THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 9, November, 1972
Printed in U.S.A.

On Fast Start-Up Data Communication Systems Using Pseudo-Random Training Sequences

By R. W. CHANG and E. Y. HO

(Manuscript received June 21, 1972)

This paper analyzes the start-up performance of automatic transversal equalizers when maximum-length pseudo-random sequences of short periods are selected as the training signals for fast start-up purposes. Single-sideband Nyquist systems are considered because they represent the limiting case of vestigial-sideband systems with small excess bandwidth. It is shown that the equalizer is capable of fast start-up except in some rare situations which can be avoided by using proper timing, phase, and equalizer initial settings. The results also show that the equalizer tap convergence rate is independent of the phase characteristic of the communication channel and of the choice of the pseudo-random sequences which have the same period.

The equalizer is set up in the training period by minimizing the mean-square error between the equalizer output and the transmitted pseudo-random sequence, which is different from the mean-square error for random data. Surprisingly, we have found that, even for pseudo-random sequences of very short periods, this start-up algorithm results in only a slight degradation in system performance. Accordingly, good system performance can be expected immediately after the system switches from the training mode to the data mode.

I. INTRODUCTION

Pseudo-random sequences have been used in the past as training signals for setting up automatic transversal equalizers during start-up periods.[1,2] For fast start-up, it is desirable to know how the equalizer settling time depends on the choice of the pseudo-random sequence, the channel characteristics, and the initial receiver conditions. These problems are examined in the first part of this paper for single-sideband Nyquist systems. We present basic theories from which the reader can work out numerical examples of special interest. The results are compared with those obtained previously[3] for a different class of training signals (isolated test pulses). An important difference between these two cases is pointed out.

When pseudo-random sequences are used, it is most convenient to adjust the equalizer tap gains to minimize the mean-square error between the equalizer output and the transmitted pseudo-random sequence. It is not immediately clear how closely this simple algorithm optimizes the data set performance for transmission of random data (because an equalizer setting optimum for pseudo-random sequence transmission is not necessarily optimum for random data transmission, particularly when pseudo-random sequences with very short periods are used for fast start-up purpose). This problem is examined in Section IV and the analysis is illustrated by examples.

Section V summarizes the results of this paper. The reader mainly interested in the conclusions and their implications may read Section V next.

II. MATHEMATICAL MODEL AND FUNDAMENTALS

An amplitude modulation data communication system with a conventional tapped delay line transversal equalizer is depicted in Fig. 1. During data transmission, the transmitter transmits the information digits, $\{d_i\}$, sequentially at time instants $t = \cdots, t_1 - T, t_1, t_1 + T, \cdots$. The equalizer output is sampled sequentially at the symbol rate to recover the information digits. Let the ith equalizer output

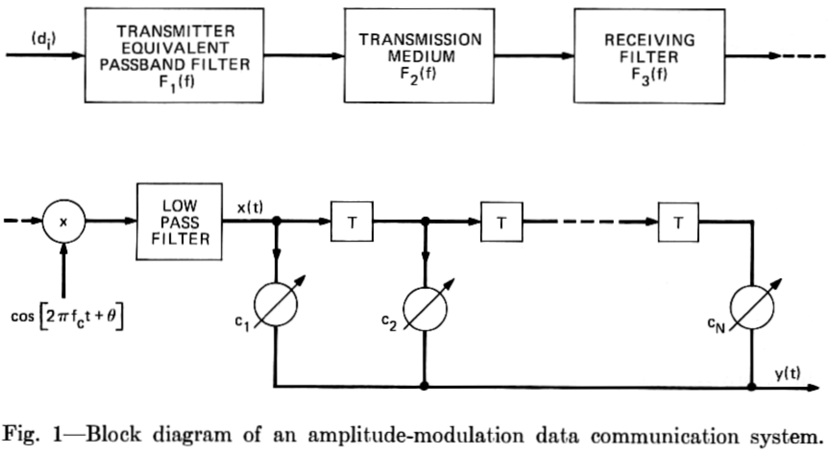

Fig. 1—Block diagram of an amplitude-modulation data communication system.

sample be y_i . We adopt the familiar mean-square error (MSE) criterion and adjust the gain controls of the equalizer to minimize the MSE between y_i and d_i . We assume $\{d_i\}$ is an ergodic process, hence the mean-square error can be written as

$$\begin{aligned} \mathcal{E} &= \lim_{\nu\to\infty} \frac{1}{\nu} \sum_{i=1}^{\nu} (y_i - d_i)^2 \\ &= \langle (y_i - d_i)^2 \rangle \end{aligned} \tag{1}$$

where $\langle x \rangle$ denotes the time average of x.

It can be seen from Fig. 1 that

$$y(t) = \sum_{k=1}^{N} C_k x(t - (k-1)T), \tag{2a}$$

and

$$x(t) = \sum_{i=-\infty}^{\infty} d_i h(t - iT), \tag{2b}$$

where $h(t)$ is the overall system (without equalizer) impulse response. For the sake of simplicity, we shall shift the time origin and use the abbreviations $y_i = y(iT)$, $x_i = x(iT)$, and $h_i = h(iT)$. Thus (1) can be written as

$$\mathcal{E} = \mathbf{C}'\mathbf{A}\mathbf{C} - 2\mathbf{C}'\mathbf{V} + \langle d_i^2 \rangle, \tag{3}$$

where

$$\mathbf{C} = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_N \end{bmatrix}, \tag{4}$$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix}, \tag{5}$$

$$\begin{aligned} a_{ij} &= \langle x_{-i+1} x_{-j+1} \rangle \\ &= \sum_m \sum_n \langle d_m\, d_n \rangle h_{-i+1-m} h_{-j+1-n} \end{aligned}$$

$$i = 1, 2, \cdots, N, \quad j = 1, 2, \cdots, N, \tag{6}$$

$$\mathbf{V} = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_N \end{bmatrix}, \tag{7}$$

and

$$\begin{aligned} V_k &= \langle x_{i-k+1}\, d_i \rangle \\ &= \sum_m \langle d_i\, d_m \rangle h_{i-k+1-m}\,, \qquad k = 1, 2, \cdots, N. \end{aligned} \tag{8}$$

Let $\partial \mathcal{E}/\partial C_i$ be the partial derivative of $\mathcal{E}$ with respect to C_i, $i = 1$ to N, and let $\partial \mathcal{E}/\partial \mathbf{C}$ represent an $N \times 1$ column vector the ith element of which is $\partial \mathcal{E}/\partial C_i$. From (3), we obtain

$$\frac{\partial \mathcal{E}}{\partial \mathbf{C}} = 2\mathbf{AC} - 2\mathbf{V}. \tag{9}$$

The optimum value of $\mathbf{C}$ that minimizes the MSE $\mathcal{E}$ is denoted $\mathbf{C}_{\text{opt}}$. It is clear from (9) that

$$\mathbf{C}_{\text{opt}} = \mathbf{A}^{-1}\mathbf{V}. \tag{10}$$

Thus, the minimum value of $\mathcal{E}$ when $\mathbf{C} = \mathbf{C}_{\text{opt}}$ is

$$\mathcal{E}_{\min} = \langle d_i^2 \rangle - \mathbf{V}'\mathbf{A}^{-1}\mathbf{V}. \tag{11}$$

Let $\mathbf{e}$ denote the difference between $\mathbf{C}$ and $\mathbf{C}_{\text{opt}}$; i.e.,

$$\mathbf{e} = \mathbf{C} - \mathbf{A}^{-1}\mathbf{V}. \tag{12}$$

Now consider the adjustment of the equalizer. As is well known, the equalizer can be adjusted in the training period by transmitting either a succession of isolated test pulses or a special class of pseudo-random sequence (pseudo-noise sequence[3]). The case of sending isolated test pulses during the training period has been considered by Chang.[4] In this paper, we examine the case of sending pseudo-noise sequences.†

In the training period, a binary pseudo-noise sequence is applied to the transmitter input. Since an adjustment is made at the end of each training sequence period and since we wish to make the largest number of adjustments during a fixed training time, we consider the shortest‡ possible pseudo-random sequence period (i.e., the case where

† The best known sequences of this type are the m-sequences (also known as maximal-length linear recurring sequences or maximum-length pseudo-random sequences).

‡ When the period of pseudo-random sequence is shorter than the length of the transversed equalizer, the analysis is difficult because A^{-1} may not exist.

the period of the pseudo-noise sequence is equal to the length of the transversal equalizer).

Let the pseudo-noise sequence be denoted by $\beta_1\beta_2\beta_3 \cdots \beta_N \beta_1\beta_2\beta_3 \cdots \beta_{k'}$, where $\beta_{k'}$ is the last bit. From (2b), the input to the equalizer can be written as

$$x(t) = \sum_{k=1}^{k'} \beta_k h(t - kT). \qquad (2b)$$

Practically speaking, we may assume that $h(t)$ is time limited. Then it can be shown that for t larger than a certain value, say t_0, $x(t)$ will be a periodic function of period NT; i.e.,

$$x(t) = x(t + NT), \quad t_0 < t \quad \text{and} \quad t + NT \leqq k'T. \qquad (13)$$

In the training period, the values of a_{ij}, V_k, $\mathcal{E}$, $\mathbf{C}_{\text{opt}}$, $\mathbf{A}$, and $\mathbf{V}$ are denoted by a^*_{ij}, V^*_k, $\mathcal{E}^*$, $\mathbf{C}^*_{\text{opt}}$, $\mathbf{A}^*$, and $\mathbf{V}^*$, respectively. From (6), we obtain

$$\begin{aligned} a^*_{ij} &= \sum_m \sum_n \langle \beta_m \beta_n \rangle h_{-i+1-m} h_{-j+1-n} \\ &= \sum_k \sum_m h_{-i+1-m} h_{-j+1-m+kN} \\ &\quad - \frac{1}{N} \sum_k \sum_{l=1}^{N-1} \sum_m h_{-i+1-m} h_{-j+1-m+kN+l} \quad \text{all} \quad i, j. \end{aligned} \qquad (6)$$

From (8)

$$\begin{aligned} V^*_k &= \sum_m \langle \beta_m \beta_i \rangle h_{i-k+1-m} \\ &= \sum_i h_{-k+1+iN} - \frac{1}{N} \sum_i \sum_{l=1}^{N-1} h_{-k+1+jN+l}, \qquad k = 1, 2, \cdots, N \\ &\qquad\qquad j = \text{integers}. \end{aligned} \qquad (8)$$

The partial derivative $\partial\mathcal{E}^*/\partial C_i$, $i = 1$ to N, can be computed from each block of N samples of $x(t)$, and the gain control C_i is adjusted by an amount proportional to $\partial\mathcal{E}^*/\partial C_i$ at the end of each block. For example, $\partial\mathcal{E}^*/\partial C_i$ is computed from x_{i+1} to x_{i+N} and C_i is adjusted after x_{i+N}. Then $\partial\mathcal{E}^*/\partial C_i$ is computed from x_{i+N+1} to x_{i+2N} and C_i is adjusted after x_{i+2N}. The optimum tap setting that minimizes the MSE, $\mathcal{E}^*$, is

$$\mathbf{C}^*_{\text{opt}} = (\mathbf{A}^*)^{-1}\mathbf{V}^*.$$

We now proceed to examine the convergence of $\mathcal{E}^*$.

III. SINGLE-SIDEBAND NYQUIST SYSTEMS

In this section, we consider single-sideband data communication systems which transmit at the Nyquist rate with $\sin x/x$ pulses (hereafter referred to as single-sideband Nyquist system). Such systems are considered because they represent the limiting case of sharp rolloff vestigial-sideband systems.

The transfer functions of the transmitting filter, transmission medium, and receiving filter (see Fig. 1) are $F_1(f)$, $F_2(f)$ and $F_3(f)$, respectively. The $F_i(f)$ are of the following form

$$F_i(f) = | F_i(f) | e^{J\beta_i(f)}, \quad i = 1, 2, 3 \tag{14}$$

where J is used to denote the imaginary number $\sqrt{-1}$.

In a single-sideband Nyquist system, $| F_1(f) |$ and $| F_3(f) |$ are specified by

$$\begin{aligned} | F_1(f) | &= 1 \qquad f_1 \leqq | f | \leqq f_2 \\ &= 0 \qquad \text{otherwise,} \end{aligned} \tag{15}$$

and

$$\begin{aligned} | F_3(f) | &= 1 \qquad f_1 \leqq | f | \leqq f_2 \\ &= 0 \qquad \text{otherwise.} \end{aligned} \tag{16}$$

In general, with lower single-sideband transmission, the carrier frequency, f_c, is set equal to f_2. Let $H(f)$ denote the Fourier transform of $h(t)$, which is the overall system impulse response at the equalizer input. It can be shown that

$$\begin{aligned} H(f) &= \tfrac{1}{2} | F_2(f - f_c) | \\ &\quad \cdot e^{J[\beta_1(f-f_c)+\beta_2(f-f_c)+\beta_3(f-f_c)-2\pi(f-f_c)t_1+\theta]}, \qquad 0 \leqq f \leqq f_2 - f_1 \\ &= \tfrac{1}{2} | F_2(f + f_c) | \\ &\quad \cdot e^{J[\beta_1(f+f_c)+\beta_2(f+f_c)+\beta_3(f+f_c)-2\pi(f+f_c)t_1-\theta]}, \qquad -(f_2 - f_1) \leqq f \leqq 0 \\ &= 0 \qquad \text{otherwise,} \end{aligned} \tag{17}$$

where θ represents demodulating carrier phase. The signaling interval is

$$T = \frac{1}{2(f_2 - f_1)}. \tag{18}$$

Since the time samples $h(iT)$, $i = \cdots 0, 1, 2, \cdots$ are taken at the Nyquist rate, we obtain from the sampling theorem, Parseval's theorem, and (17)

$$g(i-j) \triangleq \sum_{m=-\infty}^{\infty} h_{-i+1-m} h_{-j+1-m}$$

$$= 2(f_2 - f_1) \int_{-\infty}^{\infty} h(t - iT + T) h(t - jT + T)\, dt$$

$$= (f_2 - f_1) \int_0^{f_2 - f_1} [\cos 2\pi f(i-j)T][|\, F_2(f - f_c)\,|]^2\, df. \quad (19)$$

Substituting (19) into (6), we obtain

$$a_{ij}^* = \sum_k g(i - j + kN) - \frac{1}{N} \sum_k \sum_{l=1}^{N-1} g(i - j + kN + l) \quad \text{all } i, j. \quad (20)$$

It is clear from (19) and (20) that a_{ij}^* is independent of the demodulating carrier phase θ, the system timing t_1, and the phase characteristics $\beta_i(f)$ of the system. We also note that for a fixed N, a_{ij}^* is independent of the choice of the pseudo-noise sequence. [The pseudo-noise sequence $\beta_1, \beta_2, \cdots$ does not appear in (19) or (20).] Using the method in Ref. 4, it is concluded that the equalizer tap convergence rate is independent of the demodulating carrier phase θ, the system timing t_1, the phase characteristics of the system, and the choice of the pseudo-noise sequence for fixed N.

Note from (19) and (20) that a_{ij}^* depends on the amplitude characteristic $|\, F_2(f)\,|$ of the transmission medium. Since amplitude distortion is not severe in private line systems, in the following discussion we assume that

$$|\, F_2(f)\,| = 1, \quad f_1 \leqq f \leqq f_2\,. \quad (21)$$

Substituting (21) into (19), and neglecting a normalizing constant $(f_2 - f_1)^2$, we obtain

$$\begin{aligned} g(i-j) &= 1, \qquad i = j \\ &= 0, \qquad i \neq j. \end{aligned} \quad (22)$$

Substituting (22) into (20) gives

$$\begin{aligned} a_{ij}^* &= 1, \qquad i = j \\ &= -\frac{1}{N} \qquad i \neq j. \end{aligned} \quad (23)$$

Therefore, the eigenvalues of $\mathbf{A}^*$ are

$$\lambda_k = 1 + \frac{1}{N}, \qquad k = 1, 2, \cdots, N-1$$

$$\lambda_N = \frac{1}{N}. \tag{24}$$

The eigenvector $\mathbf{u}_N$ corresponding to λ_N is an $N \times 1$ vector whose elements are all unity. Since all but the last eigenvalue are equal, the equalizer can settle rapidly except in the case where the initial $\mathcal{E}^*$ contains a large component $(\mathbf{e}_0'\mathbf{u}_N)^2\lambda_N$, where $\mathbf{e}_0$ is the initial tap setting error vector[3]. Since λ_N is small and since $(\mathbf{e}_0'\mathbf{u}_N)^2$ cannot be exceptionally large with proper timing, phase, and initial equalizer settings†, it is very unlikely that $\mathcal{E}^*$ would contain a large component $(\mathbf{e}_0'\mathbf{u}_N)^2\lambda_N$. Therefore, the equalizer can settle rapidly in the training period.

IV. FURTHER ANALYSIS OF SYSTEM PERFORMANCE

At the end of the training period, the equalizer taps are set very nearly to $\mathbf{C}^*_{\text{opt}}$. The data set is then switched to the data transmission mode. Since statistics of the true data differ from those of the training pseudo-random sequence, the optimum tap settings, $\mathbf{C}^*_{\text{opt}}$, obtained for a training sequence cannot also be the optimum one for the true data. Thus, system performance degradation during the early stage of data transmission is expected, even if the data set is equipped with an adaptive equalizer. We now proceed to determine this degradation.

We assume zero-mean independent information digits (binary or multilevel). The signal level is normalized such that

$$\langle d_i^2 \rangle = 1. \tag{25}$$

The mean-square error, $\mathcal{E}$, can be obtained from (3),

$$\mathcal{E} = (\mathbf{C}^*_{\text{opt}})'\mathbf{A}(\mathbf{C}^*_{\text{opt}}) - 2(\mathbf{C}^*_{\text{opt}})'\mathbf{V} + 1, \tag{26}$$

where $\mathbf{A}$ and $\mathbf{V}$ are given by (5) and (7), respectively. From (6), (8), (19), and (25), we obtain

$$a_{ij} = g(i-j), \tag{27}$$

and

$$V_k = h_{-k+1}. \tag{28}$$

The a_{ij}^* and V_k^* can be rewritten as

$$a_{ij}^* = a_{ij} + \Delta_{ij}(N), \tag{29}$$

† For example, one may use the method described in Ref. 1.

and

$$V_k^* = V_k + \gamma_k(N), \tag{30}$$

where

$$\Delta_{ij}(N) = \sum_{k \neq 0} g(i - j + kN) - \frac{1}{N} \sum_k \sum_{l=1}^{N-1} g(i - j + kN + l), \tag{31}$$

and

$$\gamma_k(N) = \sum_{j \neq 0} h_{-k+1+jN} - \frac{1}{N} \sum_j \sum_{l=1}^{N-1} h_{-k+1+jN+l} \,. \tag{32}$$

From (26), we have

$$\begin{aligned} \mathcal{E} &= \mathbf{C}'_{\text{opt}}\mathbf{A}\mathbf{C}_{\text{opt}} - 2\mathbf{C}_{\text{opt}}\mathbf{V} + 1 + (\boldsymbol{\delta}\mathbf{C})'\mathbf{A}(\boldsymbol{\delta}\mathbf{C}) \\ &= \mathcal{E}_{\min} + (\boldsymbol{\delta}\mathbf{C})'\mathbf{A}(\boldsymbol{\delta}\mathbf{C}), \end{aligned} \tag{33}$$

where $\mathbf{C}_{\text{opt}}$ is the optimum tap setting for the true data and $\boldsymbol{\delta}\mathbf{C}$ is the difference between $\mathbf{C}^*_{\text{opt}}$ and $\mathbf{C}_{\text{opt}}$,

$$\boldsymbol{\delta}\mathbf{C} = \mathbf{C}^*_{\text{opt}} - \mathbf{C}_{\text{opt}} \,. \tag{34}$$

The last term in (33) represents the system performance degradation and is non-negative.

The mean-square error during the early stage of data transmission can now be determined from (33) and (34). As the period of the training PN sequence approaches infinity, $\lim_{N\to\infty} \Delta_{ij}(N) \to 0$, $\lim_{N\to\infty} \gamma_k(N) \to 0$, and $\lim_{N\to\infty} \boldsymbol{\delta}\mathbf{c} \to \mathbf{0}$. Hence, $\mathbf{C}^*_{\text{opt}}$ approaches $\mathbf{C}_{\text{opt}}$ asymptotically as N increases.

We now assume some specific channel characteristics and apply these formulas to determine the initial performance degradation. *Example 1:* A baseband channel with a flat amplitude characteristic and a typical quadratic delay characteristic[5] is assumed. The delay at the Nyquist frequency is taken to be $\beta_m T$ seconds. The phase characteristic is of the form

$$\beta_2(f) = 8\pi\beta_m T^3 f^3/3. \tag{35}$$

The system impulse response, $h(t)$, can be calculated from

$$h(t) = 2 \int_0^{1/2T} \cos\,(2\pi f t + \beta(f))\, df. \tag{36}$$

In this example, we consider a typical value $\beta_m = 2$. One hundred one samples of $h(t)$ (from $t_0 - 50T$ to $t_0 + 50T$) are taken with $T = 1$ and

$t_0 = 0.6T$. For a 7-tap equalizer, the minimum mean-square error attainable is 0.03347. The mean-square error obtained from (33) is 0.0393. The results for a 15-tap equalizer are 0.01484 and 0.01785, respectively. It is clear from these numbers that the performance degradation caused by using a PN sequence in the training period is negligible. This can be further illustrated by sketching the vector $\boldsymbol{\delta}\mathbf{C}$ in (34). Since the amplitude characteristic is constant, we have

$$\mathbf{A} = \mathbf{I} \tag{37a}$$

and

$$\mathbf{A}^* = \mathbf{I} - \boldsymbol{\Delta}, \tag{37b}$$

where

$$\boldsymbol{\Delta} = \begin{bmatrix} 0 & \frac{1}{N} & \frac{1}{N} & \cdots & \frac{1}{N} \\ \frac{1}{N} & 0 & \frac{1}{N} & \cdots & \frac{1}{N} \\ \frac{1}{N} & \frac{1}{N} & 0 & \cdots & \frac{1}{N} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{N} & \frac{1}{N} & \frac{1}{N} & \cdots & 0 \end{bmatrix} \tag{37c}$$

The inverse of $\mathbf{A}^*$ is found to be

$$(\mathbf{A}^*)^{-1} = \mathbf{I} + \frac{N-1}{N+1}\mathbf{I} + \frac{N^2}{N+1}\boldsymbol{\Delta}. \tag{38}$$

From (10), (34), (37a), and (38), we have

$$\boldsymbol{\delta}\mathbf{C} = \frac{N-1}{N+1}\mathbf{V} + \frac{2N}{N+1}\boldsymbol{\gamma} + \frac{N^2}{N+1}\{\boldsymbol{\Delta}\mathbf{V} + \boldsymbol{\Delta}\boldsymbol{\gamma}\}. \tag{39}$$

Substituting (28), (32), and (37c) into (39), we obtain

$$\boldsymbol{\delta}\mathbf{C} = \begin{bmatrix} \sum_{j\neq 0} h_{[(N-1)/2]+t_0+jN} \\ \vdots \\ \sum_{j\neq 0} h_{-[(N-1)/2]+t_0+jN} \end{bmatrix}. \tag{40}$$

In Fig. 2, the time samples h_k, $k = -\infty$ to ∞, of $h(t)$ are sketched.

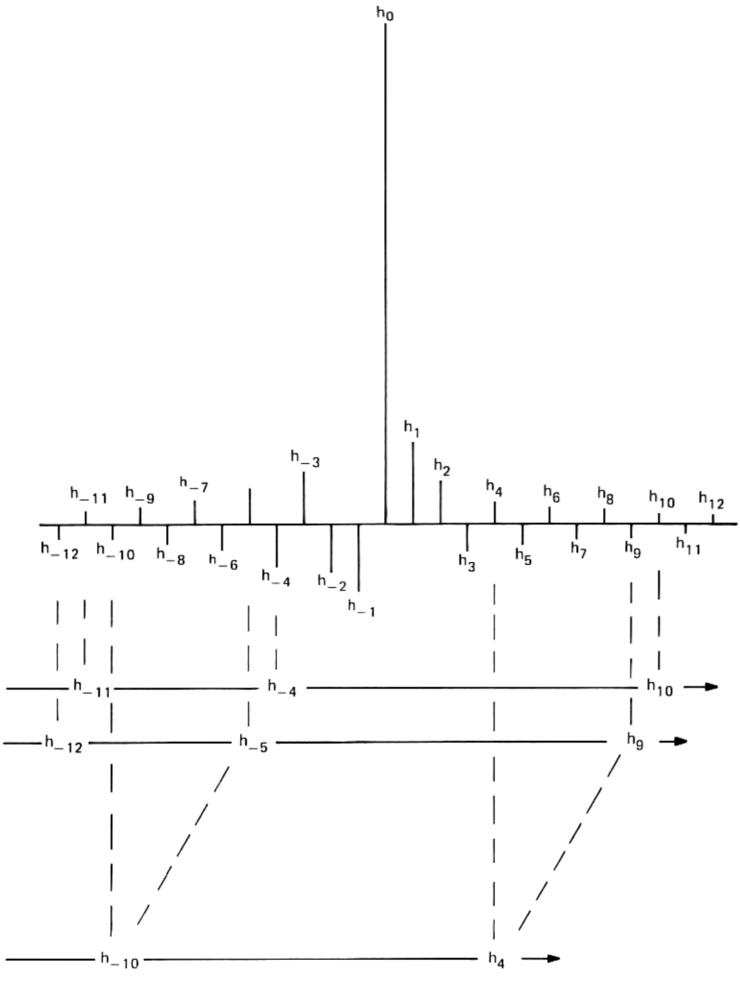

Fig. 2—Illustration of the elements of $\boldsymbol{\delta}\mathbf{C}$

For $N = 7$ (7-tap equalizer), the first element of $\boldsymbol{\delta}\mathbf{C}$ is the sum of the infinite sequence $\cdots, h_{-11}, h_{-4}, h_{10}, \cdots$; the second element of $\boldsymbol{\delta}\mathbf{C}$ is the sum of the infinite sequence $\cdots, h_{-12}, h_{-5}, h_9, \cdots$; etc. It can be seen that the large time samples h_{-3} to h_3 are not included in these sums. This explains why $\boldsymbol{\delta}\mathbf{C}$ should have small elements. By repeating the above for $N = 15$, one can easily see that the performance degradation due to the use of PN training sequence is small and that this degradation approaches zero as N increases.

Example 2: We continue example 1 but change β_m to its minimum value zero. The infinite sum $\sum_{i \neq 0} h_{-k+t_0+iN}$ can now be evaluated in closed form. By using the sampling theorem or the formulas for Psi (Digamma) functions,[6] it is obtained after some manipulations

$$\sum_{j \neq 0} h_{-k+t_0+jN} = (-1)^{-k+N} \frac{\sin \pi t_0}{\pi N} \cdot \left\{ \frac{N}{t_0 - k} - \frac{\pi}{2} \left(Ctg \frac{\pi(t_0 - k)}{2N} + tg \frac{\pi(t_0 - k)}{2N} \right) \right\} \quad (41)$$

where $0 \leqq t_0 \leqq 0.5T$ is assumed. The performance degradation can now be determined in closed form

$$(\delta \mathbf{C})'\mathbf{A}(\delta \mathbf{C}) = \sum_{k=-[N-1]/2}^{[N-1]/2} \frac{\sin^2 \pi t_0}{\pi^2 N^2} \cdot \left\{ \frac{N}{t_0 - k} - \frac{\pi}{2} \left(Ctg \frac{\pi(t_0 - k)}{2N} + tg \frac{\pi(t_0 - k)}{2N} \right) \right\}^2 \quad (42)$$

$\mathcal{E}_{\min}$ and $(\delta \mathbf{C})'\mathbf{A}(\delta \mathbf{C})$ are plotted in Fig. 3 for $N = 7$, 15, and 31 and $t_0 = 0.05$, 0.1, 0.15, 0.2 and 0.25. It can be seen that the value of $(\delta \mathbf{C})'\mathbf{A}(\delta \mathbf{C})$ is approximately an order of magnitude less than that of $\mathcal{E}_{\min}$. Also note that $(\delta \mathbf{C})'\mathbf{A}(\delta \mathbf{C})$ reduces almost by half when N is doubled. These results again show that the performance degradation caused by using PN training sequence is negligible.

V. CONCLUSIONS AND DISCUSSIONS

We have analyzed the start-up performance of a transversal equalizer for the case where a maximum-length pseudo-random sequence is used as the training signal to adjust the equalizer in the training period. The equalizer taps are adjusted by the gradient method to minimize the mean-square error, $\mathcal{E}^*$, between the equalizer output and the transmitted pseudo-random sequence. The pseudo-random sequence has been denoted $\beta_1, \beta_2, \cdots, \beta_N, \beta_1, \beta_2, \cdots$, where N is the period of the sequence. We have considered the case where N is equal to the number of taps of the equalizer. The following results are obtained:

(*i*) For a fixed N, the initial value of the mean-square error $\mathcal{E}^*$, the convergence rate of $\mathcal{E}^*$, and the minimum value of $\mathcal{E}^*$ are all independent of the specific values of the β_k's. Therefore, the same performance is obtained with any of the many pseudo-random sequences available. For example, a maximum-length pseudo-random sequence can be cyclic shifted to produce N

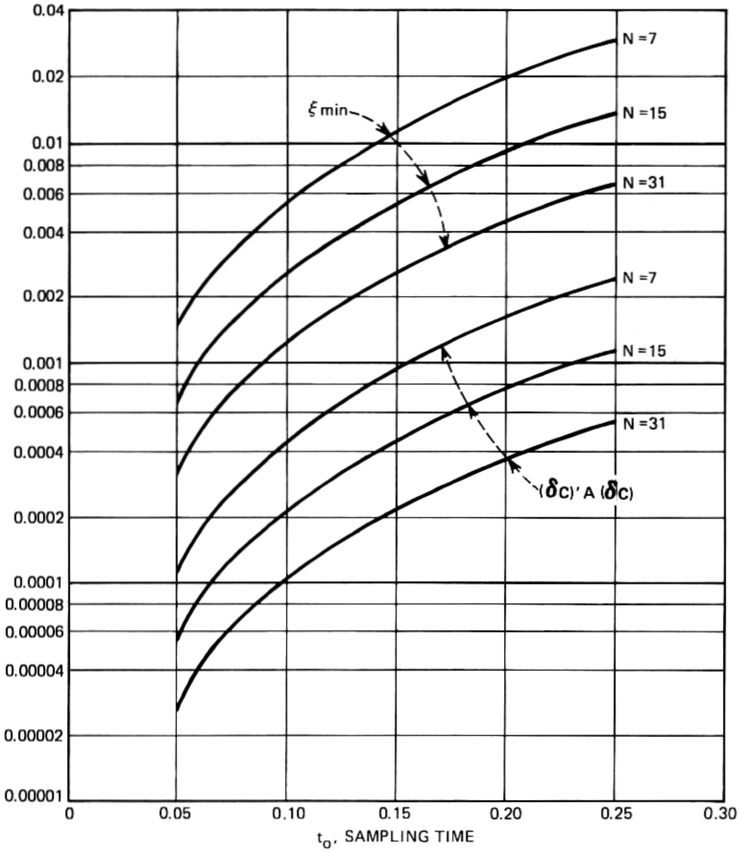

Fig. 3—Computational results of Example 2.

pseudo-random sequences. Any one of these N sequences can be used in the training period with the same result.

(*ii*) The initial value of $\mathcal{E}^*$ depends on the phase characteristic of the communication channel, and the timing and phase settings at the receiver. However, the convergence rate of $\mathcal{E}^*$ is independent of all these parameters. This result is similar to the one obtained previously[4] for the case where isolated test pulses are used as the training signal.

(*iii*) Unlike the isolated test pulse case, the eigenvalues of the correlation matrix here are not close together. For example, when the channel has a flat amplitude characteristic, the first $N - 1$ eigenvalues λ_1 to λ_{N-1} are equal to $1 + 1/N$, while the Nth eigenvalue λ_N is equal to $1/N$. The mean-square error can be decomposed into N components, each associated with one of the N eigenvalues. The tap gain adjustment reduces rapidly the $N - 1$ components associated with λ_1 to λ_{N-1}, but the component associated with λ_N decreases very slowly. Therefore, as discussed at the end of Section III, care must be exercised in setting the timing, phase, and equalizer taps at the beginning of the training period so that the component associated with λ_N has a small initial value. Note that this precaution is not required when isolated test pulses are used, because in that case the eigenvalues are all close together and the components of the mean-square error all decrease rapidly.[4]

(*iv*) The analysis shows that the tap settings obtained with maximum-length pseudo-random sequences with very short periods are nearly optimum for random data transmission. More specifically, the equalizer taps are adjusted in the training period to minimize the mean-square error $\mathcal{E}^*$ between the equalizer output and the transmitted pseudo-random sequence. When such tap settings are used for actual data transmission, the mean-square error between the equalizer output and the transmitted random data can be written as

$$\mathcal{E} = \mathcal{E}_{\min} + \epsilon$$

where $\mathcal{E}_{\min}$ is the minimum attainable value of $\mathcal{E}$, and ϵ is nonnegative, because tap settings obtained with pseudo-random sequence do not necessarily minimize $\mathcal{E}$. Formulas for computing ϵ were developed in Section IV and illustrated by numerical examples. It can be seen from these formulas and Figs. 2 and 3 that ϵ decreases rapidly as N increases (N is the period of the pseudo-random sequence and also the number of equalizer taps). The computations show that ϵ decreases approximately by the factor $1/N$ (for example, ϵ reduces approximately by half when N is doubled). The computations also show that ϵ is about an order of magnitude less than $\mathcal{E}_{\min}$. (This is so for N as small as seven.) Therefore, tap settings obtained with pseudo-random sequences are nearly optimum for actual data transmission.

(*v*) As can be seen from the computations in Section IV, $\mathcal{E}_{\min}$ is rather large when N is small. For example, for a system with typical channel delay distortion (see example 1) and S/N = 30 dB, $\mathcal{E}_{\min}$ can be 15 dB above the thermal noise level when

$N = 7$, and 12 dB above the thermal noise level when $N = 15$. These large mean-square errors are due to the fact that for single-sideband Nyquist systems the overall system impulse response decays very slowly with time. Thus, for very sharp rolloff VSB systems (such as 4-percent rolloff), it is necessary to use a large number of equalizer taps (such as 31 or more).

REFERENCES

1. Lucky, R. W., and Rudin, H. R., "An Automatic Equalizer for General-Purpose Communication Channels," B.S.T.J., *46*, No. 9 (November 1967), pp. 2179–2208.
2. Mueller, K. H., Private Communication.
3. Golomb, S. W., *Digital Communications with Space Applications*, Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1964.
4. Chang, R. W., "A New Equalizer Structure for Fast Start-Up Digital Communication," B.S.T.J., *50*, No. 6 (July–August 1971), pp. 1969–2014.
5. Lucky, R. W., Salz, J., and Weldon, E. J., Jr., *Principles of Data Communication*, New York: McGraw-Hill, 1969, pp. 68–69.
6. Gradshteyn, I. S., and Ryzhik, I. M., *Table of Integrals Series and Products*, New York and London: Academic Press, 1965.