THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 5, May-June, 1971
*Printed in U.S.A.*

# Timing Recovery in PAM Systems

By R. D. GITLIN and J. SALZ

(Manuscript received January 7, 1971)

*It is shown how various timing recovery schemes are reasonable approximations of the maximum likelihood strategy for estimating an unknown timing parameter in additive white gaussian noise. These schemes derive an appropriate error signal from the received data which is then used in a closed-loop system to change the timing phase of a voltage-controlled oscillator. The technique of stochastic approximation is utilized to cast the synchronization problem as a regression problem and to develop an estimation algorithm which rapidly converges to the desired sampling time. This estimate does not depend upon knowledge of the system impulse response, is independent of the noise distribution, is computed in real time, and can be synthesized as a feedback structure. As is characteristic of stochastic approximation algorithms, the current estimate is the sum of the previous estimate and a time-varying weighted approximation of the estimation error. The error is approximated by sampling the derivative of the received signal, and the mean-square error of the resulting estimate is minimized by optimizing the choice of the gain sequence.*

*If the receiver is provided with an ideal reference (or if the data error rate is small) it is shown that both the bias and the jitter (mean-square error) of the estimator approach zero as the number of iterations becomes large. The rate of convergence of the algorithm is derived and examples are provided which indicate that reliable synchronization information can be quickly acquired.*

## I. INTRODUCTION

The problem of symbol synchronization in digital data transmission in the presence of intersymbol interference is extremely complicated. The best sampling instants are channel dependent and are in general difficult to determine. Consequently, the problem of timing recovery in high-speed data transmission is intimately tied in with adaptive equalization. Since general methods for simultaneous optimum determination of the receiver parameters are not known, these parameters are independently determined.

Timing information is usually obtained directly from the data wave in a variety of ways.[1–3] Our objectives in this paper are:

(*i*) To indicate the optimum method (maximum likelihood) for estimating an unknown timing parameter from random data for a certain class of PAM data transmission systems;
(*ii*) To show that a variety of timing recovery methods currently in use are reasonable approximations of the optimum method, and to note that the generation of an error signal from the received signal is a feature common to these methods;
(*iii*) To demonstrate that timing recovery dynamics can often be studied and controlled through the application of stochastic approximation theory.[4–6]

Identifying the desired timing parameter as the solution of a regression equation will allow us to apply stochastic approximation theory to the symbol synchronization problem. For purposes of illustration we analyze a stochastic approximation timing recovery procedure for square-wave modulation. For this example we derive asymptotic formulas for the probability of error as a function of signal-to-noise ratio and the number of iterations used in the timing recovery loop. Since the number of iterations is directly proportional to the number of signaling intervals, insight is provided into the setup time required to achieve reliable symbol synchronization.

We finally focus on the more difficult problem of timing recovery in bandlimited PAM systems. Here timing information must be obtained in the presence of intersymbol interference as well as additive noise. A stochastic approximation algorithm is presented which derives symbol synchronization (i.e., estimates the desired sampling time) from the received data in a quick and accurate manner. The estimation algorithm developed does not require explicit knowledge of the system impulse response or the noise distribution. If the impulse response of the channel satisfies certain conditions, then the algorithm will converge in mean-square provided the gain sequence is properly chosen. Symbol synchronization is obtained by adjusting the sampling time in the following manner: at the end of each symbol interval the current estimate is taken to be the sum of the previous estimate and a weighted approximation to the actual estimation error. The desired sampling time is assumed to be that instant when the system impulse response is a maximum. For this sampling time it is shown that a reasonable approximation to the estimation error is the sampled derivative of the received signal.† When the error is small, its evolution can be described by a first-order random difference equation. At every iteration the mean-square error (mse) can be minimized by optimizing the choice of the (time-varying) weighting sequence. The optimum weighting sequence is of the form $1/(\alpha + \beta n)$, where $\alpha$ and $\beta$ are quantities which depend on the system impulse response and noise power, and $n$ is the discrete time index. Since $\alpha$ and $\beta$ are generally unknown at the receiver they may either be estimated (giving rise to an adaptive synchronization algorithm) or picked arbitrarily. In an effort to overcome the lack of knowledge of $\alpha$ and $\beta$ (in addition to simplifying the algorithm) it is tempting to use the asymptotic form of the gain $c/n$, where $c$ is a constant. However, if $\beta \ll \alpha$ then the optimum gain is essentially a constant $(1/\alpha)$ for many iterations, and for a wide range of $c$ the estimate obtained using $c/n$ is shown to be unreliable. Hence it appears that in order to obtain satisfactory performance some adaptivity to determine $\alpha$ and $\beta$ should be used in any realization of the algorithm.

Under the assumptions that the receiver error rate is small (so that an ideal reference can be assumed) and that the "eye" of the differentiated impulse response is open, the optimum mse is asymptotically of the form $1/\rho n$, where $\rho$ is a "signal-to-noise" ratio. The "signal" term is the value of the slope of the differentiated impulse response near the origin, the "noise" term is the sum of the actual noise variance and two intersymbol interference type terms. Thus the mse can be driven to zero and an example is given to illustrate how an accurate estimate can be obtained in a few signaling intervals. We show that for a $\sin x/x$ impulse response, ten iterations will drive the mean-square error to less than 0.01 of a signaling interval.

In Section II we determine the maximum likelihood estimate of an unknown timing parameter for a baseband PAM data signal which has been contaminated by white gaussian noise. Several approximations to the optimum estimator are described in Section III. The theory of stochastic approximation is introduced in Section IV, and is used both to cast the synchronization problem as a regression problem,

---

† B. R. Saltzberg[7] has suggested a technique for timing recovery which uses this approximation. His investigation is restricted to algorithms which can be realized using time-invariant devices. The algorithm we develop exploits the advantages of using time-varying elements.

and to analyze and control the dynamics of timing recovery. In Section V we discuss a timing recovery algorithm for bandlimited PAM.

## II. THE MAXIMUM LIKELIHOOD ESTIMATOR OF AN UNKNOWN TIMING PARAMETER

Consider the $L$ level data wave in additive white gaussian noise $\nu(t)$ of double-sided spectral density $N_0$,

$$V(t) = \sum_n a_n h(t - nT - \tau^*) + \nu(t), \tag{1}$$

where $\{a_n\}$ are the data symbols taking on values $\pm d, \pm 3d, \cdots \pm (L-1)d$ with equal probability, $h(t)$ is a bandlimited pulse whose peak value occurs at $\tau^*$, and $-T/2 \leqq \tau^* \leqq T/2$ is an unknown timing parameter.†

Detection of the data symbols $\{a_n\}$ is usually accomplished by first suitably filtering $V(t)$ and then sampling the output at time instants $\tau + kT$, $k = \pm 1, \pm 2, \cdots$. The resulting error rate is a function of $\tau$ in addition to other parameters. An ideal timing recovery system would supply the detector with $\tau$ which minimizes the probability of error. While this problem is conceptually straightforward, it is not analytically tractable and the structure of such an optimum timing recovery system is not generally yet known. We therefore must resort to a less utopian criterion.

Much simpler evaluation functions often used in data transmission[8] are

$$D_j(\tau - \tau^*) = \frac{1}{|\, h(\tau - \tau^*) \,|^j} \sum_{\substack{k \\ k \neq 0}} |\, h(\tau - \tau^* - kT) \,|^j \qquad j = 1 \quad \text{or} \quad 2. \tag{2}$$

Even for these relatively simple evaluation functions it is generally difficult to find the optimum $\tau$. R. W. Chang[9] derives timing recovery procedures based on minimizing a particular version of equation (2). However, for a certain class of linear distortions, namely the type that gives rise to symmetrical pulse shapes, the best $\tau$, which minimizes (2), is equal to the unknown parameter $\tau^*$. For this class of channels the problem of optimal timing recovery procedures can be cast in the language of statistical estimation theory. This is the situation treated in this section.

---

† We assume throughout that $\tau^*$ is independent of time.

The statistical problem we pose is this: determine an estimation procedure for the parameter $\tau$ based on observations made on the received signal $V(t)$ [equation (1)]. The more detailed question we wish to answer is the following. How should the observed signal, say for $T_s$ seconds, be processed such that a "good" estimate of $\tau^*$ is obtained? The answer of course depends on what one means by good. A reasonable measure of goodness is to require that the estimate maximize the likelihood function of the unknown parameter. For binary transmission this is a classical problem for which a solution is known. (See for example Ref. 3, 10, and 11.)† The extension to multilevel signaling is straightforward and we now briefly sketch the derivation. The likelihood function of the received signal is proportional to (superfluous constants are omitted)

$$L[V] \sim E\left\{\exp - \frac{1}{2N_0} \int_0^{T_s} [V(t) - s(t;\tau)]^2 \, dt\right\}_{\mathbf{a}}, \tag{3}$$

where $s(t;\tau) = \sum a_n h(t - nT - \tau)$ and $E\{\cdot\}_{\mathbf{a}}$ denotes expectation with respect to the data symbols. The expectation indicated in (2) can be carried out provided the reasonable assumption is made that the power in the data signal $s(t;\tau)$ when measured over an interval $[0, T_s]$ (large compared with a symbol duration) is independent of the data sequence and the unknown parameter $\tau$. This assumption leads to a simplified version of (3)

$$L[V] \sim E\left\{\exp\left\{\frac{1}{N_0} \int_0^{T_s} V(t)s(t;\tau)\, dt\right\}\right\}_{\mathbf{a}} \tag{4}$$

$$L(V) \sim \prod_n \left\{\frac{2}{L} \sum_{\substack{k=1 \\ k \text{ odd}}}^{L-1} \cosh\left(\frac{kd}{N_0} z_n(\tau)\right)\right\}, \tag{5}$$

where

$$z_n(\tau) = \int_0^{T_s} V(t)h(t - nT - \tau)\, dt \tag{6}$$

is recognized as the sampled (at times $nT + \tau$) output of a filter matched to $h(t)$, whose input is $V(t)$.

The maximum likelihood estimate (MLE) is obtained by differentiating $L[V]$ with respect to $\tau$ and setting the resulting expression to zero. An equivalent strategy may be obtained by differentiating any monotonic function of $L$ and a convenient such function in this appli-

† None of the references cited claims originality. It is difficult to determine where the result was written down first.

cation is the logarithmic function. From equation (5)

$$\Lambda[V] = \ln L[V] \sim \sum_n \left\{ \ln \left[ \sum_{\substack{k=1 \\ k \text{ odd}}}^{L-1} \cosh \left( \frac{kd}{N_0} z_n(\tau) \right) \right] \right\}, \tag{7a}$$

and upon differentiation we obtain

$$\frac{\partial \Lambda}{\partial \tau} = \sum_n \left\{ \frac{\sum_{k=1}^{L/2} (2k-1) \sinh \left( \frac{(2k-1)\, d}{N_0} z_n(\tau) \right)}{\sum_{k=1}^{L/2} \cosh \left( \frac{(2k-1)\, d}{N_0} z_n(\tau) \right)} \right\} \frac{d}{N_0} \frac{dz_n(\tau)}{d\tau}, \tag{7b}$$

where the bracketed term can be shown to be†

$$\frac{(L-1) \sinh \left( \frac{(L+1)\, d}{N_0} z_n(\tau) \right) - (L+1) \sinh \left( \frac{(L-1)\, d}{N_0} z_n(\tau) \right)}{\cosh \left( \frac{(L+1)\, d}{N_0} z_n(\tau) \right) - \cosh \left( \frac{(L-1)\, d}{N_0} z_n(\tau) \right)}; \tag{7c}$$

and for the typical data communication environment of a large signal-to-noise ratio the above expression becomes proportional to

$$(L-1) \tanh \left( \frac{(L+1)\, d}{N_0} z_n(\tau) \right).$$

Thus we finally have that

$$\frac{\partial \Lambda}{\partial \tau} \sim \sum_n \frac{dz_n(\tau)}{d\tau} \tanh \left( \frac{(L+1)\, d}{N_0} z_n(\tau) \right). \tag{8}$$

The optimum estimation strategy is exhibited in equation (8). The best value of $\tau$ (i.e., the MLE) makes the right-hand side of equation (8) as small as possible. The mathematical operations exhibited in equation (8) can readily be instrumented. The implementation objective would be to use the right-hand side of (8) as an error signal in a closed-loop system that iteratively adjusts $\tau$ to determine the MLE. A block diagram of this implementation is shown in Fig. 1. The received signal and its derivative are first passed through filters with identical impulse responses $h(-t)$ whose outputs are periodically sampled at times $nT + \tau$. In the undifferentiated branch, the samples are first multiplied by $(L + 1)d/N_0$ and are then passed through the memoryless nonlinearity tanh $(\cdot)$ which resembles an infinite clipper for large input values. The output from the two branches are mul-

† Note that for $L = 2$, equation (7c) becomes $(\sinh 3y - 3\sinh y)/(\cosh 3y - \cosh y) = \tanh y$, which agrees with the bracketed term in (7b).

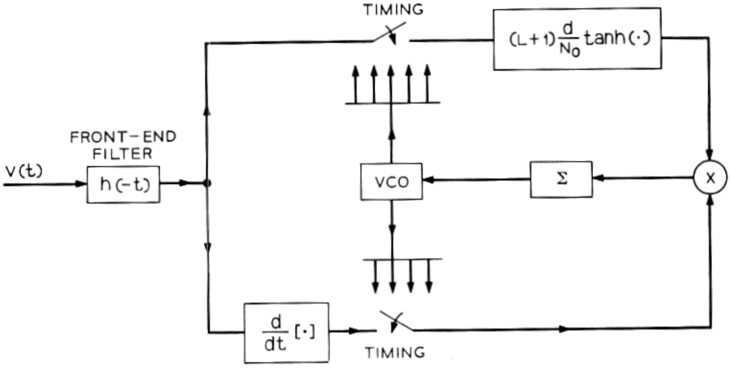


Fig. 1—Implementation of maximum likelihood strategy.

tiplied and averaged as indicated by the sum in equation (8). This then is the error signal driving a voltage-controlled oscillator which in turn determines the new timing phase.

## III. IMPLEMENTATIONS APPROXIMATING THE OPTIMUM

We now examine approximations of equations (7) and (8) leading to several simplified implementations of timing recovery systems. The first approach is to approximate $\tanh(x)$ in equation (8) by the limiter function $\text{sgn}(x)$. This approximation yields

$$\tanh\left(\frac{(L+1)\,d}{N_0}\,z_n(\tau)\right) \sim \text{sgn}\, z_n(\tau) \equiv \text{sgn}\, \hat{a}_n\,, \tag{9}$$

where $\hat{a}_n$ is the $n$th decision, or the estimate of the $n$th data symbol. The approximation (9) is a good one at large signal-to-noise ratio and in this case $\hat{a}_n$ will equal $a_n$ most of the time. When this approximation is made, the detection circuit which computes $\hat{a}_n$ from $z_n(\tau)$ is separated from the timing circuit. In the timing branch the received signal is first passed through the filter with impulse response $h(-t)$ and the output is differentiated or equivalently passed through a high-pass filter and then sampled. These samples are multiplied by the sign of the respective decisions and summed to form an error signal. The multiplication of the respective derivative samples by the sign of the decisions is clearly necessary so as to convert all the error samples to the same polarity.† Figure 2 shows this simplified version of detec-

† This is a decision directed estimation procedure. As the timing phase is acquired, the decisions become more reliable.

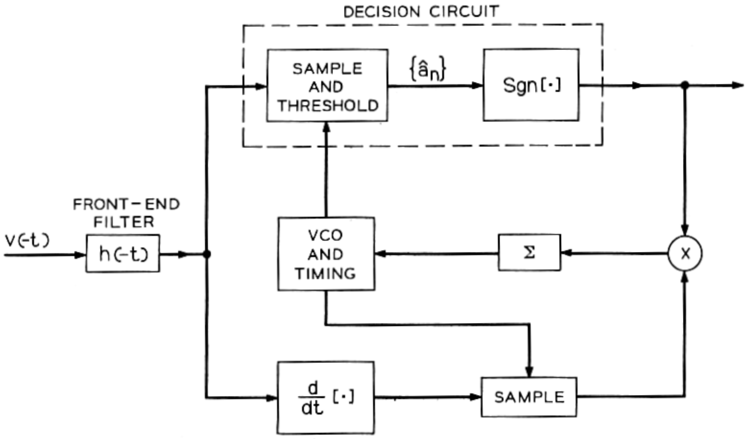


Fig. 2—An implementation approximating the ideal.

tion and timing recovery circuit. Deriving an error signal from the derivative of the received signal is very reasonable and a timing circuit based on this idea has been built and analyzed by Saltzberg.[7]

Another technique suggested from (7) is dubbed "early-late" timing recovery.[2,11] The approximations involved here are the following. First the derivative of $\Lambda[V]$ is approximated by the difference

$$\sum_k \{\ln \cosh ((kd/N_0)z_n(\tau + \Delta)) - \ln \cosh ((kd/N_0)z_n(\tau - \Delta))\}$$

$$\Delta \ll T. \qquad (10)$$

Next the nonlinear function ln [cosh $(x)$] is approximated by $|x|$. This again is a good approximation at large signal-to-noise ratio since for large $|x|$, cosh $x \to e^{|x|}$. This implementation is shown in Fig. 3. Here two clock pulses separated by $2\Delta$ sample the received wave after appropriate filtering. The respective samples are then full-wave rectified and substracted from one another. The error signal is formed by adding a number of successive differences. It appears that any even $N$th-law device may be used in place of the ln(cosh) nonlinearity in equation (7). Successful results for instance were obtained with a square-law device.[12]

A feature common to the above timing recovery systems is the generation of an error signal from the received signal. The sampling instant is then adjusted so as to decrease the magnitude of the error, a new error is computed, and the estimation continues in this manner.

The fewer the number of iterations needed to obtain a reliable estimate, the better the system. Stochastic approximation is a technique which will enable us to study and control the dynamic behavior of such iterative estimation algorithms by viewing the synchronization problem as a regression problem.

## IV. THE APPLICATION OF STOCHASTIC APPROXIMATION TO SYMBOL SYNCHRONIZATION

### 4.1 *Stochastic Approximation*

We will briefly describe the salient features of stochastic approximation, in particular the Robbins-Monro algorithm. Stochastic approximation[4–6] is a technique employed to iteratively solve regression problems. The method is an extension of the Newton-Raphson technique to a random environment, and is especially useful when the regression function is unknown. More precisely, suppose $z_n$ is a sequence of independent observations of a stationary random process and it is desired to find the value of the (non-random) parameter $\tau$ such that the regression equation,

$$E[f(z_n\,;\,\tau)] \triangleq m(\tau) = m_o\,, \tag{11}$$

is satisfied; where $E$ denotes expectation, $f(\cdot)$ is a given function, and $m(\cdot)$ is called the regression function. As mentioned above, $m(\cdot)$ is typically unknown, and we desire an algorithm which uses the data to sequentially estimate the value of $\tau$, say $\tau^*$, which satisfies (11). Robbins and Monro have shown that if (11) has a unique solution

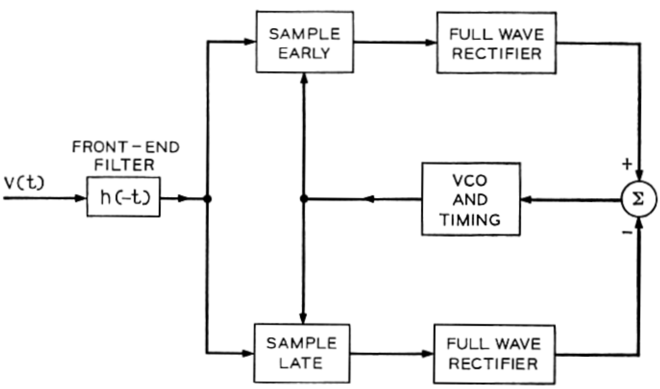


Fig. 3—Implementation of early-late timing recovery scheme.

then the estimate $\tau_n$, given by

$$\tau_{n+1} = \tau_n + c_n[f(z_n ; \tau_n) - m_o] \qquad n = 1, 2, \cdots ,$$

will converge in mean-square and with probability one to $\tau^*$, under some general conditions[4] on both the observations $z_n$ and on the positive scalar time-varying weighting sequence $c_n$. A useful interpretation of the Robbins-Monro algorithm is that the current estimate is the sum of the previous estimate and a weighted correction term, where the average (with respect to the observations) correction is the error term $m(\tau_n) - m_o$. Thus the correction term will, on the average, give an increment in the correct direction, and the estimate will converge. Alternatively, if we regard the correction term as an approximation (in a stochastic sense) to an error term, we are reminded of the deterministic error or gradient search type of algorithms. The weighting sequence $c_n$ is chosen to converge to zero fast enough so as to suppress the correction term as the estimate converges,† but slow enough so that large corrections are possible for many iterations (frequently $c_n$ is of the form $1/n$).

We now cast the synchronization problem as a regression problem, and then use the theory of stochastic approximation to develop a synchronization algorithm which has desirable dynamic properties. From (8) the optimum (maximum likelihood) timing parameter is the solution of

$$\frac{\partial}{\partial \tau} [\Lambda(z_n ; \tau)] = 0.$$

If we make the identification

$$\frac{\partial}{\partial \tau} [\Lambda(z_n ; \tau)] \leftrightarrow f(z_n ; \tau), \tag{12a}$$

and now ask for the value of $\tau$ which satisfies

$$m(\tau) \equiv E\left[\frac{\partial}{\partial \tau} \Lambda(z_n ; \tau)\right] = 0, \tag{12b}$$

then the desired [i.e., the solution of (12b)] timing parameter will be the solution of a regression equation. It is important to note that the solutions of (8) and (12b) will not, in general, be the same. However the solution of (8) is a random variable, which as the observation

† Note that even when $\tau_n$ is close to $\tau^*$, the variance of the correction term can be quite large due to the randomness of the data.

time $T_s$ becomes large converges to $\tau^*$; while the solution† of (12b) is in fact $\tau^*$. Thus if we use a Robbins-Monro algorithm to iteratively solve (12b) we are indeed generating the maximum likelihood estimate.

### 4.2 *Binary Square-Wave Modulation*

Consider applying this method to analyze a timing recovery procedure when $h(t)$ in equation (1) is a rectangular pulse of $T$ seconds duration and height $A$, where binary transmission is assumed for convenience. In this case, the observable function, equation (6), becomes

$$z_n(\tau) = \int_0^{T_s} V(t)h(t - nT - \tau)\,dt = \int_{nT+\tau}^{(n+1)T+\tau} V(t)\,dt. \tag{13}$$

As mentioned earlier, we can use a square-law device to approximate the ln cosh $(\cdot)$ nonlinearity for mathematical convenience. Thus the MLE is obtained by finding a $\tau$ such that the derivative of $\sum_n z_n^2(\tau)$ is zero. From (7a) and (13) we obtain

$$\frac{d}{d\tau}\sum_n z_n^2(\tau) = 2\sum_n [V((n+1)T+\tau) - V(nT+\tau)]z_n(\tau). \tag{14}$$

At large signal-to-noise ratio, symbol transition information is obtained from

$$d_n = V((n+1)T+\tau) - V(nT+\tau) \sim \begin{cases} 0, & a_{n+1}\cdot a_n = 1 \\ \pm 1, & a_{n+1}\cdot a_n = -1 \end{cases}. \tag{15}$$

The Robbins-Monro procedure for recursively estimating $\tau$ can now be applied by using the regression function

$$m(\tau) = E\{d_n z_n(\tau)\}. \tag{16}$$

For convenience we center the pulse $h(t)$ at $t = 0$ such that

$$h(t) = \begin{cases} A, & |t| \leqq T/2 \\ 0, & \text{elsewhere} \end{cases}.$$

and calculate

$$d_n z_n(\tau) = d_n A \sum_m a_m \int_{nT+\tau}^{(n+1)T+\tau} h(t - mT)\,dt + \int_{nT+\tau}^{(n+1)T+\tau} \nu(t)\,dt$$

† For a high signal-to-noise ratio.

$$= d_n A\left\{a_n \int_{nT+\tau}^{nT+T/2} dt + a_{n+1} \int_{nT+T/2}^{(n+1)T+\tau} dt\right\} + \nu_n$$

$$= d_n A\{a_n(T/2 - \tau) + a_{n+1}(T/2 + \tau)\} + \nu_n , \tag{17}$$

where

$$\nu_n = \int_{nT+\tau}^{(n+1)T+\tau} \nu(t)\, dt.$$

In the absence of data transitions, (17) is independent of $\tau$ while when transition occurs, i.e., $a_n \neq a_{n+1}$ ,

$$d_n z_n(\tau) = 2A\tau + \nu_n , \qquad -T/2 \leqq \tau \leqq T/2. \tag{18}$$

Using (13) the recursive procedure for estimating the unknown timing parameter, $\tau$, is now as follows: Pick an arbitrary sampling phase $\tau_0$ , $|\tau_0| \leqq T/2$, and compute the next sampling phase $\tau_1$ from the relation (assuming that a data transition occurs)

$$\begin{aligned} \tau_1 &= \tau_0 - \tfrac{1}{2}(d_0 z_0(\tau_0)) \\ &= \tau_0 - \tfrac{1}{2}(2A\tau_0 + \nu_0). \end{aligned} \tag{19}$$

The $(n + 1)$th sampling phase is then related to the $n$th by the recursion relation

$$\tau_{n+1} = \tau_n - \frac{1}{n+2}\,(d_n z_n(\tau_n)), \tag{20}$$

where we have taken $c_n$ to be $1/n+1$. For numerical evaluation purposes it is convenient to normalize (18) and work with the regression function†

$$m(\tau_n) = E[f(x_n , \tau_n)] = E[\tau_n + x_n], \tag{21}$$

where $\{x_n\}$ is a sequence of gaussian random variables with

$$E\{x_n\} = 0 \tag{22}$$

and

$$E\{x_n^2\} = \frac{N_0 T}{4A^2} = \frac{T^2}{8\rho}$$

where $\rho = A^2/2N_0 1/T$ is the signal-to-noise ratio in a bit-rate bandwidth.

† We are assuming that a linear theory applies, i.e., the sequence of $\{\tau_n\}$ rarely exceeds $|T/2|$. In practice no values of $\tau_n$ which exceeds $|T/2|$ will be accepted. Including these restrictions in the mathematical model will render equations (19), (20), and (21) nonlinear and thus mathematically intractable.

Upon substituting (21) into (20) a linear recursion relation is obtained with the well-known solution

$$\tau_n = \frac{\tau_0}{n+1} - \frac{1}{n+1} \sum_{k=0}^{k=n-1} x_k \qquad |\tau_n| \leqq T/2. \tag{23}$$

By inspection the following pertinent parameters are computed

$$\mu_n = E[\tau_n] = \frac{\tau_0}{n+1} \to 0, \quad \text{as} \quad n \to \infty$$

and

$$\sigma_n^2 = E\{\tau_n - E^2[\tau_n]\} = \text{var } \tau_n = \frac{T^2}{8\rho} \frac{n}{(n+1)^2} \to 0, \quad \text{as} \quad n \to \infty. \tag{24}$$

In evaluating (24) we assumed that the sequence of random variables $\{x_n\}$ is independent. This is not strictly true. We see from (17) that the sequence of random variables $\{x_n\}$ for fixed $\tau$ is indeed independent since each $x_n$ represents nonoverlapping integrals of the white-noise process $\nu(t)$. However, as $\tau_n$ is changed according to equation (20) the noise integrals may overlap. To include this dependence in the analysis would render this seemingly simple problem untractable mathematically. Physically we feel, however, that this dependence is weak and therefore can be neglected.

From (23) we see that $\tau_n$ possesses a truncated gaussian probability density

$$\begin{aligned} P(\tau_n) &= p_1\, \delta(\tau_n - T/2) + p_2\, \delta(\tau_n + T/2) + G(\tau_n) \qquad |\tau_n| \leqq T/2 \\ &= 0 \qquad\qquad |\tau_n| > T/2 \end{aligned} \tag{25}$$

where

$$G(\tau_n) = \frac{1}{\sqrt{2\pi}\,\sigma_n} \exp\left\{-\frac{1}{2\sigma_n^2}(\tau_n - \mu_n)^2\right\}$$

and

$$p_1 = \int_{-\infty}^{-T/2} G(\tau_n)\, d\tau_n$$

$$p_2 = \int_{T/2}^{\infty} G(\tau_n)\, d\tau_n\,.$$

Using this probability density we can compute the system error rate.

Dispensing with tedious computational details, and focusing attention on essentials, we find that the conditional error rate (conditioned on the unknown parameter $\tau_n$) for this simple system is asymptotically (large signal-to-noise ratio)

$$P_e(\tau_n) \sim \exp - \left\{\frac{A^2[T - 2 \mid \tau_n \mid]^2}{2\sigma^2}\right\} \qquad \mid \tau \mid \leqq T/2, \tag{26}$$

where

$$\sigma^2 = N_0 T.$$

When $\tau_n = 0$, we have ideal performance, as we should. When $\tau_n = \pm T/2$, we have disaster. To obtain the actual error rate we must average (26) over the permissable values of $\tau_n$. This calculation yields

$$P_e = E\{P_e(\tau_n)\} \sim p_1 + p_2 + \int_{-T/2}^{T/2} P_e(\tau_n) G(\tau_n)\, d\tau_n . \tag{27}$$

The evaluation of (27) is straightforward. In terms of the normalized random variable $\alpha = \tau_n/T$, we express (26) in the form

$$P_e(\alpha) \sim e^{-(1-2|\alpha|)^2} \qquad |\alpha| < 1/2. \tag{28}$$

In terms of the same normalized variables and the explicit values of $\mu_n$ and $\alpha_n$ [equation (24)] we write

$$G(\alpha) \sim \exp\left[-4n\left(\alpha - \frac{1}{2n}\right)^2\right] \tag{29}$$

which is valid when $n$ is large. In writing down (29) we set $\tau_0 = T/2$ (a worst initial guess).

Asymptotically, $p_1$ and $p_2$ behave as $e^{-n\rho}$ and, as we shall see shortly, can be neglected compared with the last term in (27). To conclude the error rate calculation we evaluate

$$\int_{-T/2}^{T/2} P_e(\tau_n) G(\tau_n)\, d\tau_n = \xi_n(\rho) \sim \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\rho(1-2|\alpha|)^2 - \rho 4n(\alpha - 1/2n)^2}\, d\alpha$$

$$= \int_{-\frac{1}{2}}^{0} e^{-\rho E_1(\alpha)}\, d\alpha + \int_{0}^{\frac{1}{2}} e^{-\rho E_2(\alpha)}\, d\alpha, \tag{30}$$

where

$$E_1(\alpha) = (1 + 2\alpha)^2 + 4n\left(\alpha - \frac{1}{2n}\right)^2$$

and

$$E_2(\alpha) = (1 - 2\alpha)^2 + 4n\left(\alpha - \frac{1}{2n}\right)^2.$$

Using a saddle-point technique to obtain an asymptotic approximation for the integrals, we find that

$$P_e(\rho) \sim e^{-\rho M_1(n)} + e^{-\rho M_2(n)},$$

where

$$M_1(n) \sim 1 + \frac{1}{n}$$

and

$$M_2(n) \sim 1 - \frac{3}{n}. \tag{31}$$

Combining the above asymptotic results with $p_1$ and $p_2$ we obtain finally

$$P_e \sim \exp\left\{-\rho\left(1 - \frac{3}{n}\right)\right\} \tag{32}$$

for $n$ and $\rho$ large. All the other terms have exponents larger than (32) and therefore can be neglected. For example when $n = 30$, the degradation from ideal $(n \to \infty)$ is only 0.5 dB approximately.

What this example shows is that for square-wave modulations in the presence of additive white gaussian noise, bit timing can reliably be derived in approximately 30-bit intervals.

### 4.3 *Synchronization of Bandlimited PAM*

We now consider a timing recovery algorithm for a bandlimited PAM signal. As in the previous section the synchronization problem will be cast as a regression problem. Our received signal is given by equation (1)

$$V(t) = \sum_m a_m h(t - mT - \tau^*) + \nu(t), \tag{33}$$

and as before the objective of the synchronizer is to accurately and rapidly estimate $\tau^*$. In order to extract information about $\tau^*$ we low-pass filter, differentiate, and sample the received signal. Hence the error signal is similar to that shown in Fig. 2, with the matched filter replaced by a low-pass filter. Thus the receiver does not need knowledge of the pulse $h(t)$. If we denote the derivative of $h(\cdot)$ by $g(\cdot)$, then the differentiated and sampled received signal is given by

$$\begin{aligned} V'(kT + \tau) &= \sum_m a_m g[(k - m)T + \tau - \tau^*] + \nu(kT + \tau) \\ &= \sum_m a_m g_{k-m}(\tau - \tau^*) + \nu_k , \end{aligned} \tag{34}$$

where $\tau$ is an arbitrary sampling time such that $|\tau| < T/2$, $g_{k-m}$ denotes $g((k - m)T)$, and $\nu_k$ are samples† of the differentiated noise process $\nu(t)$. As before we let $\hat{a}_k$ denote the decision made at time $kT + \tau$. Assuming that the error rate is low enough so that with high probability $\hat{a}_k = a_k$, we then have that

$$\begin{aligned}\hat{a}_k V'(kT + \tau) &= \hat{a}_k a_k g_o(\tau - \tau^*) + \hat{a}_k \sum_{m \neq k} a_m g_{k-m}(\tau - \tau^*) + \nu_k \\ &= a_k^2 g(\tau - \tau^*) + \hat{a}_k \sum_{m \neq k} a_m g_{k-m}(\tau - \tau^*) + \nu_k, \qquad (35)\end{aligned}$$

where we have noted that $g_o(\tau - \tau^*) = g(\tau - \tau^*)$. If we further assume that $\hat{a}_k$ is uncorrelated with $a_j$,‡ for $j \neq k$, then averaging (35) gives

$$m(\tau) \triangleq E[\hat{a}_k V'(kT + \tau)] = \overline{a^2} g(\tau - \tau^*), \qquad (36)$$

where

$$\overline{a^2} = \frac{d^2}{3}(L^2 - 1).$$

Now for the typical impulse response $h(t)$ and its derivative $g(t)$, shown in Figs. 4 and 5, respectively, it is true that the (regression) equation

$$g(\tau - \tau^*) = 0, \qquad |\tau - \tau^*| \leqq T/2 \qquad (37)$$

has the unique solution

$$\tau = \tau^*. \qquad (38)$$

Since the synchronization problem has been modeled as a regression problem, we again use a Robbins-Monro algorithm to sequentially estimate $\tau^*$. Denoting the $k$th estimate by $\tau_k$, we have the modified Robbins-Monro algorithm

$$\tau_{k+1} = \begin{cases} \tau_k + c_k[\hat{a}_k V'(kT + \tau_k)], & |\tau_k + c_k[\hat{a}_k V'(kT + \tau_k)]| < T/2 \\ \tau_k, & \text{otherwise.} \end{cases} \qquad (39)$$

A feedback implementation of the above algorithm is shown in Fig. 6, with $D$ denoting a delay. It is again noted that the algorithm con-

† The dependence of the noise sample on the sampling offset $\tau$ is not shown, since it is assumed that the noise is stationary.

‡ As it will be if the $a_k$'s are independent and the receiver is supplied with an ideal reference.

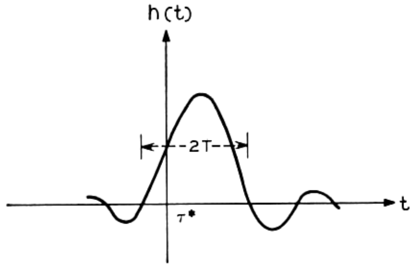


Fig. 4—A typical impulse response $h(t)$.

strains the estimate to a region of width $T$. This is consistent with the observation that any actual sampling instant will always be within $T/2$ seconds of the desired instant $\tau^*$, i.e., we may "slip" $T$ seconds but this is immaterial as far as estimating $\tau^*$ is concerned. It is by no means clear, *a priori*, that the above algorithm will converge rapidly or will converge at all. In fact the rest of this paper will consider the conditions which must be satisfied for the above algorithm to converge and the resulting rate of convergence.

## V. ANALYSIS OF THE SYNCHRONIZATION ALGORITHM

### 5.1 *The Error Equation*

In order to evaluate the proposed synchronization algorithm we will derive a difference equation for the mean-square estimation error $\overline{e_k^2}$, where

$$e_k \equiv \tau_k - \tau^*, \tag{40}$$

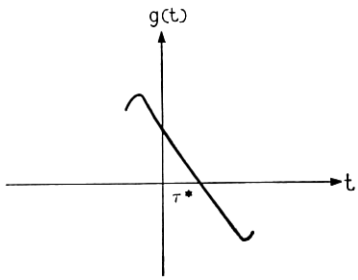


Fig. 5—The derivative of $h(t)$.

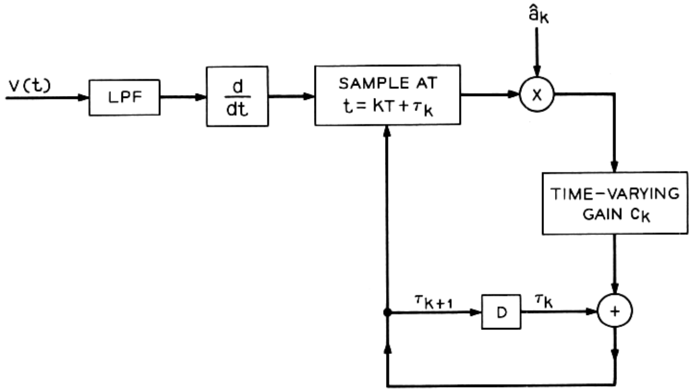


Fig. 6—A realization of the synchronization algorithm.

and the overbar denotes expectation.† In order to do this we see that from (39), and neglecting for the moment the constraining portion of the algorithm, we have

$$\begin{aligned} e_{k+1} &= \tau_{k+1} - \tau^* = \tau_k - \tau^* + c_k[\hat{a}_k V'(hT + \tau_k)] \\ &= \tau_k - \tau^* + c_k[g(\tau_k - \tau^*) + \hat{a}_k \sum_{m \neq k} a_m g_{k-m}(\tau_k - \tau^*) + \nu_k] \\ &= e_k + c_k[g(e_k) + \hat{a}_k \sum_{m \neq k} a_m g_{k-m}(e_k) + \nu_k]. \end{aligned} \tag{41}$$

We note that $g(\cdot)$ is such that, on the average, the error is decreased at each iteration, and once the estimation error is small‡ we need only keep first-order terms in a Taylor Series expansion of $g_{k-m}(e_k)$ about $(k - m)T$, i.e.,

$$g_{k-m}(e_k) \approx g_{k-m} + g'_{k-m}e_k \ , \tag{42}$$

where $g'_{k-m}$ denotes the derivative of $g(\cdot)$ evaluated at $(k - m)T$. Combining (41) and (42) yields the (approximate) first-order stochastic difference equation for the evolution of the error,

$$e_{k+1} = [1 + g'_0 c_k + c_k \hat{a}_k \sum_{m \neq k} a_m g'_{k-m}]e_k + c_k \hat{a}_k [\sum_{m \neq k} a_m g_{k-m} + \nu_k]. \tag{43}$$

Before studying the behavior of (43) we introduce the following

† We use the mean-square estimation error as a measure of performance. This is because the estimate is a nonlinear one, and thus the probability of error cannot be computed.

‡ Under this assumption we can certainly neglect the possibility that $\tau_{k+1} = \tau_k$.

notation:

$$g'_o = -\alpha \tag{44a}$$

$$\beta_k = c_k(g'_o + \hat{a}_k \sum_{m \neq k} a_m g'_{k-m}) \tag{44b}$$

$$\gamma_k = 1 + \beta_k \tag{44c}$$

$$Q_k = \hat{a}_k \sum_{m \neq k} a_m g_{k-m} + \nu_k \ , \tag{44d}$$

and using the above we rewrite (43) as

$$e_{k+1} = \gamma_k e_k + c_k Q_k \ . \tag{45}$$

Thus the error obeys a stochastic difference equation where the gain ($\gamma_k$) and the driving term ($Q_k$) are correlated. It is important to note that for the system described by (45) the probability density of the present error $e_k$ does not depend solely on a finite number of past data symbols, $a_k$ , but depends on all past and future values. This renders impossible an exact analysis of the mean-square error. However, if we assume that both $\gamma_k$ and $Q_k$ are independent sequences, then $e_k$ depends solely on past $\gamma_k$ and $Q_k$ , and we can obtain a bound on the mean-square error.† Squaring and averaging both sides of (45) gives

$$E[e_{k+1}^2] = E[\gamma_k^2 e_k^2] + 2c_k E[\gamma_k e_k Q_k] + c_k^2 E[Q_k^2]. \tag{46}$$

We now proceed to bound each of the terms on the right-hand side of (46). If we assume that the "eye" of the twice-differentiated impulse response is open, i.e.,

$$\alpha > \sum_{m \neq 0} | \ g'_m \ |, \tag{47}$$

then

$$\gamma_k = 1 - c_k(\alpha - \hat{a}_k \sum_{m \neq k} a_m g'_{k-m}) \leqq 1 - c_k(\alpha - \sum_{m \neq 0} | \ g'_m \ |) \tag{48a}$$

$$= 1 - c_k \beta, \tag{48b}$$

where $\beta$ denotes $\alpha - \sum_{m \neq 0} |g'_m|$. Using the above assumption, and the boundedness of the error, we have that

$$|E[\gamma_k Q_k e_k]| = |E[e_k] E[\gamma_k Q_k]| \leqq T/2 \ |E[\gamma_k Q_k]|, \tag{49}$$

† Despite much effort we have been unable to proceed without this assumption, but since the results which follow are intuitively satisfying and provide insight into this difficult problem they have been included in the paper.

and due to the independence of the data bits

$$E[\gamma_k Q_k] = E[((1 - c_k)\alpha + c_k \hat{a}_k \sum_{m \neq k} a_m g'_{k-m})\hat{a}_k(\sum_{i \neq k} a_i g_{k-i} + \nu_k)]$$
$$= c_k \sum_{m \neq 0} g'_m g_m = c_k 2/TG, \tag{50}$$

where† $G$ denotes $T/2 \sum_{m \neq 0} g'_m g_m$ . Finally we have

$$E[Q_k^2] = \sigma^2 + \sum_{m \neq 0} g_m^2 = \sigma^2 + P, \tag{51}$$

where $P$ denotes $\sum_{m \neq 0} g_m^2$ . Letting

$$\Delta_k = E[e_k^2], \tag{52}$$

and combining (46)–(52) we have the iterative bound

$$\Delta_{k+1} \leqq (1 - \beta c_k)^2 \Delta_k + c_k^2 M \tag{53}$$

on the mean-square error, where $M$ is the sum of $G$ and $\sigma^2 + P$. Although several assumptions have been made in obtaining (53) it is believed that the effect of the salient quantities upon the synchronization algorithm have been preserved. We now proceed to find the gain sequence which minimizes the bound of (53).

### 5.2 *The Optimum Gain Sequence*

We now find the sequence of gains, $c_k^*$, which minimize the right-hand side (RHS) of (53) for fixed $\Delta_k$ . Since we minimize a bound on the mean-square error at every iteration, this is a min-max procedure. We first find the optimum gain sequence in terms of $\Delta_k$ , and then by simultaneously iterating this equation and the bound of (53) we show that $c_k^*$ is proportional to $1/k$ for large $k$. We begin by setting to zero the derivative of the RHS of (53) with respect to $c_k$ , i.e.,

$$-\beta(1 - \beta c_k)\Delta_k + Mc_k = 0$$

or

$$c_k^* = \frac{\beta \Delta_k}{M + \beta^2 \Delta_k}. \tag{54}$$

Using (54) in (53) we have

$$\Delta_{k+1} \leqq \left(1 - \frac{\beta^2 \Delta_k}{M + \beta^2 \Delta_k}\right)^2 \Delta_k + M\left(\frac{\beta \Delta_k}{M + \beta^2 \Delta_k}\right)^2$$
$$\leqq \frac{M}{\beta} c_k^*, \tag{55}$$

† It should be noted that if $h(t)$ is an even function of time (with respect to the origin), then $g(t)$ and $g'(t)$ will be respectively odd and even time functions and $G$ will be zero.

or

$$\frac{c_{k+1}^*}{(1 - \beta c_{k+1}^*)} \leqq c_k^* . \tag{56}$$

Now if

$$(1 - \beta c_{k+1}^*) \geqq 0 \tag{57}$$

then we have the relation

$$c_{k+1}^* \leqq \frac{c_k^*}{1 + \beta c_k^*} , \tag{58}$$

which can be iterated to give†

$$c_k^* \leqq \frac{c_o^*}{1 + \beta c_o^* k} \tag{59a}$$

$$= \frac{\beta \Delta_o}{M + \beta^2 \Delta_o (k + 1)} , \tag{59b}$$

where

$$c_o^* = \frac{\beta \Delta_o}{M + \beta^2 \Delta_o} , \tag{60}$$

and $\Delta_o$ is the initial error variance. Henceforth we will interpret the sequence $c_k^*$, specified by (59) and (60) with the inequality replaced by an equality, as the optimum gain sequence. Combining (55), (59) and (60) we see that the mean-square error is bounded by

$$\Delta_k^* \leqq \frac{M \Delta_o}{M + \beta^2 \Delta_o k} \tag{61a}$$

which for large $k$ becomes

$$\Delta_k^* \leqq \frac{M}{\beta^2} \cdot \frac{1}{k} . \tag{61b}$$

Thus we see that asymptotically the minimized mean-square error is bounded by a term which decays as $1/k$, and is inversely proportional to signal-to-noise type ratio $(\beta^2/M)$.

The optimum gain, as given by (59b), depends upon the parameters $\Delta_o$, $M$, and $\beta$. Since these quantities are generally unknown it is tempting to replace $c_k^*$ by its asymptotic (large $k$) value $1/\beta(k + 1)$. Caution must be exercised in making this approximation; since $M \gg \beta^2 \Delta_o$

† Note that $\beta c_k^* \leqq \beta c_o^*/(1 + \beta c_o^* k) \leqq 1$, thus satisfying (57).

implies that the optimum gain sequence is essentially constant for many iterations, substitution of a decaying sequence could lead to an unreliable estimate (we will consider this point in Section 5.4). However if† $\beta^2\Delta_o \gg M$, then $c_k^* \approx 1/\beta(k+1)$ and we have only one unknown parameter. A possibility is to replace $\beta$ by an estimate—techniques of this sort are called adaptive estimation procedures. We now sketch a particular adaptive scheme.

### 5.3 *An Adaptive Synchronization Algorithm*

We now give a method for recursively estimating $\beta$, which can then be incorporated in an adaptive synchronization scheme. Since

$$\beta = \alpha - \sum_{m \neq 0} | g'_m |,$$

we desire a function of the received data which has $\beta$ as its average value. We note that from (34) we have

$$E[\hat{a}_k V''(kT + \tau)] = g'(\tau - \tau^*) \approx -\alpha \tag{62a}$$

(where the approximation is for small $\tau - \tau^*$), and

$$E[\hat{a}_i V''(kT + \tau)] = g'_{i-k}(\tau - \tau^*) \approx g'_{i-k}\,. \tag{62b}$$

We can then estimate $\beta$ by using a recursive stochastic approximation algorithm of the type discussed in Section 4.1. Such a scheme would twice differentiate the incoming data and then multiply the data sample by as many of the previous decisions as there are significant nonzero samples in the impulse response. Since even an approximate analysis of the above algorithm is hopelessly complex, we will consider the effect of using a gain of the form $c/k$, where $c$ is a constant to be chosen.

### 5.4 *A Suboptimum Gain*

We consider the mean-square error, as given by (53), with $c_k = c/k$. This gain is chosen since the optimum gain is asymptotically of this form. Care must be taken in choosing $c$, since the mean-square error will be shown to be a sensitive function of this parameter. Iterating (53) gives

$$\Delta_{k+1} \leqq \prod_{i=0}^{k} (1 - \beta c_i)^2 \Delta_o + \sum_{i=0}^{k} \prod_{j=i+1}^{k} (1 - \beta c_j)^2 c_i^2 M. \tag{63}$$

---

† A condition one would expect to be satisfied in practice.

The inequality

$$1 - x \leqq e^{-x} \tag{64}$$

gives

$$\prod_{j=i+1}^{k} (1 - \beta c_j)^2 \leqq \exp\left(-2 \sum_{j=i+1}^{k} \beta c_j\right); \tag{65}$$

and noting that

$$\sum_{j=i+1}^{k} \beta \frac{c}{j} \approx \beta c \int_{i+1}^{k} \frac{1}{x}\, dx = \beta c \ln\left(\frac{k}{i+1}\right)$$

results in

$$\prod_{j=i+1}^{k} (1 - \beta c_j)^2 \leqq \left(\frac{i+1}{k}\right)^{2\beta c}. \tag{66}$$

We can see that the transient behavior of the mean-square error, which is specified by the first term on the RHS of (63), will be of the form $(1/k)^{2\beta c}$. The other component of the mean-square error will be (approximately) bounded by

$$\begin{aligned}\sum_{i=0}^{k} \left(\frac{i+1}{k}\right)^{2\beta c} M \frac{c^2}{i^2} &\leqq \frac{Mc^2}{k^{2\beta c}} \sum_{i=0}^{k} (i+1)^{2(\beta c-1)} \\ &\leqq \frac{Mc^2}{(1+k)^{2\beta c}} \frac{1}{2\beta c - 1} (1+k)^{2\beta c - 1} \\ &= \frac{Mc^2}{(2\beta c - 1)(1+k)},\end{aligned} \tag{67a}$$

which results in

$$\Delta_{k+1} \leqq \left(\frac{1}{k}\right)^{2\beta c} \Delta_o + \frac{Mc^2}{(2\beta c - 1)(1+k)} \tag{67b}$$

as a bound on the mean-square error. If $2\beta c > 1$, then for large $k$ the above bound becomes

$$\Delta_{k+1} \leqq \frac{Mc^2}{(2\beta c - 1)(1+k)} \tag{67c}$$

and the mean-square error will converge at the optimum rate $(1/k)$. It is seen that care must be taken in selecting $c$, since for $c \geqq 1/2\beta$ (i.e., for $2\beta c > 1$) the quantity $Mc^2/2\beta c - 1$ has a minimum† at $c =$

† With $c = 1/\beta$, $\Delta_k \leqq M/\beta^2\ 1/k$ which is the optimum asymptotic rate of convergence.

$1/\beta$, and is infinite at both $c = 1/2\beta$ and $c = \infty$. Thus a very small step size ($c \ll 1/2\beta$) will result in an mse which converges at a less than optimum rate, while large step sizes ($c \gg 1/2\beta$) will result in a mean-square error which, while converging at the optimum rate, may be quite large for many iterations. The sensitivity of the above bound with respect to "$c$" may make the use of an adaptive procedure (which estimates $\beta$) advisable.

### 5.5 *An Example*

Consider the (minimum bandwidth) pulse

$$h(t) = A \frac{\sin \pi W t}{\pi W t} \tag{68}$$

where $W = 1/T$. It is easy to show that

$$\beta = \tfrac{1}{3} A (\pi W)^2$$

$$M = \sigma^2 + \frac{\pi^2}{3} A^2 W^2;$$

thus from (61b) the percentage minimized mean-square error is bounded by

$$\frac{\Delta_k}{T^2} < \frac{M}{T^2 \beta^2 k} = \frac{1 + \left(\frac{\sigma}{A}\right)^2 \left(\frac{\sqrt{3}}{\pi W}\right)^2}{\frac{1}{3}\pi^2 k}. \tag{69}$$

For a 30 dB signal-to-noise ($A/\sigma$) ratio, and with $W = 3000$ Hz, we see that $\Delta_k/T^2$ is less than 0.01 for $k \geqq 10$. In other words, after 10 symbols have been received, the above synchronization algorithm reduces the mean-square error to less than 1/100 of a symbol interval.

## REFERENCES


1. Bennett, W. R., and Davey, J. R., *Data Transmission,* New York: McGraw-Hill, 1965.
2. Sanger, David K., "Digital Demodulation with Data Subcarrier Tracking," Jet Propulsion Laboratory Technical Report 32–1314, August 1968.
3. Wintz, P. A., and Luecke, E. J., "Performance of Optimum and Supoptimum Synchronizers," IEEE Trans. Commun. Technology, *COM-17* (June 1969), pp. 380–389.
4. Robbins, H., and Monro, S., "A Stochastic Approximation Method," Ann. Math. Stat. *22* (1951), pp. 400–407.
5. Kiefer, J., and Wolfowitz, J., "Stochastic Estimation of the Maximum of a Regression Function," Ann. Math. Stat. *23* (1952), pp. 462–466.
6. Balakrishnan, A. V., ed., *Advances in Communication Systems,* Vol. 2, New York: Academic Press, 1966, pp. 51–106.

7. Saltzberg, B. R., "Timing Recovery for Synchronous Binary Data Transmission," B.S.T.J., *46,* No. 3 (March 1967), pp. 593–622.
8. Lucky, R. W., Salz, J., and Weldon, E. J., *Principles of Data Communication,* New York: McGraw-Hill, 1968.
9. Chang, R. W., "Joint Equalization, Carrier Acquisition, and Timing Recovery for Data Communication," International Conf. Commun., San Francisco, June 1970.
10. Murphy, J. et al., "Ultra Long-Range Telemetry Study," AL-TDR 64–67, Avionics Laboratory, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio, April 1964.
11. Mallory, P., "A Maximum Likelihood Bit Synchronizer," International Telemetering Conf. Proc., *IV* (1968), pp. 1–16.
12. Hirsch, D., private communication.