^{1/2}A} \exp\left[-\frac{1}{4\sigma_x^2}\left(\ln\frac{A}{\langle A^2\rangle^{1/2}} + 2\sigma_x^2\right)^2\right], \quad u_a,\ u_b \text{ independent}, \quad (46)$$

where we have made use of Eq. (42b).

We also consider the situation $u_a = u_b = u$, which would arise if both incoming signals were sufficiently close in frequency and space that they suffered precisely the same fluctuations at each instant of time.[23] This case is more likely to occur in a practical situation than the independent case. For dependent fluctuations, then,

$$A = \alpha A_a A_b = \alpha B_a B_b u^2, \quad (47)$$

whence

$$\langle A\rangle = \alpha\langle A_a A_b\rangle = \alpha B_a B_b, \quad (48)$$

and

$$\ln A = \ln\langle A\rangle + 2\ln u. \quad (49)$$

Since $\ln u$ has the normal distribution $f_L(u)$ as given by Eq. (44), we find that the variable $y_L = \ln A$ has the normal probability density function:

$$f_N(y_L) = \frac{1}{2\sigma_x(2\pi)^{1/2}} \exp\left[-\frac{1}{8\sigma_x^2}(y_L + 2\sigma_x^2 - \ln\langle A\rangle)^2\right]. \quad (50)$$

By variable transformation, we obtain the probability density function for A as

$$f_A(A) = \frac{1}{2\sigma_x(2\pi)^{1/2}A} \exp\left[-\frac{1}{8\sigma_x^2}\left(\ln\frac{A}{\langle A\rangle} + 2\sigma_x^2\right)^2\right], \quad u_a = u_b. \quad (51)$$

This equation appears similar to Eq. (46); we note that $\langle A^2\rangle^{1/2}$ is replaced by $\langle A\rangle$ and the effective variance has been doubled. This results in a flattening and broadening of the probability density for the case of identical disturbance to both beams, $u_a = u_b$.

For atmospheric fluctuations that vary slowly in comparison with the pulse time T (this is the usual case, see Refs. 16–18 and 21–23), the three-frequency system envelope output will be Rician during each time interval. The over-all envelope distribution in the presence of the atmosphere $f_{1A}(r)$ will therefore be a Rician smeared over all possible values of A,

$$f_{1A}(r) = \int_0^\infty f_1(r, A) f_A(A) dA, \quad (52)$$

where $f_1(r,A)$ is given by Eq. (20). In the absence of signal, the envelope probability density remains as it was before [see Eq. (22)], since the noise alone arises

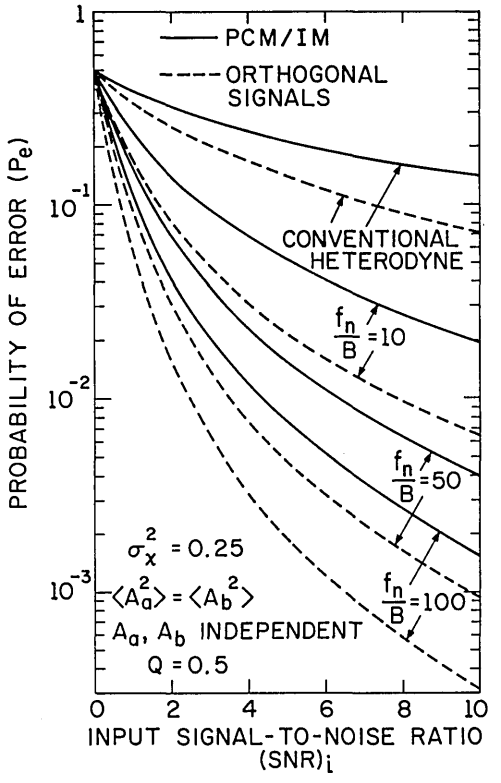


Fig. 5. Probability of error vs $(SNR)_i$ for the three-frequency binary communication system with atmospheric turbulence at the level $\sigma_x^2 = 0.25$. The input amplitudes A_a and A_b are assumed to be independent, and the noise is Gaussian. The result for the conventional heterodyne system is shown for comparison.

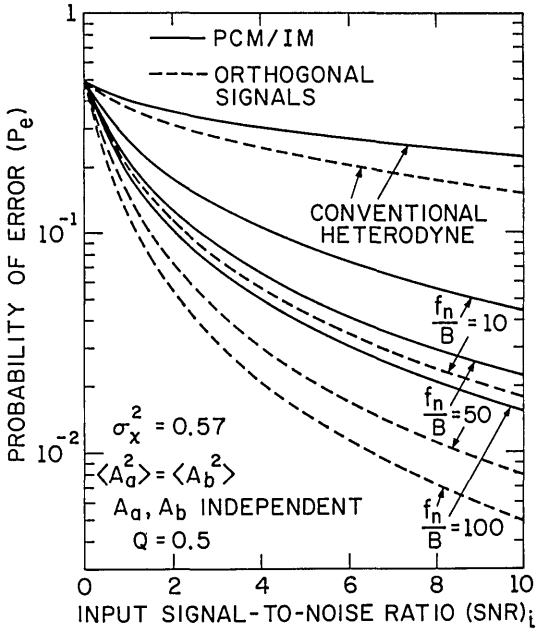


Fig. 6. Probability of error vs $(SNR)_i$ for the three-frequency binary communication system with atmospheric turbulence at the level $\sigma_x^2 = 0.57$. The input amplitudes A_a and A_b are assumed to be independent, and the noise is Gaussian. The result for the conventional heterodyne system is shown for comparison.

from the local oscillator which is unaffected by atmospheric fluctuations. Thus,

$$f_{0A}(r) = f_0(r). \tag{53}$$

V. Binary Communications (Lognormal Channel)

Under the assumptions leading to Eq. (52) and considering the various modulation schemes discussed previously, the probability of error in the presence of the lognormal turbulent atmosphere is given by

$$P_e(\text{turbulent}) = \int_0^\infty P_e(\text{quiescent}) f_A(A) dA. \tag{54}$$

This quantity was calculated using the Columbia University IBM-OS360 computer, and the results are presented in Figs. 5–8. In Figures 5 and 6, the quantities A_a and A_b were assumed to be independent with the same signal power $\langle A_a{}^2 \rangle = \langle A_b{}^2 \rangle$. The error probability curves displayed in these figures correspond to two values of the log-amplitude variance, $\sigma_\chi{}^2 = 0.25$ and $\sigma_\chi{}^2 = 0.57$. These correspond approximately to $\sigma_I = 1$ and $\sigma_I = 1.5$ (saturation value).[21,22] Other parameters are identical to those for the quiescent atmosphere as shown in Fig. 1. Figures 7 and 8 are analogous to Figs. 5 and 6, with the exception of the fact that $A_a = A_b$. For all cases, the results for conventional heterodyne operation are also shown in Figs. 5 and 6. For $\sigma_\chi \to 0$, the results properly reduce to the quiescent atmosphere data presented in Fig. 1. Computer results also indicate that the probability of error curves depend only on the signal-to-noise ratio and not on the absolute noise level in the presence of the lognormal channel, as well as in its absence.

From the graphical data presented in Figs. 1, 5, 6, 7, and 8, it is clear that orthogonal signaling formats yield better performance than nonorthogonal PCM/IM (this is also the case for direct detection[16–18]). Error probabilities are seen to increase with increasing atmospheric turbulence levels. Independent fluctuations in the two signal beams serve as a kind of diversity and thereby improve receiver performance. In all cases, furthermore, it is evident that three-frequency nonlinear heterodyne detection can provide improved performance over conventional heterodyne detection, particularly as the ratio f_n/B increases. Finally, receiver performance for the cases of phase detection with a maximum likelihood criterion[24] and phase-shift keying (PSK) have also been obtained.[7] Though PSK is definitely superior to phase detection, neither scheme provides very satisfactory performance.

VI. Future Work

The results obtained here may be extended in a number of directions. Stochastic signals, rather than sinewave signals, could be treated in the binary digital communication problem. An extensive treatment

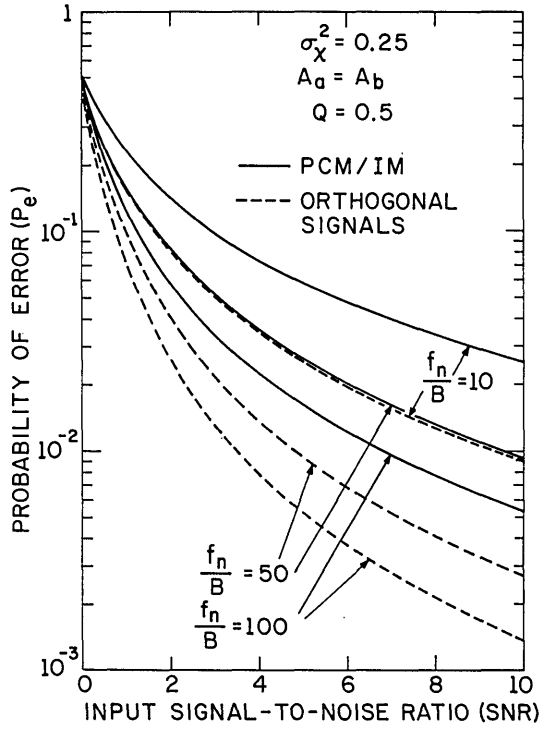


Fig. 7. Probability of error vs $(SNR)_i$ for the three-frequency binary communication system with atmospheric turbulence at the level $\sigma_\chi{}^2 = 0.25$. We assume $A_a = A_b$.

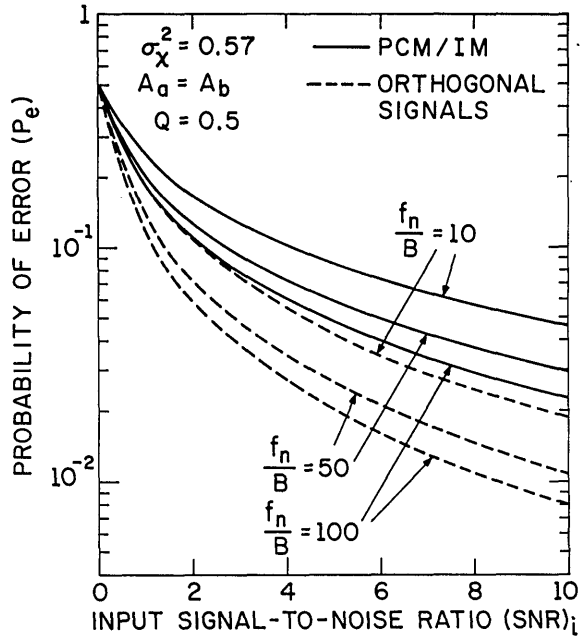


Fig. 8. Probability of error vs $(SNR)_i$ for the three-frequency binary communication system with atmospheric turbulence at the level $\sigma_\chi{}^2 = 0.57$. We assume $A_a = A_b$.

of M-ary communications is possible, as is the generalization from a single detector to an array of detectors.[16–18] Consideration could be given to the optimum matched filter detector rather than the envelope detector discussed earlier. While the present treatment consists of a per symbol analysis, prediction could be used to estimate the atmospheric turbulence level over a time period from a particular symbol, for example. In short, the usual variations possible with the conventional heterodyne system may be extended and/or modified for application to the three-frequency nonlinear heterodyne technique proposed here.

This work was supported in part by the National Science Foundation. One of us (M.C.T.) is grateful to the John Simon Guggenheim Memorial Foundation for assistance.

References

1. M. C. Teich and R. Y. Yen, App. Opt. 14, 666 (1975).
2. W. B. Davenport, Jr. and W. L. Root, *An Introduction to the Theory of Random Signals and Noise* (McGraw-Hill, New York, 1958), p. 193.
3. Ref. 2, p. 356.
4. R. G. Gallager, *Information Theory and Reliable Communication* (Wiley, New York, 1968), p. 365.
5. R. J. Schwarz and B. Friedland, *Linear Systems* (McGraw-Hill, New York, 1965), p. 299.
6. Ref. 2, pp. 93–101.
7. R. Y. Yen, "Optical Communications: An Investigation of Several Techniques," Ph.D. thesis, Columbia University, 1972.
8. M. Kac and A. J. F. Siegert, J. Appl. Phys. **18**, 383 (1947).
9. R. C. Emerson, J. Appl. Phys. **24**, 1168 (1953).
10. V. Voorhis, *Microwave Receivers* (McGraw-Hill, New York, 1948), p. 157.
11. S. O. Rice, "Mathematical Analysis of Random Noise," in *Selected Papers on Noise and Stochastic Processes*, N. Wax, Ed. (Dover, New York, 1954), pp. 133–294.
12. S. Stein and J. J. Jones, *Modern Communication Principles* (McGraw-Hill, New York, 1967), pp. 138–139.
13. W. K. Pratt, *Laser Communication Systems* (Wiley, New York, 1969), pp. 224–229.
14. Ref. 2, pp. 352–355.
15. D. L. Fried and G. E. Mevers, J. Opt. Soc. Am. **55**, 740 (1965).
16. M. C. Teich and S. Rosenberg, Appl. Opt. **12**, 2616 (1973).
17. S. Rosenberg and M. C. Teich, Appl. Opt. **12**, 2625 (1973).
18. S. Rosenberg and M. C. Teich, IEEE Trans. Inform. Theory **IT-19**, 807 (1973).
19. Ref. 12, pp. 286–309.
20. V. I. Tatarski, *Wave Propagation in a Turbulent Medium* (McGraw-Hill, New York,1961).
21. P. Diament and M. C. Teich, J. Opt. Soc. Am. **60**, 1489 (1970).
22. R. S. Lawrence and J. W. Strohbehn, Proc. IEEE **58**, 1523 (1970).
23. D. L. Fried, Appl. Opt. **10**, 721 (1971).
24. Ref. 2, p. 167.