XX. Communications Systems Research: Communications Systems Development

TELECOMMUNICATIONS DIVISION

A. On Estimating the Phase of a Square Wave in White Noise, *S. Butman*

1. Introduction

Square waves are to be used in the JPL sequential ranging system for locating distant spacecraft such as *Mariner* Mars 1969 (SPS 37-53, Vol. II, Chapter III-A). The system operates by transmitting and receiving, in succession, square-wave components whose frequencies are successively halved. The first, or highest-frequency, component provides the most precise range estimate within an unknown integer multiple of the component wavelength or period. However, each succeeding component removes half of the ambiguity left by its predecessors. The process terminates when the balance of the range ambiguity becomes discernible from other considerations.

Range measurements are obtained by estimating the phase or time delay of the received noise-corrupted target return relative to a locally generated noiseless replica of the square wave. Specifically, the received signal is correlated with two square-wave replicas spaced one-quarter period apart, with analogy to the optimum estimator for the phase of a sine wave (Ref. 1). The two correlator outputs are then combined (in a nonlinear manner) to give the required phase estimate. This is the optimum method for determining the range through tracking.

The purpose here is to determine the functional form of the optimum (maximum-likelihood) processing of the outputs of the two correlators and the accuracy of the resulting estimate. One measure of accuracy is given by the signal-to-noise ratio (SNR) out of the correlators. The sum of the output SNRs is a function of the unknown phase of the received signal, ranging from a high equal to the theoretical maximum to a low of one half of the theoretically maximum SNR, or -3 dB. This amounts to an average SNR which is 1.8 dB below the theoretical maximum, where the average is taken with respect to a uniform *a priori* phase distribution between 0 and 2π. Such an *a priori* distribution is justifiable when there is no *a priori* phase information, as would be the case during acquisition. This raises the question of whether there may not be a better choice of the two correlator functions.

A general two-correlator estimation scheme is, therefore, considered from the point of view of maximizing the average SNR during acquisition. It is found that the best two correlators for this purpose are the sine and cosine waves, even though the received signal is not sinusoidal. The sum of the SNRs is then phase-independent and is only 1.0 dB below the theoretical maximum when the received signal is a square wave. Moreover, the processing of the above correlator outputs to give the maximum-likelihood phase estimate is also independent of the structure of the ranging signal, being of the same form for all signals that it is for the sine-wave phase estimator.

2. Formulation

Let $s(t-\tau)$ denote a square wave of unit amplitude and period T that has been delayed by an amount τ, $-T/2 < \tau \leqq T/2$, and observed in the presence of additive gaussian white noise $n(t)$ of one-sided spectral density N_0, in watts/hertz, as

$$z(t) = s(t-\tau) + n(t), \qquad 0 \leqq t \leqq MT \tag{1}$$

where MT is the length of the observation time which, for convenience, is taken to be an integral number of periods. It is assumed that $s(t-\tau)$ is present during the entire observation time, starting on or before $t = 0$ and extending to $t = MT$ or beyond. It is also assumed that the *a priori* probability density $p(\tau)$ is uniform on $(-T/2, T/2]$ and that the amplitude of the signal or, equivalently, the value of N_0 is known exactly.

3. Estimation of τ Using the Outputs of Two Square-Wave Correlators Separated by One-Quarter Period

When $z(t)$ is correlated with the locally generated square waves $s(t)$ and $s[t+(T/4)]$, the correlator outputs will be

$$x = \frac{1}{MT}\int_0^{MT} z(t)\, s(t)\, dt \tag{2}$$

$$y = \frac{1}{MT}\int_0^{MT} z(t)\, s[t+(T/4)]\, dt \tag{3}$$

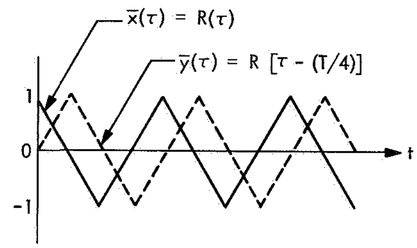
Substitution of Eq. (1) into Eqs. (2) and (3) immediately shows that

$$x = \bar{x}(\tau) + n_x \tag{4}$$

$$y = \bar{y}(\tau) + n_y \tag{5}$$

where, as shown in the following sketch, $\bar{x}(\tau) = R(\tau)$, and $\bar{y}(\tau) = R[\tau - (T/4)]$, with $R(\tau)$ being the autocorrelation function of $s(t)$ defined by

$$R(\tau) = \frac{1}{T}\int_0^{T} s(t)\, s(t+\tau)\, dt \tag{6}$$



Also,

$$n_x = \frac{1}{MT}\int_0^{MT} n(t)\, s(t)\, dt \tag{7}$$

$$n_y = \frac{1}{MT}\int_0^{MT} n(t)\, s[t+(T/4)]\, dt$$

are zero-mean gaussian random variables of variance $E[n_x^2] = E[n_y^2] = \sigma^2 = N_0/2MT$, where E is the expectation or averaging operator. They are statistically independent because they have zero cross covariance, $E[n_x n_y] = 0$.

In vector notation, we have $\mathbf{z} = \mathrm{col}(x, y)$, $\bar{\mathbf{z}}(\tau) = \mathrm{col}[\bar{x}(\tau), \bar{y}(\tau)]$, and $\mathbf{n} = \mathrm{col}(n_x, n_y)$, where $E[\mathbf{n}\mathbf{n}^T] = \sigma^2\mathbf{I}$ is the covariance matrix of the noise, $\mathbf{I}$ is the two-dimensional identity matrix, and the superscript T denotes transpose. Now, $\mathbf{z}$ is conditionally normal with conditional mean $E[\mathbf{z}|\tau] = \bar{\mathbf{z}}(\tau)$ and covariance matrix $E\{[\mathbf{z} - \bar{\mathbf{z}}(\tau)][\mathbf{z} - \bar{\mathbf{z}}(\tau)]^T|\tau\} = \sigma^2\mathbf{I}$. Consequently, the conditional probability density $p(\mathbf{z}|\tau) = p[\mathbf{z}|\bar{\mathbf{z}}(\tau]$ is

$$p(\mathbf{z}|\tau) = (2\pi\sigma^2)^{-1}\exp\left[-\frac{\|\mathbf{z} - \bar{\mathbf{z}}(\tau)\|^2}{2\sigma^2}\right] \tag{8}$$

or

$$p(x, y|\tau) = (2\pi\sigma^2)^{-1}\exp\left\{-\frac{[x - \bar{x}(\tau)]^2 + [y - \bar{y}(\tau)]^2}{2\sigma^2}\right\} \tag{9}$$

where $\|\cdot\|$ denotes the Euclidian norm.

The *a posteriori* probability density, as given by Bayes' rule, is

$$p(\tau|\mathbf{z}) = \frac{p(\mathbf{z}|\tau)}{Tp(\mathbf{z})} \tag{10}$$

where $p(\tau) = 1/T$ is the assumed *a priori* density. It is obvious from Eq. (10) that the most probable *a posteriori* estimate $\hat{\tau}$ that maximizes $p(\tau|\mathbf{z})$ over $-T/2 < \tau \leq T/2$ also maximizes $p(\mathbf{z}|\tau)$ and is, therefore, identical to the maximum-likelihood estimate, given $\mathbf{z}$. However, from Eq. (8) or (9) it is clear that $p(\mathbf{z}|\tau)$ is greatest when $\|\mathbf{z} - \bar{\mathbf{z}}(\tau)\| = \{[x - \bar{x}(\tau)]^2 + [y - \bar{y}(\tau)]^2\}^{1/2}$ is least.

Geometrically, this implies that $\hat{\tau}$ must be selected such that $\hat{\mathbf{z}} = \bar{\mathbf{z}}(\hat{\tau})$ is the closest point, from the set of possible points $\bar{\mathbf{Z}} = \mathbf{z}(\tau), \tau \in (-T/2, T/2]$, to the observed point $\mathbf{z}$. To determine $\hat{\mathbf{z}}$ analytically would be difficult, since it would be necessary to minimize $\|\mathbf{z} - \bar{\mathbf{z}}\|$ over $\bar{\mathbf{z}} \in \bar{\mathbf{Z}}$, where $\bar{\mathbf{Z}}$ is the locus of points described parametrically by

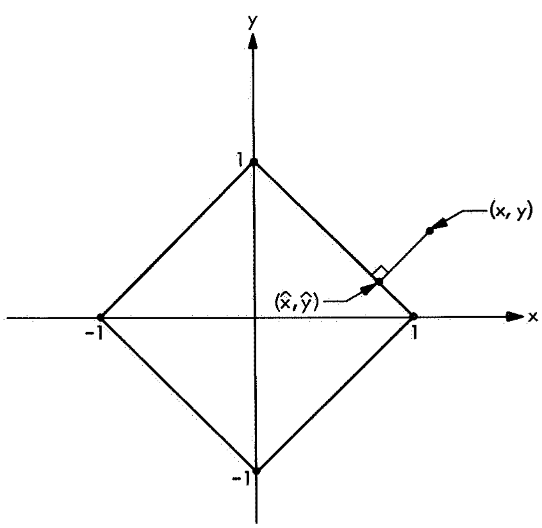
$$\bar{x} = R(\tau) = 1 - (4|\tau|/T), \qquad |\tau| \leq T/2 \tag{11}$$

$$\bar{y} = R[\tau - (T/4)] = \begin{cases} 4\tau/T & |\tau| \leq T/4 \\ 2\,\mathrm{sgn}\,\tau - (4|\tau|/T), & T/4 < |\tau| \leq T/2 \end{cases} \tag{12}$$

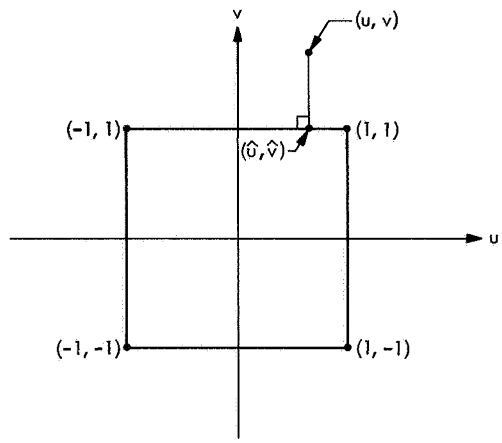
which can be combined into the simpler, but not analytic, constraint equation

$$|\bar{x}| + |\bar{y}| = 1 \tag{13}$$

Equation (13) describes the two-dimensional square of side $2^{1/2}$ shown below:



The following geometry results when the axes are rotated 45 deg using the transformation $u = 2^{-1/2}(x - y)$, $v = 2^{-1/2}(x + y)$:



In vector notation, we have $\mathbf{w} = \mathbf{U}\mathbf{z}$, where

$$\left.\begin{aligned} \mathbf{U} &= 2^{-1/2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ \mathbf{U}^{-1} &= 2^{-1/2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \end{aligned}\right\} \tag{14}$$

is the orthogonal matrix defining the rotation, and $\mathbf{w} = \mathrm{col}(u, v)$ denotes a vector in the new coordinates. Referring to the above sketch, it is easy to see that the point $\hat{\mathbf{w}} = \bar{\mathbf{w}}(\hat{\tau})$ that is nearest to the received point $\mathbf{w}$ is given by

$$\hat{u} = 2^{-1/2}\,\mathrm{sgn}\,2^{1/2}u, \qquad \hat{v} = 2^{-1/2}\,\mathrm{sat}\,2^{1/2}v, \qquad |u| > |v|$$

$$\hat{u} = 2^{-1/2}\,\mathrm{sat}\,2^{1/2}u, \qquad \hat{v} = 2^{-1/2}\,\mathrm{sgn}\,2^{1/2}v, \qquad |u| < |v|$$

where $\mathrm{sat}\,u = u$ for $|u| < 1$, $\mathrm{sat}\,u = \mathrm{sgn}\,u$ for $|u| \geq 1$, $\mathrm{sgn}\,u = 1$ for $u \geq 0$, and $\mathrm{sgn}\,u = -1$ for $u < 0$.

The region $|u| \geq |v|$ corresponds to the region $\mathrm{sgn}\,x = -\mathrm{sgn}\,y$; similarly, $|u| < |v|$ corresponds to the region $\mathrm{sgn}\,x = \mathrm{sgn}\,y$. Therefore,

$$\hat{x} = 2^{-1/2}(\hat{u} + \hat{v})$$

$$= \begin{cases} \frac{1}{2}(\mathrm{sgn}\,2^{1/2}u + \mathrm{sat}\,2^{1/2}v), & |u| > |v| \\ \frac{1}{2}(\mathrm{sgn}\,2^{1/2}v + \mathrm{sat}\,2^{1/2}u), & |u| < |v| \end{cases}$$

becomes

$$\hat{x} = \begin{cases} \tfrac{1}{2}[\operatorname{sgn}(x-y) + \operatorname{sat}(x+y)], & \operatorname{sgn} x = -\operatorname{sgn} y \\ \tfrac{1}{2}[\operatorname{sgn}(x+y) + \operatorname{sat}(x-y)], & \operatorname{sgn} x = \operatorname{sgn} y \end{cases}$$

$$= \tfrac{1}{2}[1 + \operatorname{sat}(|x| - |y|)] \operatorname{sgn} x \qquad (15)$$

Next, from Eq. (11) we have $|\hat{\tau}| = (1-\hat{x})\,T/4$, and from Eq. (12) it is evident that $\operatorname{sgn}\hat{\tau} = \operatorname{sgn}\hat{y}$. Also, $\operatorname{sgn}\hat{y} = \operatorname{sgn} y$. Therefore,

$$\frac{\hat{\tau}}{T} = \tfrac{1}{4}\{1 - \tfrac{1}{2}[1 + \operatorname{sat}(|x| - |y|)] \operatorname{sgn} x\} \operatorname{sgn} y \qquad (16)$$

$$= \tfrac{1}{8}[2 \operatorname{sgn} y - \operatorname{sgn} x \operatorname{sgn} y - \operatorname{sat}(x \operatorname{sgn} y - y \operatorname{sgn} x)] \qquad (17)$$

The last expression for $\hat{\tau}$ can be implemented easily in either digital or analog fashion. The pieces of analog equipment required, in addition to the correlators, are three multipliers, three adders, two hard limiters, and one soft limiter. The complete mechanization is shown in Fig. 1.

The distribution of the true value of the delay τ about the maximum-likelihood estimate $\hat{\tau}$ is given by the *a posteriori* probability density $p(\tau|\hat{\tau})$, which is related to the conditional probability density $p(\hat{\tau}|\tau)$ through Bayes' formula:

$$p(\tau|\hat{\tau}) = \frac{p(\hat{\tau}|\tau)\,p(\tau)}{\int_{-T/2}^{T/2} p(\hat{\tau}|\tau)\,p(\tau)\,d\tau} = \frac{p(\hat{\tau}|\tau)}{\int_{-T/2}^{T/2} p(\hat{\tau}|\tau)\,d\tau} \qquad (18)$$

since $p(\tau) = 1/T$. The conditional probability density $p(\hat{\tau}|\tau)$, on the other hand, determines the distribution of the maximum-likelihood estimate $\hat{\tau}$ about the true value τ. It is clear from Bayes' formula that the plot of $p(\tau|\hat{\tau})$ versus τ is of the same shape as that of $p(\hat{\tau}|\tau)$ versus τ [but not the same shape as that of $p(\hat{\tau}|\tau)$ versus $\hat{\tau}$].

It is a straightforward matter to determine $p(\hat{\tau}|\tau)$ analytically. Thus, the probability of obtaining $\hat{\tau} = 0$ is equal to the probability of decoding $\mathbf{w} = \operatorname{col}(u, v)$ as the corner $\hat{\mathbf{w}} = (2^{-1/2}, 2^{-1/2})$. This happens if, and only if, $u > 2^{-1/2}$

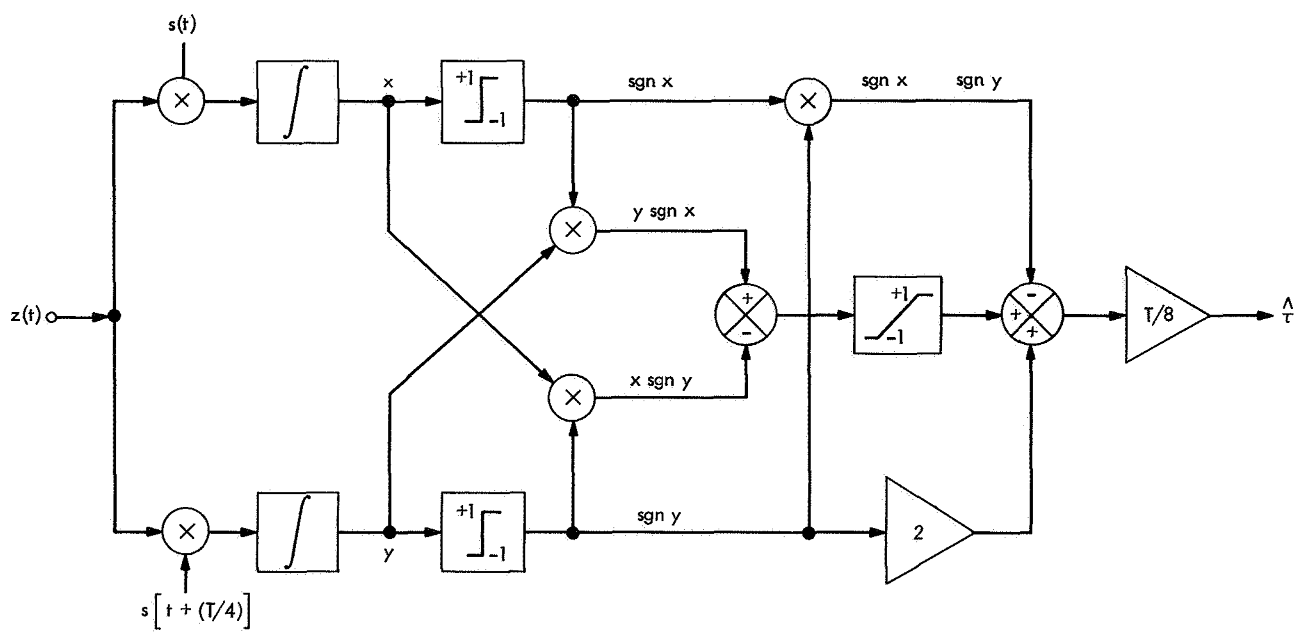


Fig. 1. Range estimator mechanization

and $v > 2^{-1/2}$; hence,

$$\Pr\{\hat{\tau} = 0|\tau\} = \Pr\{u > 2^{-1/2}, v > 2^{-1/2}|\hat{u}(\tau), \hat{v}(\tau)\}$$
$$= \int_{2^{-1/2}}^{\infty}\int_{2^{-1/2}}^{\infty} \exp\left[-\frac{(u-\bar{u})^2 + (v+\bar{v})^2}{2\sigma^2}\right] du\, dv$$
$$= Q[2^{-1/2} - \bar{u}(\tau)]\, Q[2^{-1/2} - \bar{v}(\tau)] \qquad (19)$$

where

$$Q(a) = (2\pi\sigma^2)^{-1/2}\int_a^{\infty} \exp(-b^2/2\sigma^2)\, db \qquad (20)$$

and $\Pr\{\cdot\}$ denotes probability (as opposed to probability density). In the present context, $p(\hat{\tau} = 0|\tau) = \delta(\hat{\tau})\Pr\{\hat{\tau} = 0|\tau\}$,

$$\bar{u}(\tau) = 2^{-1/2}[\bar{x}(\tau) - \bar{y}(\tau)]$$
$$= 2^{-1/2}\left[1 - 4\left|\frac{\tau}{T}\right| - \left(1 - \left|1 - 4\left|\frac{\tau}{T}\right|\right|\right)\operatorname{sgn}\tau\right] \qquad (21)$$

$$\bar{v}(\tau) = 2^{-1/2}[\bar{x}(\tau) + \bar{y}(\tau)]$$
$$= 2^{-1/2}\left[1 - 4\left|\frac{\tau}{T}\right| + \left(1 - \left|1 - 4\left|\frac{\tau}{T}\right|\right|\right)\operatorname{sgn}\tau\right] \qquad (22)$$

Similarly,

$$\Pr\{\hat{\tau} = (T/4)|\tau\} = Q[2^{-1/2} + \bar{u}(\tau)]\, Q[2^{-1/2} - \bar{v}(\tau)]$$
$$\Pr\{\hat{\tau} = (T/2)|\tau\} = Q[2^{-1/2} + \bar{u}(\tau)]\, Q[2^{-1/2} + \bar{v}(\tau)]$$
$$\Pr\{\hat{\tau} = -(T/4)|\tau\} = Q[2^{-1/2} - \bar{u}(\tau)]\, Q[2^{-1/2} + \bar{v}(\tau)]$$

Next, we determine $p(\hat{\tau}|\tau)$ for $-(T/4) < \tau < 0$. In this region, $\hat{u} = 2^{-1/2}$ and $|\hat{v}| < 2^{-1/2}$ with $\hat{v} = 2^{-1/2}[1 + (8\hat{\tau}/T)]$. Therefore,

$$p(\hat{\tau}|\tau) = 2^{-1/2}\frac{8}{T}\, p\{\hat{u} = 2^{-1/2}, \hat{v} = 2^{-1/2}[1 + (8\hat{\tau}/T)]|\tau\}$$

However, $\hat{u} = 2^{-1/2}$ and $\hat{v} = v$ if, and only if, $u > |v| = |\hat{v}|$; hence,

$$p(2^{-1/2}, v) = Q(|\hat{v}| - \bar{u}) \times (2\pi\sigma^2)^{-1/2}\exp[-(\hat{v} - \bar{v})^2/2\sigma^2]$$

and

$$p(\hat{\tau}|\tau) = (4\pi\sigma^2)^{-1/2}\frac{8}{T}\, Q[2^{-1/2}|1 + (8\hat{\tau}/T)| - \bar{u}] \times \exp\left[-\frac{1 - (8\hat{\tau}/T) - 2^{1/2}\bar{v}}{4\sigma^2}\right]$$

A similar procedure is used for the remaining regions. The end result is

$$p(\hat{\tau}|\tau) = \delta(\hat{\tau})\, Q[2^{-1/2} - \bar{u}(\tau)]\, Q[2^{-1/2} - \bar{v}(\tau)] + \delta[\hat{\tau} - (T/4)]\, Q[2^{-1/2} + \bar{u}(\tau)]\, Q[2^{-1/2} - \bar{v}(\tau)]$$
$$+ \delta[\hat{\tau} + (T/4)]\, Q[2^{-1/2} - \bar{u}(\tau)]\, Q[2^{-1/2} + \bar{v}(\tau)] + \delta[\hat{\tau} - (T/2)]\, Q[2^{-1/2} + \bar{u}(\tau)]\, Q[2^{-1/2} + \bar{v}(\tau)]$$
$$+ (4\pi\sigma^2)^{-1/2}\frac{8}{T}\begin{cases} Q\left[2^{-1/2}\left|1 + \frac{8\hat{\tau}}{T}\right| - \bar{u}(\tau)\right]\exp\left\{-\frac{[1 + (8\hat{\tau}/T) - 2^{1/2}\bar{v}(\tau)]^2}{4\sigma^2}\right., & \hat{\tau}\epsilon\left(-\frac{T}{4}, 0\right) \\ Q\left[2^{-1/2}\left|1 + \frac{8\hat{\tau}}{T}\right| - \bar{v}(\tau)\right]\exp\left\{-\frac{[1 - (8\hat{\tau}/T) - 2^{1/2}\bar{u}(\tau)]^2}{4\sigma^2}\right., & \hat{\tau}\epsilon\left(0, \frac{T}{4}\right) \\ Q\left[2^{-1/2}\left|3 - \frac{8\hat{\tau}}{T}\right| + \bar{u}(\tau)\right]\exp\left\{-\frac{[3 - 8\hat{\tau}/T) - 2^{1/2}\bar{v}(\tau)]^2}{4\sigma^2}\right., & \hat{\tau}\epsilon\left(\frac{T}{4}, \frac{T}{2}\right) \\ Q\left[2^{-1/2}\left|3 - \frac{8\hat{\tau}}{T}\right| + \bar{v}(\tau)\right]\exp\left\{-\frac{[3 + (8\hat{\tau}/T) - 2^{1/2}\bar{u}(\tau)]^2}{4\sigma^2}\right., & \hat{\tau}\epsilon\left(-\frac{T}{2}, -\frac{T}{4}\right) \end{cases} \qquad (23)$$

The conditional probability density $p(\hat{\tau}|\tau)$ is plotted in Fig. 2 for several values of σ and $\tau \in [0, T/8]$. The plots for $\tau \in [T/8, T/4]$ are then obtained by reflecting the original set of graphs about the axis $\hat{\tau} = T/8$. Similarly, $p(\hat{\tau}|\tau)$ for $\tau \in [T/4, 3T/8]$ is a reflection about $\hat{\tau} = T/4$ of the plot of $p(\hat{\tau}|\tau)$ for $\tau \in [T/8, T/4]$, etc.

The *a posteriori* probability density $p(\tau|\hat{\tau})$ is plotted in Fig. 3 and satisfies the same conditions of symmetry as $p(\hat{\tau}|\tau)$. However, it should be observed that $p(\tau|\hat{\tau})$ has no delta functions even though $p(\hat{\tau}|\tau)$ does. Since $p(\hat{\tau})$ has delta functions at the same places as $p(\hat{\tau}|\tau)$, the delta functions cancel in $p(\tau|\hat{\tau}) = p(\hat{\tau}|\tau)/Tp(\hat{\tau})$.

4. The General Two-Correlator Problem

The preceding discussion was concerned with making the best estimate, given the outputs x and y of the two orthogonal square-wave correlators. It is logical to inquire now whether there may not exist a better choice of correlators (assuming, of course, that the correlator outputs can always be processed in an optimum manner).

One measure of the performance of a correlator is the SNR:

$$\rho_x(\tau) = [\bar{x}(\tau)/\sigma]^2 \tag{24}$$

$$\rho_y(\tau) = [\bar{y}(\tau)/\sigma]^2 \tag{25}$$

The sum of the two SNRs is then

$$\rho_z(\tau) = [\|\bar{\mathbf{z}}(\tau)\|/\sigma]^2 \tag{26}$$

For the square-wave correlators, $\bar{\mathbf{z}}(\tau)$ is on the square of side $2^{1/2}$ and $\rho_z(\tau)$ varies from a maximum of $1/\sigma^2$ to a minimum of $1/2\sigma^2$ with periodicity $T/4$. The average value of $\rho_z(\tau)$ when τ is uniform on $(-T/2, T/2]$ is, therefore,

$$\bar{\rho}_z = \frac{1}{T}\int_{-T/2}^{T/2} \rho_z(\tau)\,dr \tag{27}$$

$$= 2/3\sigma^2 \tag{28}$$

This average value would be obtained during acquisition, when there is no knowledge of τ other than that it is equi-probable on $(-T/2, T/2]$. However, once an estimate $\hat{\tau}$ has been obtained, the receiver can readjust the local zero reference to $\hat{\tau}$ and obtain $\rho_z(\tau - \hat{\tau})$ with an *a posteriori* average of

$$\bar{\rho}_z(\tau) = \int_{\hat{\tau}-(T/2)}^{\hat{\tau}+(T/2)} \rho_z(\tau - \hat{\tau})\,\rho(\tau|\hat{\tau})\,d\tau \tag{29}$$

The continual updating of the local zero reference to the latest estimate $\hat{\tau}$ is known as *tracking* and can be implemented with a phase-locked loop. As the estimate $\hat{\tau}$ improves, $p(\tau|\hat{\tau})$ approaches $\delta(\tau - \hat{\tau})$ and

$$\bar{\rho}_z(\hat{\tau}) \rightarrow \rho_z(0) = 1/\sigma^2$$

which is the theoretical maximum. Therefore, the use of square-wave correlators is optimum during tracking. However, it is still necessary to determine the best two correlators for acquisition purposes.

5. Optimum Correlators for Acquisition

Suppose that $r(t)$ is correlated with some pair of periodic, orthonormal, but otherwise arbitrary, time functions $f(t)$ and $h(t)$. Then, the correlator outputs are

$$f = \frac{1}{MT}\int_0^{MT} r(t)\,f(t)\,dt \tag{30}$$

$$= \frac{1}{MT}\int_0^{MT} s(t-\tau)\,f(t) + \frac{1}{T}\int_0^{T} n(t)\,f(t)$$

or

$$f = \phi_{fs}(\tau) + n_f \tag{31}$$

and

$$h = \phi_{hs}(\tau) + n_h \tag{32}$$

where $\phi_{fs}(\tau)$ and $\phi_{hs}(\tau)$ are cross-correlation functions, while n_f and n_h are independent zero-mean gaussian random variables of variance σ^2. Consequently,

$$\rho_f(\tau) = \phi_{fs}^2(\tau)/\sigma^2$$

$$\rho_h(\tau) = \phi_{hs}^2(\tau)/\sigma^2$$

$$\rho = \rho_f + \rho_h \tag{33}$$

$$\bar{\rho} = \frac{1}{T}\int_{-T/2}^{T/2} \rho(\tau)\,dr$$

$$= \bar{\rho}_f + \bar{\rho}_h \tag{34}$$

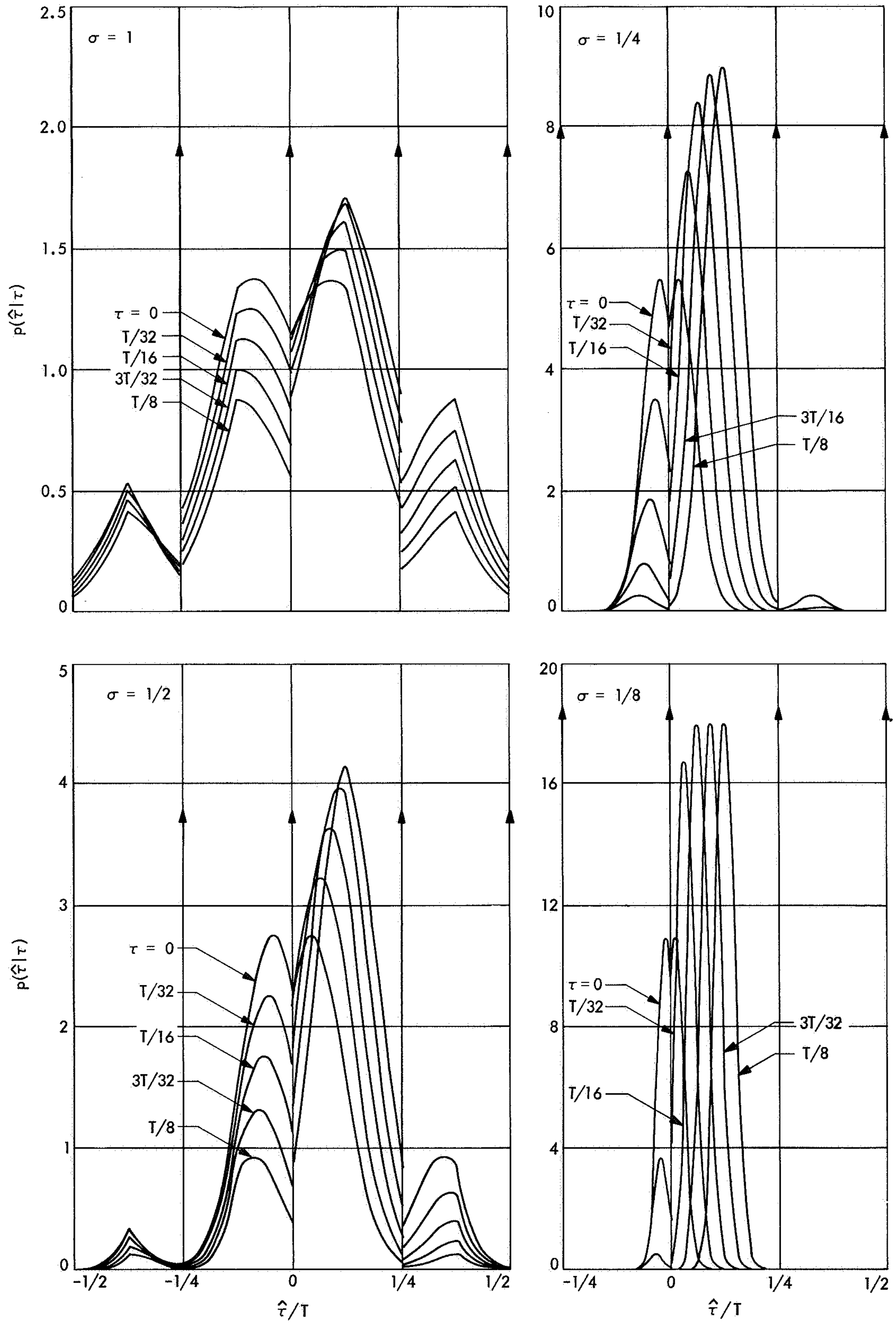


Fig. 2. Conditional probability density $p(\hat{\tau}|\tau)$ vs $\hat{\tau}/T$ for various values of σ

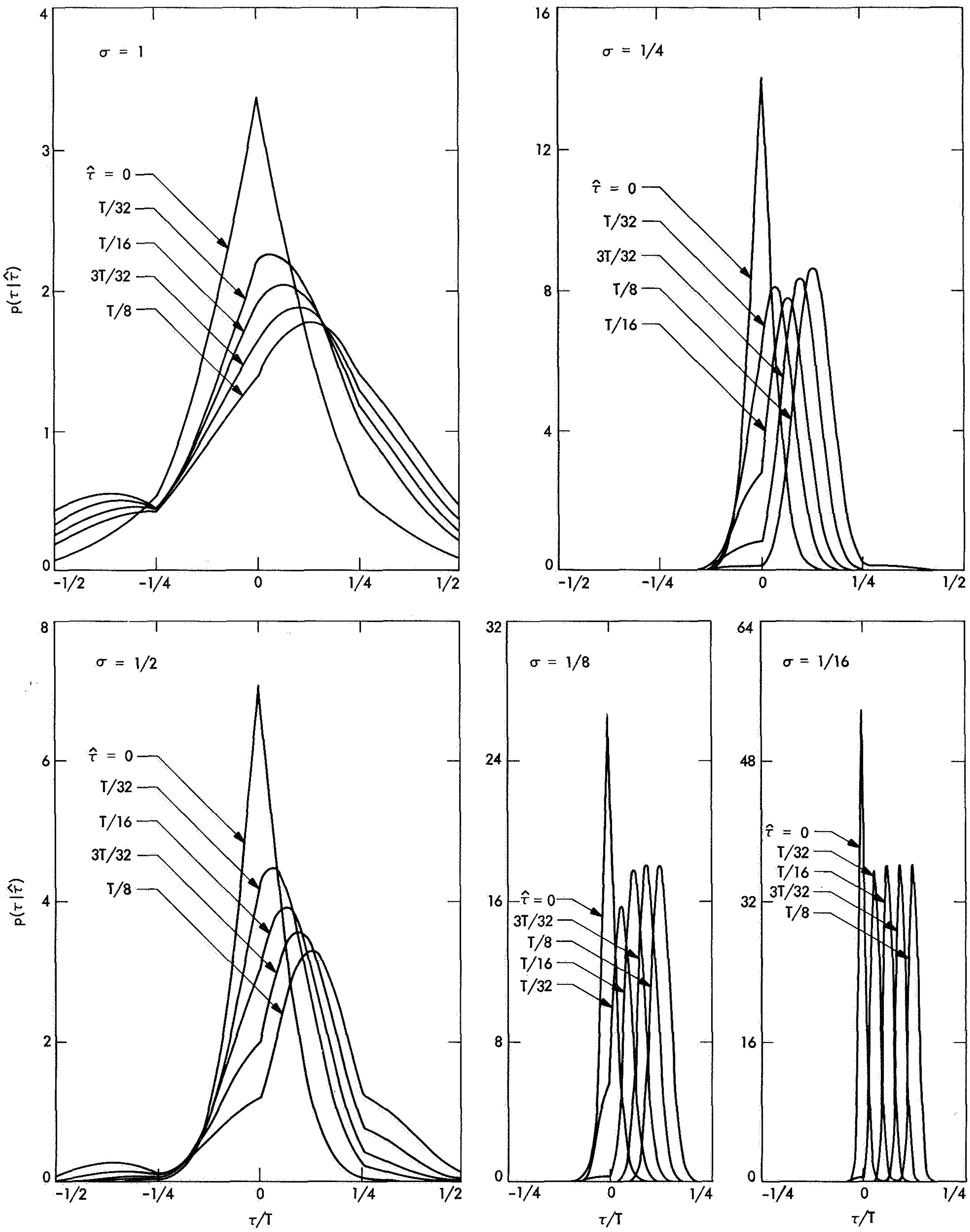


Fig. 3. A *posteriori* probability density $p(\tau|\hat{\tau})$ vs $\hat{\tau}/T$ for various values of σ

Now,

$$\bar{\rho}_f = \frac{1}{\sigma^2 M^2 T^3} \int_{-T/2}^{T/2} \int_0^{MT} \int_0^{MT} f(t)\, s(t-\tau)\, s(u-\tau)\, f(u)\, dt\, du\, d\tau$$

$$= \frac{1}{\sigma^2} \left(\frac{1}{MT}\right)^2 \int_0^{MT} \int_0^{MT} f(t)\, R(t-u)\, f(u)\, du\, dt \tag{35}$$

Similarly,

$$\bar{\rho}_h = \frac{1}{\sigma^2} \left(\frac{1}{MT}\right)^2 \int_0^{MT} \int_0^{MT} h(t)\, R(t-u)\, h(u)\, dt\, du \tag{36}$$

However, $R(t)$ is an even periodic function of period T with a positive spectrum:

$$R(t-u) = \sum_k^{\infty} s_k^2 \cos \frac{2\pi k}{T}(t-u) \tag{37}$$

Also, $f(t)$ and $h(t)$ are periodic functions of the general form

$$f(t) = \sum_{k=1}^{\infty} \left(a_k \sin \frac{2\pi k}{T} t + d_k \cos \frac{2\pi k}{T} t \right) \tag{38}$$

$$h(t) = \sum_{k=1}^{\infty} \left(c_k \sin \frac{2\pi k}{T} t + d_k \cos \frac{2\pi k}{T} t \right) \tag{39}$$

with

$$\frac{1}{MT} \int_0^{MT} f^2(t)\, dt = \sum_{k=1}^{\infty} (a_k^2 + b_k^2) = 1 \tag{40}$$

$$\frac{1}{MT} \int_0^{MT} h^2(t)\, dt = \sum_{k=1}^{\infty} (c_k^2 + d_k^2) = 1 \tag{41}$$

Therefore,

$$\bar{\rho}_f = \frac{1}{\sigma^2} \sum_k s_k^2 \frac{1}{MT} \int_0^{MT} \int_0^{MT} f(u)\, f(t) \cos 2\pi k (t-u)\, dt\, du$$

$$= \frac{1}{2\sigma^2} \sum_k s_k^2 (a_k^2 + b_k^2)$$

$$\leq \frac{1}{2\sigma^2} s_{\max}^2 \tag{42}$$

with equality if, and only if,

$$f(t) = a \sin \frac{2\pi m}{T} t + b \cos \frac{2\pi m}{T} t$$

where m is the subscript of $s_{\max}$. Consequently, the optimum choices for $f(t)$ and $h(t)$ are

$$f(t) = 2^{1/2} \cos \left(\frac{2\pi m}{T} t + \theta \right) \tag{43}$$

$$h(t) = 2^{1/2} \sin \left(\frac{2\pi m}{T} t + \theta \right) \tag{44}$$

where θ is arbitrary and $\bar{\rho} = s_m^2/\sigma^2$. In addition, $\rho_f(\tau)$ is now independent of τ; hence,

$$\rho(\tau) = \bar{\rho} = s_m^2/\sigma^2 \tag{45}$$

In the case of a square wave, $s_m^2 = s_1^2 = 8/\pi^2$, and we obtain

$$\rho(\tau) = \bar{\rho} = 8/\pi^2\sigma^2 \tag{46}$$

which shows that the theoretical maximum of $1/\sigma^2$ is unachievable in the square-wave case.

For this choice of correlators, the maximum likelihood-estimate is well-known to be (Ref. 1)

$$\hat{\tau} = \frac{T}{2\pi} \arctan(y/x) + \frac{T}{2}(1 - \operatorname{sgn} x) \tag{47}$$

where

$$x = \frac{1}{MT} \int_0^{MT} r(t) \cos(2\pi t/T)\, dt$$

$$y = \frac{1}{MT} \int_0^{MT} r(t) \sin(2\pi t/T)\, dt$$

are the correlator outputs.

6. Mean-Square Error

The probability densities $p(\tau|\hat{\tau})$ and $p(\hat{\tau}|\tau)$ contain all of the statistical information about the performance of the estimator. Of particular interest, however, is the *a posteriori* mean-square error

$$E\left[\left(\frac{\tau-\hat{\tau}}{T}\right)^2 \middle| \hat{\tau}\right] = \int_{\hat{\tau}-(T/2)}^{\hat{\tau}+(T/2)} \left(\frac{\tau-\hat{\tau}}{T}\right)^2 p(\tau|\hat{\tau})\, d\tau$$

since it gives the scatter of the true value of the delay τ about the maximum-likelihood estimate $\hat{\tau}$. It is easy to show that this *a posteriori* mean-square error approaches $\sigma^2/32$ as σ^2 goes to zero, provided $\hat{\tau} \neq \pm kT/4$, $k=0,1,2,3$. This is because $p(\tau|\hat{\tau})$ tends to a gaussian density of mean $\hat{\tau}/T$ and variance $\sigma^2/32$. At $\hat{\tau} = \pm kT/4$, however, $p(\tau|\hat{\tau})$ is proportional to $Q(4 \cdot 2^{1/2}|\tau - \hat{\tau}|/T)$. Thus, if $\sigma^2 \rightarrow 0$ and $x = (\tau - \hat{\tau})/T$, we can write

$$E\left[\left(\frac{\tau-\hat{\tau}}{T}\right)^2 \middle| \hat{\tau} = \pm kT/4\right] = \frac{\int_{-\infty}^{\infty} x^2 Q(4 \cdot 2^{1/2}|x|)\, dx}{\int_{-\infty}^{\infty} Q(4 \cdot 2^{1/2}|x|)\, dx}$$

$$= \frac{\int_{0}^{\infty} x^2 Q(4 \cdot 2^{1/2} x)\, dx}{\int_{0}^{\infty} Q(4 \cdot 2^{1/2} x)\, dx}$$

Integrating by parts and noting that

$$Q'(4 \cdot 2^{1/2} x) = (2\pi\sigma^2)^{-1/2} \exp(-32\, x^2/\sigma^2)$$

we obtain

$$E\left[\left(\frac{\tau-\hat{\tau}}{T}\right)^2 \middle| \hat{\tau} = \pm kT/4\right] = \frac{2}{3}\, \frac{\int_{0}^{\infty} x^3 \exp(-32\, x^2/\sigma^2)\, dx}{\int_{0}^{\infty} x \exp(-32\, x^2)\, dx}$$

$$= \sigma^2/48$$

The above analytical results have also been verified numerically on a general-purpose digital computer. Numerical values of the *a posteriori* mean-square error in units of $\sigma^2/32$ are tabulated in Table 1 for $\hat{\tau} = kT/32$, $k = 0, \cdots, 4$, and $\sigma = 2^{-k}$, $k = 0, \cdots, 4$.

Reference

1. Viterbi, A. J., *Principles of Coherent Communication*, pp. 129. McGraw Hill Book Company, Inc., New York, 1966.

Table 1. *A posteriori* mean-square errors

$\hat{\tau}/T$	*A posteriori* mean-square error for indicated σ				
	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$
0	0.865	0.703	0.667	0.667	0.667
1/32	1.300	1.250	1.050	0.905	0.910
1/16	1.430	1.600	0.910	0.910	0.995
3/32	1.560	2.300	0.900	0.995	0.995
1/8	1.390	3.700	2.000	1.000	1.000

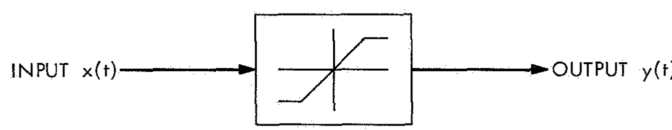
B. Analysis of Narrow-Band Signals Through the Band-Pass Soft Limiter, *R. C. Tausworthe*

1. Introduction

Several authors (Refs. 1–3) have examined the output SNR characteristics of the so-called "soft" limiter, giving several approximations for the output signal and noise terms as functions of the input parameters. The ensuing article illustrates that, under a widely accepted model of the soft limiter, the output signal power and signal suppression can be found exactly in terms of the hard-limiter signal-suppression function. The output noise is correspondingly then well approximated.

2. Limiter Suppression Factor

In the discussion below, we shall assume that the following device input



is a narrow-band waveform consisting of a signal immersed in gaussian noise of variance σ_N^2:

$$x(t) = V(t) \sin[\omega_0 t + \theta(t)] + n(t)$$
$$= V \sin\phi + n(t) \qquad (1)$$

where

$$V = V(t)$$
$$\phi = \omega_0 t + \theta(t)$$

It has been shown (SPS 37-44, Vol. IV, pp. 303–307) that the portion of the limiter output due to input signal is

$$G(V\sin\phi) = E[y(x)|V\sin\phi]$$
$$= c_1 \sin\phi + c_2 \sin 2\phi + \cdots \qquad (2)$$

in which the coefficient

$$c_k = \frac{1}{\pi}\int_{-\pi}^{\pi} G(V\sin\phi)\sin k\phi\, d\phi \qquad (3)$$

represents the amplitude of the signal in the kth harmonic zone. For the hard limiter, the c_k have been evaluated as

$$c_k = L\left(\frac{2}{\pi}\right)^{1/2}\left(\frac{v}{k\pi}\right)\int_{-\pi}^{\pi}\cos\phi\cos k\phi\exp\left\{-\frac{1}{2}v^2\sin^2\phi\right\}d\phi$$

$$= L\left(\frac{2}{\pi}\right)^{1/2}\left(\frac{v}{k}\right)\exp\frac{-v^2}{4}\left[I_{(k-1)/2}\left(\frac{v^2}{4}\right)+I_{(k+1)/2}\left(\frac{v^2}{4}\right)\right] \qquad (4)$$

for odd k, in terms of the parameter $v = V(t)/\sigma_N$ and the modified Bessel functions of the first kind (Ref. 4).

With an input SNR of ρ the *hard-limiter suppression factor* $\alpha^2(\rho)$ is defined as the ratio of the fundamental signal output power to what it would be if noise were absent. When $V(t)$ is a constant amplitude,

$$\rho = \frac{1}{2}v^2,$$

so

$$\alpha^2(\rho) = \frac{\pi}{4}\rho e^{-\rho}\left[I_0\left(\frac{\rho}{2}\right)+I_1\left(\frac{\rho}{2}\right)\right]^2 \qquad (5)$$

An excellent approximation for α^2 (Ref. 5) is

$$\alpha^2(\rho) = \frac{0.7854\rho + 0.4768\rho^2}{1 + 1.024\rho + 0.4768\rho^2} \qquad (6)$$

If, however, $V(t)$ is time varying, then the input SNR is

$$\rho = E\left(\frac{1}{2}v^2\right)$$

and

$$\alpha^2(\rho) = E\left[\alpha^2\left(\frac{1}{2}v^2\right)\right] \qquad (7)$$

Suppression is probably computed with least difficulty in this case through the approximation given in Eq. (6).

3. Soft Limiter Model

We shall take as the model of the soft limiter the function plotted in Fig. 4:

$$y = L\,\mathrm{erf}\left[\left(\frac{K\pi^{1/2}}{2L}\right)x\right] = L\,\mathrm{erf}(Bx) \qquad (8)$$

where erf(x) is the well-known error function (Ref. 4)

$$\mathrm{erf}\, x = \frac{2}{\pi^{1/2}}\int_0^x \exp(-t^2)\,dt \qquad (9)$$

and $B = K\pi^{1/2}/2L$.

Our model is thus seen to possess the following characteristics: For values of x much less than $2L/K\pi^{1/2}$, the device acts as a linear amplifier with voltage gain K. For inputs x much larger than $2L/K\pi^{1/2}$, signal limiting occurs, with the limit level L. Further, as $K \to \infty$ for fixed L, the device becomes a hard limiter, and as $L \to \infty$ for fixed K, the device becomes a linear amplifier. The soft limiter model we have chosen thus degenerates to previously analyzed devices in limiting cases.

Evaluation of the limiter performance thus now depends only upon finding $G(V\sin\phi)$ and its Fourier coefficients for the assumed characteristic. In the present case $G(V\sin\phi)$ takes the form

$$G(V\sin\phi) = \frac{L}{\sigma_N(2\pi)^{1/2}}\int_{-\infty}^{+\infty}\mathrm{erf}\,B(V\sin\phi + n) \times\exp\left\{\frac{-n^2}{2\sigma_N^2}\right\}dn \qquad (10)$$

Although the results to follow are quite general, we shall evaluate only the behavior in the fundamental output zone:

$$c_1 = \left(\frac{L}{\pi}\right)\frac{1}{\sigma_N(2\pi)^{1/2}}\int_{-\infty}^{+\infty}\exp\left\{\frac{-n^2}{2\sigma_N^2}\right\}dn \times\int_{-\pi}^{\pi}\sin\phi\,\mathrm{erf}\left[B(V\sin\phi+n)\right]d\phi \qquad (11)$$

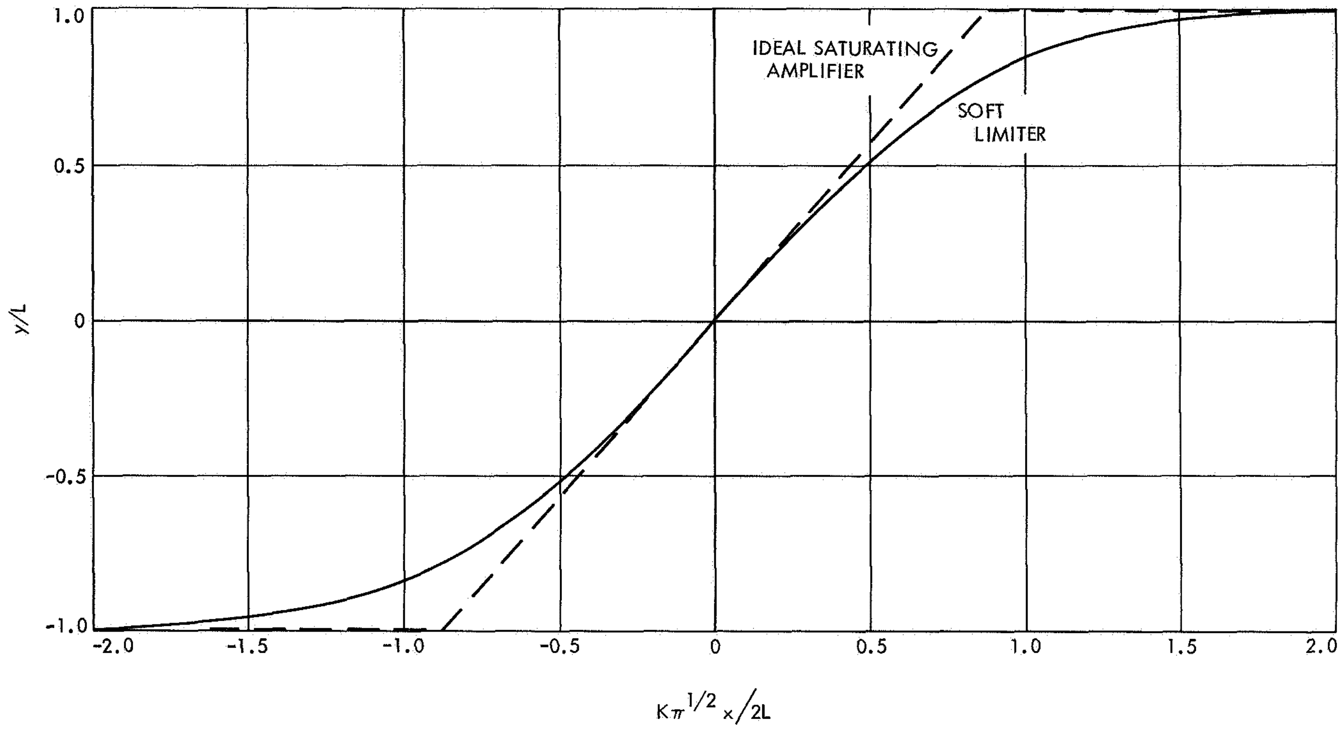


Fig. 4. Soft limiter model characteristics

The inner integral can be integrated by parts to give

$$F = \frac{2BV}{\pi^{1/2}} \int_{-\pi}^{\pi} \cos^2\phi \exp\{-B^2(V\sin\phi + n)^2\}\, d\phi \qquad (12)$$

which, when inserted back into the expression for c, produces the relation

$$c_1 = \frac{2^{1/2}BLV}{\sigma_N \pi^2} \int_{-\pi}^{\pi} \cos^2\phi \int_{-\infty}^{+\infty} \exp\left\{-\left[\left(B^2 + \frac{1}{2\sigma_N^2}\right)n^2 + 2B^2Vn\sin\phi + B^2V^2\sin^2\phi\right]\right\} dn\, d\phi \qquad (13)$$

The inner integral is tabulated (Ref. 4):

$$\int_{-\infty}^{+\infty} \exp\{-(at^2 + 2bt + c)\}\, dt = \left(\frac{\pi}{a}\right)^{1/2} \exp\left(\frac{b^2 - ac}{a}\right) \qquad (14)$$

Mere substitution thus provides

$$c_1 = L\left(\frac{2}{\pi}\right)^{1/2} \frac{v}{\pi} \int_{-\infty}^{+\infty} \cos^2\phi \exp\left\{-\frac{1}{2}v^2\sin^2\phi\right\} d\phi \qquad (15)$$

in terms of the parameter ratio

$$v^2 = \frac{2B^2V^2}{1 + 2B^2\sigma_N^2} = \frac{V^2}{\sigma_N^2 + \dfrac{2}{\pi}\left(\dfrac{L}{K}\right)^2} \qquad (16)$$

But the form of c_1 is now recognized to involve the same integral as that of the hard limiter, except with a different v. As a consequence, the results for a soft limiter are expressible in terms of the hard limiter suppression factor α^2. For example, the device output power P_s, considering $V(t) = V$ as a constant, is

$$P_s = \frac{8}{\pi^2} L^2 \alpha^2 \left[\frac{\rho}{1 + \left(\dfrac{2L}{\pi^{1/2}VK}\right)^2 \rho}\right] \qquad (17)$$

and, if $V(t)$ is time varying, P_s is

$$P_s = \frac{8}{\pi^2} L^2 E\left\{\alpha^2 \left[\frac{\dfrac{V^2(t)}{2\sigma_N^2}}{1 + \dfrac{2}{\pi}\left(\dfrac{L}{K\sigma_N}\right)^2}\right]\right\} \qquad (18)$$

Here again we can define a signal suppression factor α_s^2 as the ratio of signal output powers with and without noise. Because of the last equation above, we see this can be written as

$$\alpha_s^2 = \frac{\alpha^2\left[\dfrac{\rho}{1+\left(\dfrac{2L}{\pi^{1/2}VK}\right)^2\rho}\right]}{\alpha^2\left[\left(\dfrac{\pi^{1/2}VK}{2L}\right)^2\right]} \tag{19}$$

in the simpler, constant-V case. This function appears plotted in Fig. 5 for various values of VK/L.

Note in the limiting case

$$\left(\frac{VK}{L}\right) \to \infty \qquad \text{(approaching a hard limiter)}$$

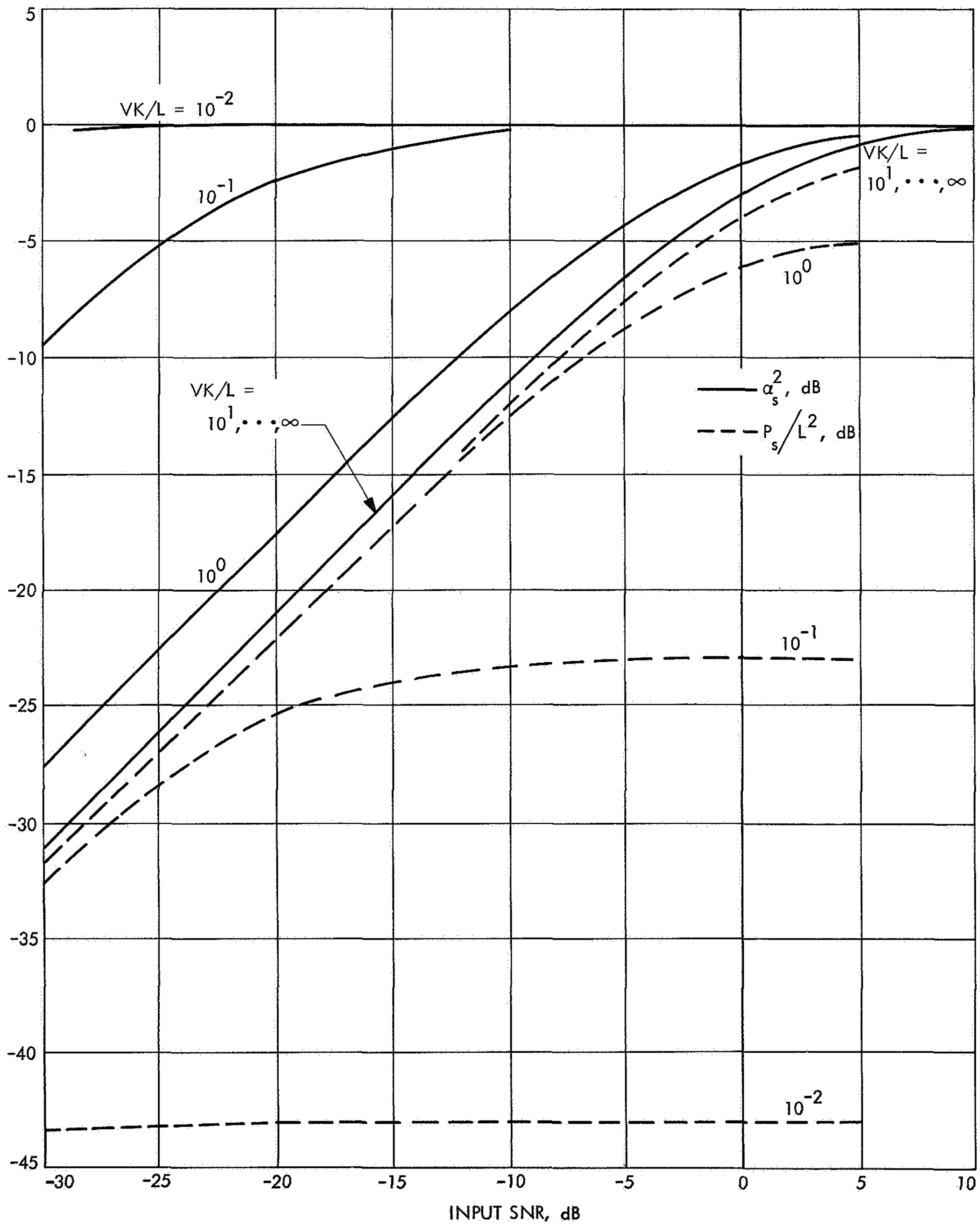


Fig. 5. Suppression factor α_s^2 and normalized output signal power P_s/L^2 as functions of input SNR

that

$$\alpha_s^2 \rightarrow \alpha^2 \qquad (20)$$

as it should, and as

$$\left(\frac{VK}{L}\right) \rightarrow 0 \qquad \text{(approaching a linear amplifier)}$$

for a fixed K, that

$$P_s \rightarrow E\left[\frac{K^2V^2(t)}{2}\right] = K^2 P_{sig} \qquad (21)$$

as it should.

4. Noise Output Power

In the linear region, the output SNR equals that input to the device

$$\rho_s \rightarrow \rho \qquad \text{(linear region)} \qquad (22)$$

whereas, when severe clipping is taking place,

$$\rho_s \rightarrow \rho_l = \frac{\alpha^2(\rho)}{1-\alpha^2(\rho)} \qquad \text{(limiting region)} \qquad (23)$$

(considering now only the constant-V case). The crossover between these two conditions begins at the point when the input begins to saturate.

Considering that the noise may be decomposed into independent in-phase and quadrature-phase terms

$$n(t) = n_c \cos\phi + n_s \sin\phi \qquad (24)$$

in which $\sigma_c^2 = \sigma_s^2 = \sigma_N^2$, then it is immediate that $x(t)$ takes the form

$$x(t) = V_{eq}(t) \sin\phi_{eq} \qquad (25)$$

with the amplitude function

$$V_{eq}^2(t) = (V + n_s)^2 + n_c^2 \qquad (26)$$

The amplitude of the output fundamental term is

$$a_1 = \frac{1}{\pi}\int_{\pi}^{\pi} \text{erf}\,(BV_{eq}\sin\phi)\sin\phi\, d\phi \qquad (27)$$

which can be integrated by parts to produce

$$a_1 = \left(\frac{2}{\pi}\right)^{1/2} (2^{1/2}BV_{eq}) \exp\left(\frac{-B^2V_{eq}^2}{2}\right) \times\left[I_0\left(\frac{B^2V_{eq}}{2}\right) + I_1\left(\frac{B^2V_{eq}}{2}\right)\right] \qquad (28)$$

This expression is the same as c_1 for the signal output portion only, except for the substitution $v = 2^{1/2}BV_{eq}$. The total limiter output power is

$$P_{s+n} = \frac{1}{2}E(a_1^2) = \frac{8}{\pi^2}E\left[\alpha^2(B^2V_{eq}^2)\right] \qquad (29)$$

Asymptotically for very large and very small values of $B^2V_{eq}^2$, the value of P_{s+n} behaves as

$$P_{s+n} \sim \frac{8}{\pi^2}L^2\alpha^2\left[B^2E(V_{eq}^2)\right] = \frac{8L^2}{\pi^2}\alpha^2\left[\left(\frac{\pi^{1/2}KV}{2L}\right)^2\frac{(1+\rho)}{\rho}\right] \qquad (30)$$

Thus an asymptotically correct approximate expression for the output SNR of the device is

$$\rho_l = \frac{\alpha^2\left[\dfrac{\rho}{1+\left(\dfrac{2L}{\pi^{1/2}VK}\right)^2\rho}\right]}{\alpha^2\left[\dfrac{(1+\rho)}{\left(\dfrac{2L}{\pi^{1/2}VK}\right)^2\rho}\right] - \alpha^2\left[\dfrac{\rho}{1+\left(\dfrac{2L}{\pi^{1/2}VK}\right)^2\rho}\right]} \qquad (31)$$

Finally, of interest is the ratio Γ_s of the input and output signal-to-noise densities at the fundamental frequency; this function is needed when the limiter output filter is considerably narrower than the bandwidth of the input process. It is clear that in the linear region, the SNR is preserved so that

$$\frac{N_l}{P_s} = \frac{N_0}{P_{sig}} \qquad \text{(linear region)} \qquad (32)$$

i.e., $\Gamma_s = 1$. At the other extreme, it has been shown (Ref. 5) that

$$\frac{N_0}{P_s} = \frac{\Gamma(\rho)N_0}{P_{sig}} \qquad \text{(limiting region)} \qquad (33)$$

where $\Gamma(\rho)$ is approximately

$$\Gamma(\rho) = \frac{1+\rho}{0.862+\rho} \tag{34}$$

In the transition region, Γ_s lies somewhere between 1 and Γ. Thus a simple asymptotic approximation to the true behavior can be expressed in the form

$$\Gamma_s = \frac{1 + aP_{\text{in}}\Gamma}{1 + aP_{\text{in}}} \tag{35}$$

in which the parameter a can be chosen to make a good fit in the transition region. To match the same type of crossover that we notice between P_s and P_{s+n}, we can take

$$a = \left(\frac{\pi^{1/2}K}{2L}\right)^2$$

to provide

$$\Gamma_s = \frac{\rho + \left(\frac{\pi^{1/2}KV}{2L}\right)^2 (1+\rho)\,\Gamma(\rho)}{\rho + \left(\frac{\pi^{1/2}KV}{2L}\right)^2 (1+\rho)} \tag{36}$$

5. Conclusions

Depending on the parameter VK/L, the soft limiter performs in varying degrees between the characteristics of a linear amplifier and a hard limiter. The performance parameters are furthermore expressible in terms of the hard-limiter suppression function under a change of variables. Such parameters include the output signal and noise powers, signal suppression factor, output SNR, and output signal-noise-density ratio.

References

1. Galejs, J., "Signal-to-Noise Ratios in Smooth Limiters," *IEEE Trans. Inf. Theory*, Vol. IT-1, June 1959, pp. 79–85.
2. Baum, R. F., "The Correlation Function of Smoothly Limited Gaussian Noise," *IEEE Trans. Inf. Theory*, Vol. IT-3, Sept. 1959, pp. 193–197.
3. Deutch, R., *Nonlinear Transformations of Random Processes*, Prentiss Hall, Inc., 1962, pp. 24–25.
4. "Handbook of Mathematical Functions," National Bureau of Standards, *Appl. Math Ser. 55*, June 1964.
5. Tausworthe, R. C., *Theory and Practical Design of Phase-Locked Receivers*, Technical Report 32-819, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 16, 1966.