Joint Optimization of Automatic Equalization and Carrier Acquisition for Digital Communication

By ROBERT W. CHANG

(Manuscript received January 22, 1970)

In this paper, we analyze single-sideband amplitude-modulation digital communication systems to develop a method for jointly and optimally setting the carrier phase and the automatic transversal equalizer of such systems. Mean-square equalization error is used as the performance criterion. We develop a simple receiver structure and study the convergence of the method. Exact locations of the stationary points in the parameter space are determined and the classifications of the stationary points are obtained. We show that the mean-square equalization error has only global minima and saddlepoints, but not local minima and maxima. Thus, the mean-square equalization error will converge to the absolute minimum by the proposed method, regardless of the initial settings of the parameters. A simple condition on the step sizes of the adjustments is also obtained which ensures the convergence of the process. Explicit formulas of the joint optimum parameter settings and of the corresponding minimum mean-square error are obtained. For illustration purposes, a single-sideband digital communication system using a five- or nine-tap transversal equalizer is simulated on a computer. Both theory and simulation show that the equalization error depends critically on the carrier phase when the number of equalizer taps is not large, and that the minimum equalization error can be obtained by using the proposed method.

I. INTRODUCTION

In single-sideband amplitude-modulation digital communication systems with transversal filter equalization,[1,2] the adjustment of the carrier phase is critical to the system's performance when the number of equalizer taps is not large. In this paper, a method is proposed for setting the carrier phase jointly with the automatic equalizer to minimize the mean-square equalization error.

We formulate a mathematical model of this study in Section II; in Sections III and IV, we analyze the system and develop a receiver structure. The problem of convergence is studied in Sections V and VI to determine if the equalization error will converge to the absolute minimum by the proposed method and whether such convergence depends on the initial settings of the parameters. We also consider step sizes of the adjustments. In Section VII, we derive explicit formulas for evaluating the system's performance. A voiceband data communication system is simulated on a computer to test the proposed method and the results are described in Section VIII.

II. MATHEMATICAL MODEL

As shown in Fig. 1, a single-sideband amplitude-modulation system is considered. When an impulse $\delta(t)$ is applied to the transmitter input, a signal $a(t)$ is received at the receiving filter output. The Fourier transform of $a(t)$ is denoted by $A(f)$. (The Fourier transform of a function will be consistently denoted by the appropriate capital letter.) It is assumed that $A(f)$ is band-limited between f_1 and f_2, that is,

$$A(f) \neq 0, \quad \text{only for} \quad f_1 < |f| < f_2 . \tag{1}$$

The signal $a(t)$ is demodulated as shown in Fig. 1, where the demodulating carrier frequency is f_c. In single-sideband systems

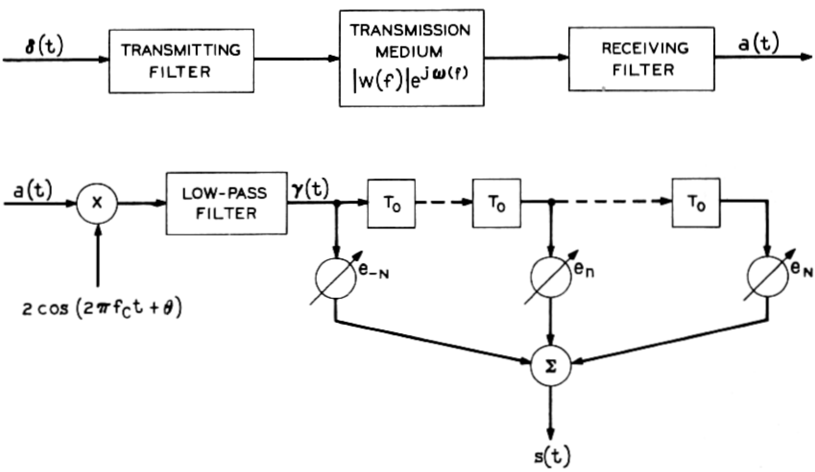

Fig. 1—An amplitude modulation system with coherent detection and transversal filter equalization.

$$f_c \leqq f_1 \tag{2}$$

or

$$f_c \geqq f_2 . \tag{3}$$

The demodulating carrier phase is denoted by θ. The demodulator output is a signal $\gamma(t)$ with Fourier transform $\Gamma(f)$. Since $A(f)$ is band-limited, $\Gamma(f)$ is also band-limited, that is,

$$\Gamma(f) = 0, \qquad |f| > f_0 \tag{4}$$

where f_0 is $f_2 - f_1$ when f_c is f_1 or f_2 (f_0 is larger than $f_2 - f_1$ if f_c is less than f_1 or larger than f_2).

As shown in Fig. 1, the channel is equalized by a conventional transversal equalizer consisting of $2N + 1$ taps with gains e_n, $n = -N$ to N, spaced at T_0-second intervals, where

$$T_0 = \frac{1}{2f_0} \quad \text{seconds.} \tag{5}$$

When an impulse $\delta(t)$ is applied at the transmitter input, the transversal equalizer output is $s(t)$. Clearly $s(t)$ is the overall impulse response of the system. The receiver parameters (that is, the demodulating carrier phase θ and the equalizer tap gains e_n, $n = -N$ to N) will be set to minimize the difference between the overall impulse response $s(t)$ and a desired impulse response $q(t)$. The familiar mean-square error criterion is used. That is, θ and e_n, $n = -N$ to N, will be jointly set to minimize the mean-square error

$$\mathcal{E}_0 = \int_{-\infty}^{\infty} [s(t) - q(t)]^2 \, dt. \tag{6}$$

For brevity, the tap gains e_n, $n = -N$ to N, will be abbreviated $\{e_n\}$ in the sequel.

III. ANALYSIS

In this section, we analyze the system to develop a receiver structure for jointly setting θ and $\{e_n\}$ by the method of steepest descent.

In analyzing carrier signals, it is most convenient to use Hilbert transform techniques. As is well known,[1] the demodulator output $\gamma(t)$ is related to the input $a(t)$ by

$$\gamma(t) = \cos(2\pi f_c t + \theta)a(t) + \sin(2\pi f_c t + \theta)\hat{a}(t) \tag{7}$$

where $\hat{a}(t)$ is the Hilbert transform of $a(t)$. {When dealing with lengthy time functions, we shall sometimes use the sign $\mathcal{H}$, that is, $\mathcal{H}[a(t)] = \hat{a}(t)$.}

It is seen from Fig. 1 that

$$s(t) = \sum_{n=-N}^{N} e_n \gamma(t - nT_0 - NT_0). \tag{8}$$

We shall use the partial derivatives $\partial \mathcal{E}_0/\partial\theta$ and $\partial \mathcal{E}_0/\partial e_n$ in the method of steepest descent. Since $q(t)$ is independent of θ, we obtain from equation (6)

$$\frac{\partial \mathcal{E}_0}{\partial \theta} = \int_{-\infty}^{\infty} 2[s(t) - q(t)] \frac{\partial s(t)}{\partial \theta} \, dt. \tag{9}$$

In writing equation (9), the order of differentiation and integration has been exchanged [the underlying conditions for making such an exchange are easily satisfied by $s(t)$ and $q(t)$ encountered in communication systems]. Substituting equation (7) into equation (8) and taking the partial derivative give

$$\frac{\partial s(t)}{\partial \theta} = \sum_{n=-N}^{N} e_n \{ -\sin [2\pi f_c(t - nT_0 - NT_0) + \theta] a(t - nT_0 - NT_0) + \cos [2\pi f_c(t - nT_0 - NT_0) + \theta] \hat{a}(t - nT_0 - NT_0) \}. \tag{10}$$

From equations (8) and (7),

$$\hat{s}(t) = \sum_{n=-N}^{N} e_n [\mathcal{H}\{ \cos [2\pi f_c(t - nT_0 - NT_0) + \theta] a(t - nT_0 - NT_0) \} + \mathcal{H}\{ \sin [2\pi f_c(t - nT_0 - NT_0) + \theta] \hat{a}(t - nT_0 - NT_0) \}]. \tag{11}$$

In single-sideband modulation, we have either inequality (2) or inequality (3). Let us consider inequality (2) first. When inequality (2) holds, the frequency spectrum of $a(t)$, $A(f)$, does not overlap the spectra of $\cos 2\pi f_c t$ and $\sin 2\pi f_c t$. Furthermore, $A(f)$ occupies a higher frequency band; therefore, equation (11) becomes

$$\mathcal{H}[\cos (2\pi f_c t + \theta) a(t)] = \cos (2\pi f_c t + \theta) \hat{a}(t),$$

and

$$\mathcal{H}[\sin (2\pi f_c t + \theta) \hat{a}(t)] = -\sin (2\pi f_c t + \theta) a(t), \qquad f_c \leqq f_1 .$$

Substituting the above into equation (11) gives

$$\hat{s}(t) = \sum_{n=-N}^{N} e_n\{\cos[2\pi f_c(t - nT_0 - NT_0) + \theta]\hat{a}(t - nT_0 - NT_0)$$
$$- \sin[2\pi f_c(t - nT_0 - NT_0) + \theta]a(t - nT_0 - NT_0)\},$$
$$f_c \leqq f_1 . \qquad (12)$$

Comparing equation (12) with equation (10) shows that

$$\hat{s}(t) = \frac{\partial s(t)}{\partial \theta}, \qquad f_c \leqq f_1 . \qquad (13)$$

Substituting equation (13) into equation (9) gives

$$\frac{\partial \mathcal{E}_0}{\partial \theta} = \int_{-\infty}^{\infty} 2[s(t) - q(t)]\hat{s}(t)\,dt, \qquad f_c \leqq f_1 . \qquad (14)$$

Since a function and its Hilbert transform are orthogonal, equation (14) reduces to

$$\frac{\partial \mathcal{E}_0}{\partial \theta} = -2\int_{-\infty}^{\infty} q(t)\hat{s}(t)\,dt, \qquad f_c \leqq f_1 . \qquad (15)$$

Thus, $\partial\mathcal{E}_0/\partial\theta$ can be generated by correlating $q(t)$ with $\hat{s}(t)$. Note that the transversal equalizer output will be $\hat{s}(t)$ instead of $s(t)$ if the demodulating carrier $\cos(2\pi f_c t + \theta)$ is replaced by $\cos[2\pi f_c t + \theta + \pi/2]$. However, even though $\hat{s}(t)$ can be generated in this simple fashion, we prefer not to generate $\hat{s}(t)$ because the system must be used instead to generate $s(t)$ to compute the other partial derivatives $\partial\mathcal{E}_0/\partial e_n$. Therefore, we convert equation (15) into the form

$$\frac{\partial \mathcal{E}_0}{\partial \theta} = 2\int_{-\infty}^{\infty} \hat{q}(t)s(t)\,dt, \qquad f_c \leqq f_1 . \qquad (16)$$

This step can be verified by Parseval's theorem. Now we need only to correlate $s(t)$ with an easily generated $\hat{q}(t)$ to obtain $\partial\mathcal{E}_0/\partial\theta$.

The above is for the case $f_c \leqq f_1$. In the other case, $f_c \geqq f_2$; the frequency spectrum of $a(t)$ occupies a frequency band lower than f_c; therefore, the two equations above equation (12) should be rewritten as

$$\mathcal{H}[\cos(2\pi f_c t + \theta)a(t)] = \sin(2\pi f_c t + \theta)a(t),$$
$$\mathcal{H}[\sin(2\pi f_c t + \theta)\hat{a}(t)] = -\cos(2\pi f_c t + \theta)\hat{a}(t), \qquad f_c \geqq f_2 .$$

Repeating the steps from equation (12) to equation (16), we get

$$\frac{\partial \mathcal{E}_0}{\partial \theta} = -2\int_{-\infty}^{\infty} s(t)\hat{q}(t)\,dt, \qquad f_c \geqq f_2 . \qquad (17)$$

Note from equations (17) and (16) that the sign of the correlator output must be reversed when one shifts the carrier frequency from one side of $A(f)$ to the other side.

About the equalizer tap gains, it is seen from equations (6) and (8) that

$$\frac{\partial \mathcal{E}_0}{\partial e_n} = \int_{-\infty}^{\infty} 2[s(t) - q(t)] \frac{\partial s(t)}{\partial e_n} \, dt,$$

$$= \int_{-\infty}^{\infty} 2[s(t) - q(t)]\gamma(t - nT_0 - NT_0) \, dt, \qquad n = -N \text{ to } N. \tag{18}$$

Thus, $\partial \mathcal{E}_0 / \partial e_n$ can be generated by correlating the error signal $[s(t) - q(t)]$ with the output $\gamma(t - nT_0 - NT_0)$ of the nth tap. This is the concept introduced in Ref. 3 where the problem of setting the tap gains $\{e_n\}$ was considered.

IV. RECEIVER STRUCTURE

It has been shown in the previous section that $\partial \mathcal{E}_0 / \partial \theta$ can be easily generated simultaneously with $\partial \mathcal{E}_0 / \partial e_n$, $n = -N$ to N. Therefore, the method of steepest descent can be used to adjust simultaneously θ and $\{e_n\}$. In the training period prior to data transmission, isolated test pulses are transmitted. For instance, $\delta(t)$ in Fig. 1 may be one of the test pulses. The transmission of $\delta(t)$ generates a signal $s(t)$ at the equalizer output. From $s(t)$ the partial derivatives $\partial \mathcal{E}_0 / \partial \theta$ and $\partial \mathcal{E}_0 / \partial e_n$, $n = -N$ to N, are computed. The parameters θ and $\{e_n\}$ are then changed by amounts proportional to the partial derivatives, that is,

$$\theta_{\text{new}} = \theta_{\text{old}} - \alpha \frac{\partial \mathcal{E}_0}{\partial \theta},$$

$$(e_n)_{\text{new}} = (e_n)_{\text{old}} - \beta \frac{\partial \mathcal{E}_0}{\partial e_n}, \qquad n = -N \text{ to } N,$$

where α and β are positive proportional constants which may vary from one adjustment to another (the choice of their values will be considered in Section VI). After the changes are all made, another test pulse is transmitted and the process is repeated. The process is terminated after a prefixed number of test pulses.

The receiver structure is shown in Fig. 2. The partial derivatives $\partial \mathcal{E}_0 / \partial \theta$ and $\partial \mathcal{E}_0 / \partial e_n$ are computed according to equations (17) and (18), respectively ($f_c \geqq f_2$ is assumed). If $f_c \leqq f_1$, the correlator output in Fig. 2 will be $-\partial \mathcal{E}_0 / \partial \theta$.

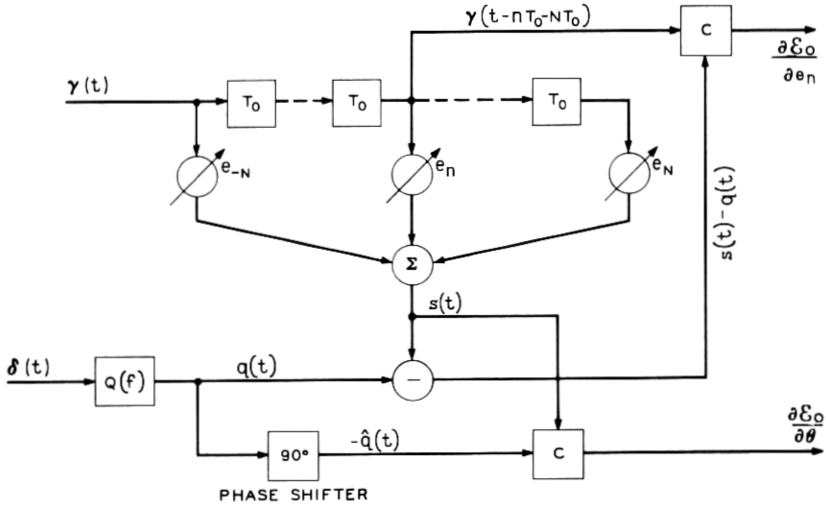

Fig. 2—Block diagram of the joint method. (C in the block denotes correlator.)

V. FURTHER ANALYSIS

We now analyze the system to answer the questions: (*i*) Will the mean-square error converge to the absolute minimum by the proposed adjustments? (*ii*) Is the convergence ensured regardless of how the parameters θ and $\{e_n\}$ are set prior to the starting of the process? (*iii*) What is the minimum mean-square error that can be obtained by the proposed method? To answer these questions, it is necessary to locate the stationary points of $\mathcal{E}_0$, to distinguish between various types of stationary points, to determine conditions under which $\mathcal{E}_0$ will converge to a global minimum, to derive explicit solutions of $\mathcal{E}_0$ when the parameters are set jointly or independently, and to obtain numerical data by simulating a real typical channel. In this section, we determine the location of the stationary points and distinguish between various types of stationary points.

It is noted from equations (16) and (17) that $\partial\mathcal{E}_0/\partial\theta$ changes sign when the carrier frequency f_c is shifted from one side of $A(f)$ (that is, $f_c \leqq f_1$) to the other side ($f_c \geqq f_2$). To avoid this complication, a quantity ρ defined by

$$\begin{aligned} \rho &= \theta, \quad \text{when} \quad f_c \leqq f_1 , \\ &= -\theta, \quad \text{when} \quad f_c \geqq f_2 , \end{aligned} \tag{19}$$

will be used instead of θ in the following. By this change, the results obtained in the sequel will hold for both $f_c \leqq f_1$ and $f_c \geqq f_2$. Hence, the position of f_c will not be identified further.

Since it is more convenient to use matrix notations in the following, time samples will be used. Let

$$\begin{aligned} s_k &= s(kT_0 + NT_0), \\ q_k &= q(kT_0 + NT_0), \\ \gamma_k &= \gamma(-kT_0). \end{aligned} \tag{20}$$

Then equation (6) becomes

$$\mathcal{E}_0 = \frac{1}{2f_0}\,\mathcal{E}$$

where

$$\mathcal{E} = \sum_{k=-\infty}^{\infty} (s_k - q_k)^2. \tag{21}$$

Since f_0 is a constant, minimizing $\mathcal{E}$ minimizes $\mathcal{E}_0$.

It can be easily shown that equation (21) can be written in the following matrix form

$$\mathcal{E} = \mathbf{E}'\mathbf{y}\mathbf{E} - 2\mathbf{E}'\mathbf{v} + \sum_{k=-\infty}^{\infty} q_k^2 \tag{22}$$

where the prime denotes transpose, the vector $\mathbf{E}$ and $\mathbf{v}$ and the matrix $\mathbf{y}$ are defined by

$$\mathbf{E} = \begin{bmatrix} e_{-N} \\ e_{-N+1} \\ \vdots \\ e_N \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} v_{-N} \\ v_{-N+1} \\ \vdots \\ v_N \end{bmatrix}, \tag{23}$$

$$\mathbf{y} = \begin{bmatrix} y_{-N,-N} & \cdots & y_{-N,N} \\ \vdots & & \vdots \\ y_{N,-N} & \cdots & y_{N,N} \end{bmatrix}, \tag{24}$$

$$y_{h,i} = \sum_{k=-\infty}^{\infty} \gamma_{h-k}\gamma_{i-k}, \tag{25}$$

$$v_n = \sum_{k=-\infty}^{\infty} q_k\gamma_{n-k}. \tag{26}$$

Clearly, $\mathbf{E}$ and $\mathbf{v}$ are $(2N + 1) \times 1$ vectors, and $\mathbf{y}$ is a $(2N + 1) \times (2N + 1)$ matrix. One can easily show that $\mathbf{y}$ is nonsingular.

We now evaluate the partial derivatives $\partial \mathcal{E}/\partial \rho$ and $\partial \mathcal{E}/\partial e_n$, $n = -N$ to N, to determine the optimum parameter settings. From equation (22),

$$\frac{\partial \mathcal{E}}{\partial \rho} = \frac{\partial}{\partial \rho} \mathbf{E}'\mathbf{y}\mathbf{E} - 2 \frac{\partial}{\partial \rho} \mathbf{E}'\mathbf{v}. \tag{27}$$

To evaluate the terms in equation (27), we note from sampling theorem that equation (25) can be written as

$$y_{h,i} = 2f_0 \int_{-\infty}^{\infty} \gamma(t)\gamma[t - (h - i)T_0]\, dt. \tag{28}$$

By Parseval's theorem, equation (28) becomes

$$y_{h,i} = 2f_0 \int_{-\infty}^{\infty} | \Gamma(f) |^2 \exp [-j2\pi f(h - i)T_0]\, df. \tag{29}$$

For single-sideband systems, the demodulating carrier phase ρ appears only in the phase characteristic of $\Gamma(f)$, while the amplitude characteristic $| \Gamma(f) |$ of $\Gamma(f)$ is independent of ρ. Therefore, from equation (29), $y_{h,i}$ is independent of ρ and

$$\frac{\partial}{\partial \rho} \mathbf{E}'\mathbf{y}\mathbf{E} = 0. \tag{30}$$

This greatly simplifies the results. (It is important to note that this simplification is not possible for double-sideband and vestigial-sideband systems because ρ appears in $| \Gamma(f) |$ in such systems.) Substituting equation (30) into equation (27) gives

$$\frac{\partial \mathcal{E}}{\partial \rho} = -2 \frac{\partial}{\partial \rho} \mathbf{E}'\mathbf{v}. \tag{31}$$

To evaluate equation (31), we convert the elements v_n of $\mathbf{v}$ into explicit functions of ρ. For single-sideband systems, $\Gamma(f)$ can be decomposed into the form

$$\begin{aligned} \Gamma(f) &= H(f) \exp(-j\rho), \qquad f \geqq 0; \\ &= H(f) \exp(j\rho), \qquad f \leqq 0; \end{aligned} \tag{32}$$

where $H(f)$ is independent of ρ.* From equations (26), (20), (4), and

* $H(f)$ is the Fourier transform of $\gamma(t)$ when $\rho = 0$. The amplitude and phase of $H(f)$ will be denoted by $|H(f)|$ and $\eta(f)$, respectively, that is, $H(f) = |H(f)| \exp [j\eta(f)]$.

(32), it can be shown that

$$v_n = \mu_n \exp(-j\rho) + \nu_n \exp(j\rho), \qquad n = -N \text{ to } N \tag{33}$$

where

$$\mu_n = \sum_{k=-\infty}^{\infty} q_k \int_0^{f_0} H(f) \exp[j2\pi f(k-n)T_0]\, df \tag{34}$$

and

$$\nu_n = \sum_{k=-\infty}^{\infty} q_k \int_{-f_0}^{0} H(f) \exp[j2\pi f(k-n)T_0]\, df. \tag{35}$$

Note that μ_n and ν_n are independent of ρ. From equation (33), we can evaluate the term $\partial/\partial\rho \mathbf{E}'\mathbf{v}$ in equation (31) to obtain

$$\frac{\partial \mathcal{E}}{\partial \rho} = 2j[\exp(-j\rho)\mathbf{E}'\boldsymbol{\mu} - \exp(j\rho)\mathbf{E}'\boldsymbol{\nu}] \tag{36}$$

where

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_{-N} \\ \mu_{-N+1} \\ \vdots \\ \mu_N \end{bmatrix}, \qquad \boldsymbol{\nu} = \begin{bmatrix} \nu_{-N} \\ \nu_{-N+1} \\ \vdots \\ \nu_N \end{bmatrix}. \tag{37}$$

The above is for $\partial\mathcal{E}/\partial\rho$. The other partial derivatives can be obtained from equation (22) as

$$\frac{\partial \mathcal{E}}{\partial \mathbf{E}} = 2\mathbf{yE} - 2\mathbf{v}. \tag{38}$$

A necessary condition for a specific ρ and $\mathbf{E}$ to be jointly optimum (that is, to jointly minimize $\mathcal{E}$) is that they satisfy

$$\frac{\partial \mathcal{E}}{\partial \rho} = 0 \tag{39}$$

and

$$\frac{\partial \mathcal{E}}{\partial \mathbf{E}} = \mathbf{0}. \tag{40}$$

There are special cases where the optimum setting of ρ is arbitrary (for instance, if an infinite length tapped delay line is used, the taps can always be adjusted to reduce the mean-square error $\mathcal{E}$ to zero regardless of how ρ is set). We shall not consider such special cases here. This implies, as can be shown from equations (36), (38), (39), and (40), that the special cases where $\mathbf{u}'\mathbf{y}^{-1}\mathbf{u}$ is zero are disregarded.

From equation (38), equation (40) can be written as

$$\mathbf{E} = \mathbf{y}^{-1}\mathbf{v}. \tag{41}$$

From equations (36) and (41), equation (39) can be written as

$$\exp(-j\rho)\mathbf{u}'\mathbf{y}^{-1}\mathbf{v} - \exp(j\rho)\mathbf{v}'\mathbf{y}^{-1}\mathbf{v} = 0. \tag{42}$$

One can show from equation (33) that equation (42) is equivalent to

$$\frac{\mathbf{u}'\mathbf{y}^{-1}\mathbf{u}}{\mathbf{v}'\mathbf{y}^{-1}\mathbf{v}} = \exp(j4\rho). \tag{43}$$

It can be shown from equations (34) and (35) that

$$\mathrm{Re}\,[\mathbf{u}'\mathbf{y}^{-1}\mathbf{u}] = \mathrm{Re}\,[\mathbf{v}'y^{-1}\mathbf{v}],$$

and

$$\mathrm{Im}\,[\mathbf{u}'\mathbf{y}^{-1}\mathbf{u}] = -\mathrm{Im}\,[\mathbf{v}'\mathbf{y}^{-1}\mathbf{v}], \tag{44}$$

where Re and Im denote real and imaginary parts of the complex number, respectively. From equation (44), one can show that equation (43) is satisfied if and only if

$$\rho = m_0\frac{\pi}{2} + \tfrac{1}{2}\tan^{-1}\frac{\mathrm{Im}\,[\mathbf{u}'\mathbf{y}^{-1}\mathbf{u}]}{\mathrm{Re}\,[\mathbf{u}'\mathbf{y}^{-1}\mathbf{u}]} \tag{45}$$

where m_0 can be any integer including zero. Substituting equation (45) into equation (33) and substituting the resulting equation into equation (41) give

$$\mathbf{E} = \exp\left[-j\left(m_0\frac{\pi}{2} + \tfrac{1}{2}\tan^{-1}\frac{\mathrm{Im}\,[\mathbf{u}'\mathbf{y}^{-1}\mathbf{u}]}{\mathrm{Re}\,[\mathbf{u}'\mathbf{y}^{-1}\mathbf{u}]}\right)\right]\mathbf{y}^{-1}\mathbf{u}$$
$$+ \exp\left[j\left(m_0\frac{\pi}{2} + \tfrac{1}{2}\tan^{-1}\frac{\mathrm{Im}\,[\mathbf{u}'\mathbf{y}^{-1}\mathbf{u}]}{\mathrm{Re}\,[\mathbf{u}'\mathbf{y}^{-1}\mathbf{u}]}\right)\right]\mathbf{y}^{-1}\mathbf{v}. \tag{46}$$

The value of ρ and $\mathbf{E}$ which satisfies the necessary conditions (39) and (40) is given by equations (45) and (46), respectively. It is clearly seen from equations (45) and (46) that, since m_0 can be any integer, there is more than one set of solutions of ρ and $\mathbf{E}$ from conditions (39) and (40). In the following, we determine which of these solutions actually minimizes $\mathcal{E}_0$.

In conventional terminology,[4] each solution of conditions (39) and

(40) is a stationary point of $\mathcal{E}_0$. We shall determine which of the stationary points is a global minimum.

We first identify the minima, maxima, and saddlepoints. As is well known,[4] a sufficient condition for a stationary point to be a local minimum is that the quadratic form

$$F = \mathbf{h}'\mathbf{Q}\mathbf{h} \tag{47}$$

be positive-definite at that stationary point, where $\mathbf{Q}$, known as the Hessian matrix, is

$$\mathbf{Q} = \begin{bmatrix} \dfrac{\partial^2 \mathcal{E}_0}{\partial \rho^2} & \dfrac{\partial^2 \mathcal{E}_0}{\partial \rho\, \partial e_{-N}} & \cdots & \dfrac{\partial^2 \mathcal{E}_0}{\partial \rho\, \partial e_N} \\ \dfrac{\partial^2 \mathcal{E}_0}{\partial e_{-N}\, \partial \rho} & \dfrac{\partial^2 \mathcal{E}_0}{\partial e_{-N}^2} & \cdots & \dfrac{\partial^2 \mathcal{E}_0}{\partial e_{-N}\, \partial e_N} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial^2 \mathcal{E}_0}{\partial e_N\, \partial \rho} & \dfrac{\partial^2 \mathcal{E}_0}{\partial e_N\, \partial e_{-N}} & \cdots & \dfrac{\partial^2 \mathcal{E}_0}{\partial e_N^2} \end{bmatrix}$$

and

$$\mathbf{h}' = [h_1 \quad h_2 \cdots h_{(2N+2)}].$$

A stationary point is a local maximum if F is negative-definite at that point. At a saddlepoint, F is indefinite. To evaluate F, note from equations (16), (17), and (19) that

$$\frac{\partial \mathcal{E}_0}{\partial \rho} = 2 \int_{-\infty}^{\infty} \hat{q}(t)s(t)\, dt \tag{48}$$

and

$$\frac{\partial s(t)}{\partial \rho} = \hat{s}(t). \tag{49}$$

From equations (48), (49), (8), and (18), we get

$$\frac{\partial^2 \mathcal{E}_0}{\partial \rho^2} = 2 \int_{-\infty}^{\infty} q(t)s(t)\, dt; \tag{50}$$

$$\frac{\partial^2 \mathcal{E}_0}{\partial \rho\, \partial e_n} = 2 \int_{-\infty}^{\infty} \gamma(t - nT_0 - NT_0)\hat{q}(t)\, dt, \qquad n = -N \text{ to } N; \tag{51}$$

and

$$\frac{\partial^2 \mathcal{E}_0}{\partial e_i\, \partial e_j} = 2 \int_{-\infty}^{\infty} \gamma(t - iT_0 - NT_0)\gamma(t - jT_0 - NT_0)\, dt, \qquad i, j = -N \text{ to } N. \tag{52}$$

Substituting equations (50), (51), and (52) into equation (47) and rearranging, we obtain after some steps

$$F = 2 \int_{-\infty}^{\infty} \left[h_1 \hat{q}(t) + \sum_{n=-N}^{N} h_{2+n+N} \gamma(t - nT_0 - NT_0) \right]^2 dt$$
$$+ 2h_1^2 \int_{-\infty}^{\infty} q(t)s(t)\, dt$$
$$- 2h_1^2 \int_{-\infty}^{\infty} q^2(t)\, dt. \qquad (53)$$

Now we may substitute equations (45) and (46) into equation (53) and evaluate the resulting expression over all possible h_1 to $h_{(2N+2)}$ to determine whether F is positive-definite, negative-definite, or indefinite at a given stationary point. This determines if that point is a local minimum, a local maximum, or a saddlepoint. While these steps are important in the analysis, they are rather complex and are therefore given in Appendix A. It is also shown in that appendix that all the local minima are equal and hence all are global minima. The results in Appendix A are summarized in the following proposition.

Proposition 1. $\mathcal{E}_0$ has a global minimum when ρ and $\mathbf{E}$ are given, respectively, by equations (45) and (46), with m_0 in these equation being an even integer. $\mathcal{E}_0$ has a saddlepoint when ρ and $\mathbf{E}$ are given, respectively, by equations (45) and (46), with m_0 in these equations being an odd integer.

It is seen from this proposition that there is an infinite number of global minima, each one corresponding to an even integer m_0. The distance in ρ between two adjacent global minima is therefore π. There is also an infinite number of saddlepoints, each one corresponding to an odd integer m_0. The distance in ρ between two adjacent saddlepoints is also π. The distance in ρ between a saddlepoint and its adjacent global minimum is $\pi/2$. It is instructive to illustrate these with an example and a figure. Consider a transversal equalizer with only one tap e_0 (such a single tap serves as an automatic gain control and the problem is to jointly set the automatic gain control and carrier phase to minimize the mean-square error). For simplicity, suppose that the term $\tan^{-1} [\mathrm{Im}\, (\mathbf{u}'\mathbf{y}^{-1}\mathbf{u})]/[\mathrm{Re}\, (\mathbf{u}'\mathbf{y}^{-1}\mathbf{u})]$ in equation (45) turned out to be zero. Then from the proposition $\mathcal{E}_0$ has global minima at $\rho = 0, \pm\pi \pm 2\pi, \cdots$, and $\mathcal{E}_0$ has saddlepoints at $\rho = \pm\pi/2, \pm 3\pi/2, \cdots$. The global minima at $\rho = 0$ and π are illustrated by points 1 and 3 in Fig. 3 and the saddlepoint at $\rho = \pi/2$ is illustrated by point 2.

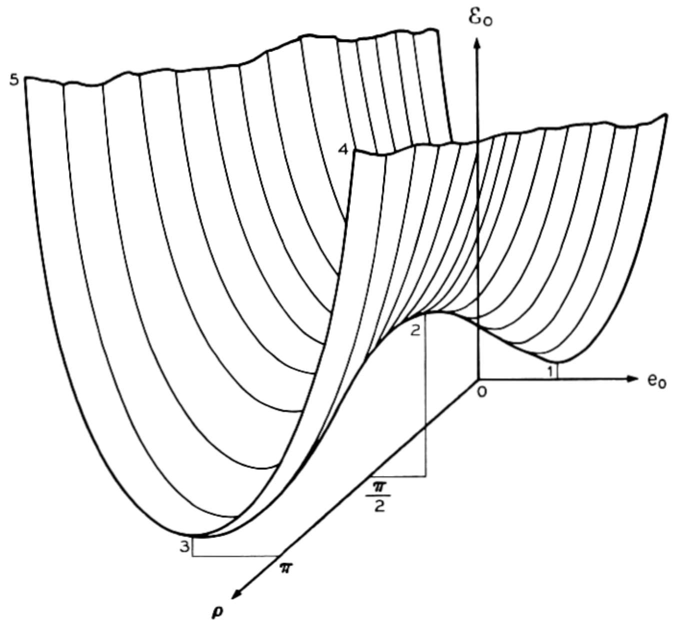

Fig. 3—An example illustrating the global minima and saddlepoints of $\mathcal{E}_0$.

The curved surface in the figure illustrates the variation of $\mathcal{E}_0$ with ρ and e_0. For instance, when ρ is fixed at π, $\mathcal{E}_0$ varies with e_0 as shown by the convex curve passing through points 4, 3, and 5. (For fixed ρ, $\mathcal{E}_0$ is a convex function of the tap gains e_{-N} to e_N.) As can be seen, point 2 is a saddlepoint because varying e_0 away from point 2 with ρ constant increases $\mathcal{E}_0$, while varying ρ away from point 2 with e_0 constant decreases $\mathcal{E}_0$.

To summarize, in this section we have located the global minima and saddlepoints and proved that there is no local minimum or maximum. There is also no valley[5] since the global minima are all distinct. These results will be used in the next section for the study of convergence and in Sections VII and VIII for the computation and comparison of performances.

VI. A CONDITION OF CONVERGENCE

As described in Section IV, after a test pulse is transmitted and the partial derivatives are computed, the carrier phase and the tap gains are adjusted according to the equation

$$\theta_{\text{new}} = \theta_{\text{old}} - \alpha \frac{\partial \mathcal{E}_0}{\partial \theta}, \tag{54}$$

$$(e_n)_{\text{new}} = (e_n)_{\text{old}} - \beta \frac{\partial \mathcal{E}_0}{\partial e_n}, \tag{55}$$

where α and β are positive proportional constants which may vary from one adjustment to another. As is well known,[5] the process always converges when α and β are sufficiently small, but may not converge if α and β are too large. In this section, we derive a condition on α and β which ensures the convergence of the process.

To facilitate analysis, we shall, as in Section V, replace θ with ρ, $\mathcal{E}_0$ with $\mathcal{E}$, the tap gains e_{-N} to e_N with the vector $\mathbf{E}$, and the partial derivatives $\partial\mathcal{E}/\partial e_{-N}$ to $\partial\mathcal{E}/\partial e_N$ with the vector $\partial\mathcal{E}/\partial\mathbf{E}$. The adjustment made after the kth test pulse will be called the kth adjustment ($k = 1, 2, \cdots$). The values of ρ, $\mathbf{E}$, $\mathcal{E}$, $\partial\mathcal{E}/\partial\rho$, and $\partial\mathcal{E}/\partial\mathbf{E}$ prior to the kth adjustment will be denoted by ρ_k, $\mathbf{E}_k$, $\mathcal{E}(\rho_k, \mathbf{E}_k)$, $(\partial/\partial\rho)\mathcal{E}(\rho_k, \mathbf{E}_k)$ and $(\partial/\partial\mathbf{E})\mathcal{E}(\rho_k, \mathbf{E}_k)$, respectively. The α and β used in the kth adjustment will be called α_k and β_k, respectively (note that $\alpha_k > 0$ and $\beta_k > 0$ for all k so that the adjustments will be made in the negative gradient direction). The values of ρ, $\mathbf{E}$, $\mathcal{E}$, $\partial\mathcal{E}/\partial\rho$, and $\partial\mathcal{E}/\partial\mathbf{E}$ after the kth adjustment will be denoted by ρ_{k+1}, $\mathbf{E}_{k+1}$, $\mathcal{E}(\rho_{k+1}, \mathbf{E}_{k+1})$, $(\partial/\partial\rho)\,\mathcal{E}(\rho_{k+1}, \mathbf{E}_{k+1})$, and $(\partial/\partial\mathbf{E})\,\mathcal{E}(\rho_{k+1}, \mathbf{E}_{k+1})$, respectively.

With the above notations, equations (53) and (54) can be written as

$$\rho_{k+1} = \rho_k - \alpha_k \frac{\partial}{\partial\rho}\,\mathcal{E}(\rho_k, \mathbf{E}_k), \tag{56}$$

$$\mathbf{E}_{k+1} = \mathbf{E}_k - \beta_k \frac{\partial}{\partial\mathbf{E}}\,\mathcal{E}(\rho_k, \mathbf{E}_k). \tag{57}$$

The decrease in mean-square error due to the kth adjustment is denoted by $\Delta\mathcal{E}_k$, that is,

$$\Delta\mathcal{E}_k = \mathcal{E}(\rho_k, \mathbf{E}_k) - \mathcal{E}(\rho_{k+1}, \mathbf{E}_{k+1}). \tag{58}$$

Clearly $\Delta\mathcal{E}_k$ approaches zero when the partial derivatives approach zero. A stronger statement that may sometimes hold is that "$\Delta\mathcal{E}_k$ approaches zero *only* when the partial derivatives approach zero." By this statement is meant that for every $\epsilon > 0$, we can find a $\delta > 0$ such that

$$|\Delta\mathcal{E}_k| \geqq \delta \tag{59}$$

if

$$\left[\frac{\partial}{\partial \rho}\,\mathcal{E}(\rho_k\,,\,\mathbf{E}_k)\right]^2 + \left[\frac{\partial}{\partial \mathbf{E}}\,\mathcal{E}(\rho_k\,,\,\mathbf{E}_k)\right]'\left[\frac{\partial}{\partial \mathbf{E}}\,\mathcal{E}(\rho_k\,,\,\mathbf{E}_k)\right] \geqq \epsilon. \tag{60}$$

We shall say that $\mathcal{E}$ converges to a stationary point if for every $\epsilon > 0$ there is an N such that

$$\left[\frac{\partial}{\partial \rho}\,\mathcal{E}(\rho_k\,,\,\mathbf{E}_k)\right]^2 + \left[\frac{\partial}{\partial \mathbf{E}}\,\mathcal{E}(\rho_k\,,\,\mathbf{E}_k)\right]'\left[\frac{\partial}{\partial \mathbf{E}}\,\mathcal{E}(\rho_k\,,\,\mathbf{E}_k)\right] < \epsilon \tag{61}$$

for all $k > N$.

The following lemma is needed in our discussion.

Lemma. If $\Delta\mathcal{E}_k > 0$ for all k and $\Delta\mathcal{E}_k$ approaches zero only when the partial derivatives approach zero, $\mathcal{E}$ must converge to a stationary point.

Proof: The proof of this lemma is simple. Assume that $\mathcal{E}$ does not converge to a stationary point, that is, for an $\epsilon > 0$ it is not possible to find an N such that equation (61) holds for all $k > N$. Then equation (59) must hold for an infinite number of k's. From this and the assumptions in the lemma, we see that $\Delta\mathcal{E}_k \geqq \delta > 0$ for infinite number of k's. This implies that $\mathcal{E}$ reduces without bound as k increases, contradicting the fact that the mean-square error cannot be less than zero. Hence the lemma holds.

The lemma provides a means of determining α_k and β_k. According to the lemma, $\mathcal{E}$ will converge to a stationary point if we can determine α_k and β_k such that $\Delta\mathcal{E}_k > 0$ for all k and $\Delta\mathcal{E}_k$ approaches zero only when the partial derivatives approach zero. Theoretically, such α_k and β_k can be determined using known mathematical techniques.[5] But unfortunately these techniques were developed for use with computers and are too complicated to be used in a receiver (for a discussion, see Appendix B). In the following proposition, it is shown that α_k and β_k can be determined rather easily if $\mathbf{E}$ and ρ are set in some alternate fashion.

Proposition 2. Let the set of positive integers be divided arbitrarily into two disjoint subsets $\mathcal{K}_1$ and $\mathcal{K}_2$, each containing an infinite number of positive integers. Let $\alpha_k = 0$ when $k \in \mathcal{K}_1$, and $\beta_k = 0$ when $k \in \mathcal{K}_2$. Let λ_1 denote the maximum eigenvalue of $\mathbf{y}$ ($\lambda_1 > 0$ since $\mathbf{y}$ is positive definite), and let $[\int_{-\infty}^{\infty} s^2(t)\,dt]_{\mathbf{E}_k}$ denote the value of $\int_{-\infty}^{\infty} s^2(t)\,dt$ when $\mathbf{E} = \mathbf{E}_k$. The mean-square error $\mathcal{E}$ will converge to a stationary point if

$$0 < \beta_k < \frac{1}{\lambda_1} \tag{62}$$

for $k \varepsilon \mathcal{K}_1$, and

$$0 < \alpha_k < \frac{2^{\frac{1}{2}}}{2f_0\left[\int_{-\infty}^{\infty} s^2(t)\, dt\right]_{\mathbf{E}_k} + 2f_0 \int_{-\infty}^{\infty} q^2(t)\, dt} \tag{63}$$

for $k \varepsilon \mathcal{K}_2$.

The proof of this proposition is complex and is given in Appendix C. The proposition states that ρ and $\mathbf{E}$ may be adjusted in any alternating fashion† and the mean-square error $\mathcal{E}$ will converge to a stationary point if β_k satisfies equation (62) during the adjustment of $\mathbf{E}$, and α_k satisfies equation (63) during the adjustment of ρ. The term λ_1 in equation (62) and the terms $[\int_{-\infty}^{\infty} s^2(t)\, dt]_{\mathbf{E}_k}$ and $\int_{-\infty}^{\infty} q^2(t)\, dt$ in equation (63) can be estimated or measured. Consider first λ_1, the maximum eigenvalue of $\mathbf{y}$. There are various methods of estimating the maximum eigenvalue of a matrix.[6,7] For example, it is possible to estimate λ_1 from amplitude characteristics of the transmission medium. [Note from Fig. 1 that the amplitude and phase characteristics of the transmission medium are denoted by $| W(f) |$ and $\omega(f)$, respectively.] It is seen from equation (29) that the elements $y_{h,i}$ of $\mathbf{y}$ can be computed from $| \Gamma(f) |$. For single-sideband systems, $| \Gamma(f) |$ depends on $| W(f) |$, but not on $\omega(f)$ and the demodulating carrier phase ρ. Thus, if statistics of $| W(f) |$ are available (for example, in a voiceband system), $y_{h,i}$ can be computed from equation (29) and λ_1 can be estimated. The maximum possible value of λ_1 then can be used instead of λ_1 in equation (62).

The factor $\int_{-\infty}^{\infty} q^2(t)dt$ in equation (63) is known because the desired signal $q(t)$ is given. The other factor $[\int_{-\infty}^{\infty} s^2(t)dt]_{\mathbf{E}_k}$ is simply the energy of the signal $s(t)$ prior to the kth adjustment, and can be easily measured at the equalizer output.

Summarizing the above, a condition of convergence has been described in the lemma. Based on the lemma, a specific condition of convergence has been obtained in Proposition 2. The upper bound in equation (62) can be estimated from *a priori* channel statistics and the upper bound in equation (63) can be easily determined prior to each adjustment. Thus, α_k and β_k can be set accordingly prior to the adjustments to ensure that $\mathcal{E}$ will converge to a stationary point.

It has been shown in the previous section that a stationary point must be a global minimum or a saddlepoint. Thus, $\mathcal{E}$ may converge

† For example, one may fix ρ and adjust $\mathbf{E}$ until $\partial\mathcal{E}/\partial\mathbf{E}$ approaches zero, then fix $\mathbf{E}$ and adjust ρ until $\partial\mathcal{E}/\partial\rho$ approaches zero, and repeat the cycle until both $\partial\mathcal{E}/\partial\mathbf{E}$ and $\partial\mathcal{E}/\partial\rho$ approach zero.

to a saddlepoint instead of a global minimum. Fortunately, such a possibility is remote. A great advantage of gradient methods is that they will inherently stay away from saddlepoints.[5] It has been found by researchers that gradient search computer program avoids saddlepoints so dependably that the only way they could test their program for exploring the neighborhood of a pass was to start the search there. Wilde and Beightler suggested the reason by sketching a bimodal surface to show that only one gradient line out of the infinite number possible actually passes through the saddlepoint.[5] The other gradient lines all lead directly to a minimum or a maximum. Hence the possibility of converging to a saddlepoint is remote.

It has been shown in the previous section that the distance in ρ between a saddlepoint and its adjacent global minima is $\pi/2$. From this, several tests can be devised for distinguishing between global minima and saddlepoints. The following one is particularly simple. At the design stage of the system, one may compute the value of $\mathcal{E}$ at global minima and saddlepoints (the detailed steps in Section VII may be used. The value of $\mathcal{E}$ at global minima is equal to $\mathcal{E}_{min}$ in equation (64), while the value of $\mathcal{E}$ at saddlepoints is equal to $\mathcal{E}_{ind}$ evaluated at $\Delta\rho = \pi/2$.) As illustrated in Section 8.1 and Fig. 4, the value of $\mathcal{E}$ at saddlepoints can be many times larger than that at global minima. Consequently, a threshold can be set up such that when $\mathcal{E}$ converges to a value above the threshold it may be concluded that a saddlepoint

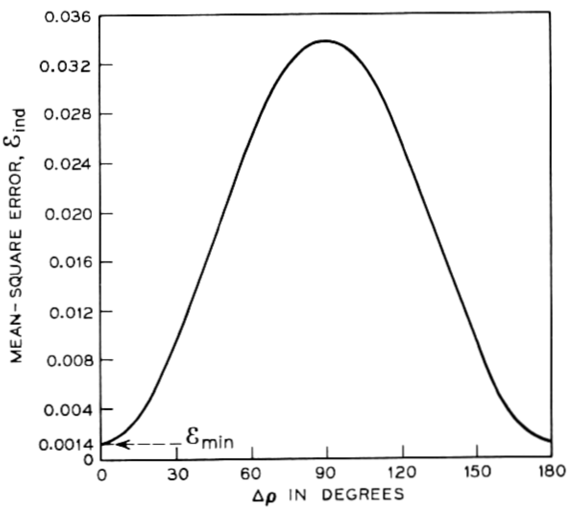

Fig. 4—Variation of $\mathcal{E}_{ind}$ with Δp. (See example in Section VIII.)

is reached. Then ρ will be shifted by $\pi/2$ and $\mathbf{E}$ readjusted. If $\mathcal{E}$ converges to a value below the threshold, the process is terminated. This and other possible tests all require additional circuitry. Since the possibility of converging to a saddlepoint is remote, one should decide from actual trials whether such a test should be employed.

VII. PERFORMANCE COMPARISON

In the previous sections, we have considered jointly setting ρ and $\mathbf{E}$ for minimizing the mean-square error $\mathcal{E}$. It has been shown in Proposition 1 that the joint optimum settings of ρ and $\mathbf{E}$ are given, respectively, by equation (45) and (46) (with m_0 in these equations being an even integer). Substituting these optimum settings into equation (22), we obtain the minimum mean-square error

$$\mathcal{E}_{\min} = \sum_{k=-\infty}^{\infty} q_k^2 - 2 \mid \underline{\mathbf{u}}'\mathbf{y}^{-1}\underline{\mathbf{u}} \mid - 2\underline{\mathbf{u}}'\mathbf{y}^{-1}\mathbf{v}. \tag{64}$$

It has been shown that this $\mathcal{E}_{\min}$ can be obtained by setting ρ and $\mathbf{E}$ jointly. Now we wish to compare $\mathcal{E}_{\min}$ with what can be obtained by another scheme. In single-sideband systems, it is possible to transmit a carrier pilot to the receiver to demodulate the received signal. The demodulating carrier phase therefore is

$$\rho = \rho_c + \rho_f,$$

where ρ_c is the phase of the received carrier pilot and ρ_f is a fixed phase shift that is sometimes introduced for signal shaping. With ρ fixed at $\rho_c + \rho_f$, $\mathbf{E}$ can be adjusted to minimize $\mathcal{E}$. The value of $\mathcal{E}$ thus obtained will be denoted by $\mathcal{E}_{\text{ind}}$, where the subscript "ind" indicates that ρ and $\mathbf{E}$ are set independently. We now compare $\mathbf{E}_{\text{ind}}$ with $\mathbf{E}_{\min}$.

To determine $\mathbf{E}_{\text{ind}}$, let the difference between $\rho_c + \rho_f$ and the optimum setting of ρ be denoted by $\Delta\rho$. From equation (45),

$$\Delta\rho = \rho_c + \rho_f - \left[m_0 \frac{\pi}{2} + \tfrac{1}{2} \tan^{-1} \frac{\text{Im}\,[\underline{\mathbf{u}}'\mathbf{y}^{-1}\underline{\mathbf{u}}]}{\text{Re}\,[\underline{\mathbf{u}}'\mathbf{y}^{-1}\underline{\mathbf{u}}]} \right] \tag{65}$$

where m_0 can be any even integer (including zero). Since $\mathbf{E}$ is set to minimize $\mathcal{E}$ after ρ is set to $\rho_c + \rho_f$, $\mathbf{E}$ is equal to $\mathbf{y}^{-1}\mathbf{v}$ evaluated at $\rho = \rho_c + \rho_f$, that is,

$$\mathbf{E} = \mathbf{y}^{-1}\mathbf{v}_c \tag{66}$$

where $\mathbf{v}_c$ is $\mathbf{v}$ evaluated at $\rho = \rho_c + \rho_f$. Substituting equation (66) into equation (22) gives

$$\mathcal{E}_{\text{ind}} = \sum_{k=-\infty}^{\infty} q_k^2 - \mathbf{v}_c'\mathbf{y}^{-1}\mathbf{v}_c \,. \tag{67}$$

It can easily be shown from equation (33) that

$$\mathbf{v}_c'\mathbf{y}^{-1}\mathbf{v}_c = \exp\left[-j2(\rho_c + \rho_f)\right]\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\mu} + 2\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\nu} + \exp\left[j2(\rho_c + \rho_f)\right]\boldsymbol{\nu}'\mathbf{y}^{-1}\boldsymbol{\nu}. \tag{68}$$

Substituting equation (65) into equation (68), and substituting the resulting equation into equation (67), one can obtain after some manipulations that

$$\mathcal{E}_{\text{ind}} = \sum_{k=-\infty}^{\infty} q_k^2 - 2(\cos 2\,\Delta\rho)\,|\,\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\mu}\,| - 2\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\nu}. \tag{69}$$

From equations (64) and (69), the difference between $\mathcal{E}_{\text{ind}}$ and $\mathcal{E}_{\text{min}}$ is

$$\mathcal{E}_{\text{ind}} - \mathcal{E}_{\text{min}} = 2[1 - \cos 2\Delta\rho]\,|\,\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\mu}\,|. \tag{70}$$

Note that $\boldsymbol{\mu}'$ and $\mathbf{y}^{-1}$ are independent of $\Delta\rho$. Thus, the only term in equation (70) that depends on $\Delta\rho$ is $\cos 2\Delta\rho$. Now we can make the following observations.

First, $\mathcal{E}_{\text{ind}} - \mathcal{E}_{\text{min}}$ is nonnegative, meaning that independently setting ρ and $\mathbf{E}$ increases the mean-square error. Second, $\mathcal{E}_{\text{ind}} - \mathcal{E}_{\text{min}}$ varies periodically with $\Delta\rho$ with a period of π. Third, because of the nature of cosine, $\mathcal{E}_{\text{ind}} - \mathcal{E}_{\text{min}}$ is small when $\Delta\rho$ is small, but increases rapidly when $\Delta\rho$ increases (note the factor two in $\cos 2\Delta\rho$).

For a given system, one may compute $\mathcal{E}_{\text{ind}}$, $\mathcal{E}_{\text{min}}$, and their difference from equations (64) and (69). The computation can best be carried out by a computer program in the following steps (see also the example in the next section). First, specify f_0, N, and the desired signal $q(t)$. Determine the time samples $\{q_k\}$ of $q(t)$. Second, determine $H(f)$ from transfer functions of the transmitting filter, the transmission medium, and the receiving filter. Third, compute the elements $y_{h,i}$ of $\mathbf{y}$ from the equation

$$y_{h,i} = 4f_0 \int_0^{f_0} |\,H(f)\,|^2 \cos\left[2\pi f(h - i)T_0\right] df. \tag{71}$$

Compute the elements $a_{h,i}$ of $\mathbf{y}^{-1}$ from $y_{h,i}$. Fourth, compute the terms $|\,\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\mu}\,|$ and $\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\nu}$ in equations (64) and (69) using the explicit equations

$$|\,\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\mu}\,| = [(B_1 - B_2)^2 + 4B_3^2]^{\frac{1}{2}}, \tag{72}$$

$$\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\nu} = B_1 + B_2\,, \tag{73}$$

where

$$B_1 = \sum_{h=-N}^{N} \sum_{i=-N}^{N} a_{h,i}\xi_h\xi_i , \tag{74}$$

$$B_2 = \sum_{h=-N}^{N} \sum_{i=-N}^{N} a_{h,i}\zeta_h\zeta_i , \tag{75}$$

$$B_3 = \sum_{h=-N}^{N} \sum_{i=-N}^{N} a_{h,i}\xi_h\zeta_i , \tag{76}$$

$$\xi_i = \sum_{k=-\infty}^{\infty} q_k \int_0^{f_0} | H(f) | \cos [\eta(f) + 2\pi f(kT_0 - iT_0)] \, df, \tag{77}$$

$$\zeta_i = \sum_{k=-\infty}^{\infty} q_k \int_0^{f_0} | H(f) | \sin [\eta(f) + 2\pi f(kT_0 - iT_0)] \, df. \tag{78}$$

Finally, compute $\mathcal{E}_{\min}$ from equation (64) and $\mathcal{E}_{\text{ind}}$ from equation (69).

VIII. COMPUTER SIMULATIONS

We now apply the previous results to a single-sideband partial-response voiceband data communication system using a five- or nine-tap transversal equalizer. Since the results are largely similar, we shall describe only the five-tap case.

Two computer programs have been written for this system. In the first program, $\mathcal{E}_{\min}$ and $\mathcal{E}_{\text{ind}}$ are computed using the formulas in the previous section. In the second one, $\mathcal{E}_{\min}$ and $\mathcal{E}_{\text{ind}}$ are obtained using the method of steepest descent. The results of these two programs are described separately in the following subsections.

8.1 *Comparison of* $\mathcal{E}_{\min}$ *and* $\mathcal{E}_{\text{ind}}$ *Using the Explicit Formulas in Section VII*

Consider the voiceband data communication system described in Ref. 8. The desired signal $q(t)$ of such a system is the class IV partial-response signal,[9] that is,

$$q(t) = \frac{2^{\frac{1}{2}}\pi \sin 2\pi f_0(t - t_0)}{[2\pi f_0(t - t_0)]^2 - \pi^2} , \tag{79}$$

where a time delay t_0 is included to take into account the time delay in the channel. From Ref. 8,

$$f_0 = 1200 \text{ Hz}. \tag{80}$$

Hence, $T_0 = 1/2f_0 = 1/2400$ seconds. We shall consider a transversal equalizer with five taps, that is, $N = 2$. It has been defined that $H(f)$ is the Fourier transform of $\gamma(t)$ when $\rho = 0$. The amplitude and phase of $H(f)$ have been denoted by $|H(f)|$ and $\eta(f)$, respectively, that is,

$$H(f) = |H(f)| \exp [j\eta(f)]. \tag{81}$$

It is assumed that the amplitude and phase distortions in the filters are negligible compared with those in the transmission medium. Consequently, $|H(f)|$ and $\eta(f)$ can be determined from amplitude and phase characteristics of the transmission medium (which is a voice channel in this case). From Ref. 10 and Fig. 4 of Ref. 8, a typical $|H(f)|$ is found to be

$$|H(f)| = [\sin 2\pi T_0 f]10^{0.3T_0(f_0-2f)}, \qquad 0 \leqq f \leqq f_0 . \tag{82}$$

The component $\sin 2\pi T_0 f$ in equation (82) represents the desired amplitude characteristic of the class IV partial-response signal, while the term $10^{0.3T_0(f_0-2f)}$ represents typical amplitude distortion in a voice channel. The delay characteristic of five links of K carrier shown in Fig. 24 of Ref. 11 is representative of the delay characteristic of a voice channel. Therefore, it will be used here and

$$\eta(f) = 9.89 \sin [2\pi(f + f_0)\cdot 0.00019 - 2.203], \qquad 0 \leqq f \leqq f_0 . \tag{83}$$

Constant phases and time delays in the transmission medium and the filters are omitted in writing $\eta(f)$ because their values are not available and because they only change the phase and time origins in the computations. However, such an omission makes it impossible to determine the phase ρ_c of the received carrier pilot because the carrier pilot does not travel the same path as the signal. The signal is transmitted through the transmitting and receiving filters, while the carrier pilot is transmitted outside of these filters (these filters theoretically should have infinite attenuations at the carrier frequency). Furthermore, the carrier pilot is recovered at the receiver through a separate narrowband filter or a phase-lock loop, while the signal is demodulated and passes through a low-pass filter. With these differences and without detailed phase characteristics of the filters, it is not possible to determine ρ_c here. Therefore, we shall simply leave ρ_c and $\Delta\rho$ as variables and compute the variation of $\mathcal{E}_{\text{ind}}$ with $\Delta\rho$. (It should be noted that $\Delta\rho$ may assume small values in some real systems.)

The variation of $\mathcal{E}_{\text{ind}}$ with $\Delta\rho$ has been determined for various values of t_0. The curve obtained at $t_0 = -2T_0$ is typical and is presented in Fig. 4. The value of $\mathcal{E}_{\min}$ is also indicated in the figure. It can be seen that $\mathcal{E}_{\text{ind}}$ can be as large as 0.034, while $\mathcal{E}_{\min}$ is only 0.0014. Thus, the mean-square error can increase 0.034/0.0014 = 24.3 times if the demodulating carrier phase ρ is not set properly. Note that this large increase is obtained for a five-tap transversal equalizer. A similar result has been obtained for a nine-tap transversal equalizer. These results demonstrate that the mean-square error depends critically on the carrier phase setting when the number of equalizer taps is not large. (It should be noted, however, that when a large number of taps are used, the mean-square error will not be sensitive to the carrier phase setting.)

8.2 *Mean-Square Errors Obtained by Method of Steepest Descent*

A computer program has been written to simulate the system described in Section 8.1. The receiver parameters are adjusted by the method of steepest descent described in Section IV. Five hundred test pulses are used in each training period. The initial setting of the center tap e_0 is unity, while the initial settings of the other taps are zero. The parameter t_0 is fixed at $-2T_0$ as mentioned in Section 8.1. The mean-square errors obtained with various initial settings of ρ are shown in Fig. 5.

The points marked by "×" in Fig. 5 are obtained by adjusting ρ and **E** jointly using $\alpha_k = 0.523$ and $\beta_k = 0.1$ for all k. For instance, point A is obtained by initially setting ρ to 40 degrees above the optimum

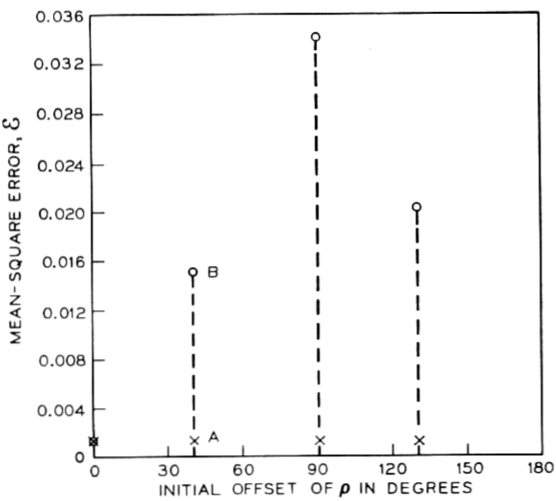

Fig. 5—Mean-square errors obtained by method of steepest descent.

ρ and then adjusting ρ and **E** jointly by the method of steepest descent. It can be seen from these points that the mean-square error $\mathcal{E}$ always converges to the value of $\mathcal{E}_{min}$ determined in Section 8.1. This demonstrates the fact that $\mathcal{E}_{min}$ can be obtained by jointly setting ρ and **E**.

The other points marked by circles are obtained by fixing ρ and adjusting **E** only ($\alpha_k = 0$ and $\beta_k = 0.1$ for all k). For instance, point B is obtained by initially setting ρ to 40 degrees above the optimum ρ and then adjusting **E** only by the method of steepest descent. It can be seen, by comparing these circled points with the value of $\mathcal{E}_{ind}$ in Fig. 4, that the mean-square error $\mathcal{E}$ converges, as expected, to $\mathcal{E}_{ind}$.

Only 500 test pulses are used in each training period because $\mathcal{E}$ converges within this period in all cases. For instance, the convergence of $\mathcal{E}$ to the values shown by points A and B in Fig. 5 are illustrated, respectively, by curves A and B in Fig. 6. It can be seen that $\mathcal{E}$ converges rapidly to the final values.

IX. SUMMARY AND CONCLUSIONS

It is shown in this paper that in single-sideband systems the transversal equalizer and the carrier phase can be set jointly by the method of steepest descent to minimize the mean-square equalization error.

The system is analyzed and a receiver structure is developed. The receiver structure is theoretically as simple as a conventional one.

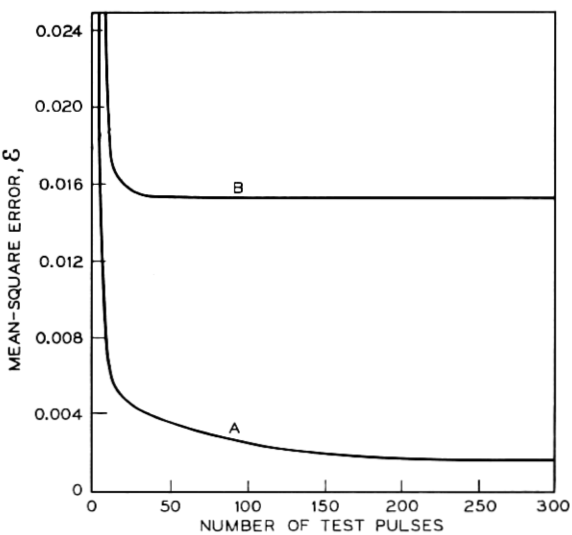

Fig. 6—An illustration of convergence.

A well-known problem of the method of steepest descent is that the function to be minimized may converge to a local minimum instead of to a global minimum. To prove that this troublesome problem does not arise here, the variation of the mean-square error in the parameter space is analyzed. Exact locations of the stationary points in the parameter space are determined and the classifications of the stationary points are obtained. It is shown that the mean-square error has only global minima and saddlepoints, but not local minimum or maximum. Thus, it is not possible for the adjustments to converge to a local minimum, regardless of the initial parameter settings. This completely eliminates the problem.

It is also shown that two adjacent global minima (or two adjacent saddlepoints) are separated by 180 degrees in demodulating carrier phase, while a global minimum is separated from its adjacent saddlepoints by 90 degrees in demodulating carrier phase. From this, a test is described to determine whether a global minimum or a saddlepoint is reached by the adjustments and to correct the settings if a saddlepoint is reached. This test may not be necessary because both previous experience and theoretical considerations have shown that the method of steepest descent inherently stays away from saddlepoints.

The choice of the step sizes of the adjustments is also considered. There are methods for determining the step sizes; however, they require complicated computations. As an alternative, it is shown in this paper that the step sizes can be easily determined if the equalizer and the demodulating carrier phase are adjusted in different steps (the steps can be alternated in any manner).

Closed-form expressions of the joint optimum parameter settings and of the corresponding minimum mean-square error are obtained for computation of system performance. For illustration purposes, a single-sideband data communication system using a five- or nine-tap transversal equalizer is simulated on a computer. Both theoretical and simulation results show that the equalization error can increase by ten times or more when the carrier phase is not properly set. These demonstrate that when the number of equalizer taps is not large, the equalization error depends critically on the carrier phase setting. The computer simulation also verifies the theory that the equalization error can be minimized by using the joint method described in this paper.

X. ACKNOWLEDGMENTS

I wish to thank Miss A. C. Weingartner for writing a computer program to simulate the data communication system described in Section

VIII, and Miss C. A. Reichenstein for the computer program used in Section 8.1.

APPENDIX A

Classification of Stationary Points

In this appendix, we prove the statements that:

(*i*) F is positive definite when ρ and $\mathbf{E}$ are given, respectively, by equations (45) and (46), and m_0 in these equations is an even integer;
(*ii*) F is indefinite when ρ and $\mathbf{E}$ are given, respectively, by equations (45) and (46), and m_0 in these equations is an odd integer; and
(*iii*) All the local minima of $\mathcal{E}_0$ are equal.

Consider the first statement. Let Ω be the set of all $\mathbf{h}$ except $\mathbf{h} = \mathbf{0}$. Let Ω be decomposed into two disjoint sets Ω_1 and Ω_2, where Ω_1 contains the h's with $h_1 \neq 0$, and Ω_2 contains the $\mathbf{h}$'s with $h_1 = 0$. Consider first Ω_1. Since $h_1 \neq 0$ in Ω_1, equation (53) can be written as

$$F = 2h_1^2 \int_{-\infty}^{\infty} [\sigma(t) - q(t)]^2 \, dt + 2h_1^2 \int_{-\infty}^{\infty} q(t)s(t) \, dt - 2h_1^2 \int_{-\infty}^{\infty} q^2(t) \, dt \tag{84}$$

where

$$\sigma(t) = \sum_{n=-N}^{N} e_n' \hat{\gamma}(t - nT_0 - NT_0) \tag{85}$$

and

$$e_n' = \frac{h_{2+n+N}}{h_1}, \qquad n = -N \quad \text{to} \quad N. \tag{86}$$

For the sake of brevity, we shall denote the entire right side of equation (45) by ρ_0 and the entire right side of equation (46) by $\mathbf{E}_0$. For any function G, the symbol $[G]_{\rho_0}$ denotes the value of G evaluated at $\rho = \rho_0$, the symbol $[G]_{\rho_0, \mathbf{E}_0}$ denotes the value of G evaluated at $\rho = \rho_0$ and $\mathbf{E} = \mathbf{E}_0$, and the symbol Min G denotes the minimum value of G in Ω_1. These symbols and notations may be used jointly. For instance, Min $[G]_{\rho_0, \mathbf{E}_0}$ denotes the minimum value of $[G]_{\rho_0, \mathbf{E}_0}$ in Ω_1.

Since $h_1 \neq 0$ in Ω_1, $[F]_{\rho_0, \mathbf{E}_0} > 0$ in Ω_1 if and only if $[F/2h_1^2]_{\rho_0, \mathbf{E}_0} > 0$ in Ω_1, or if and only if Min $[F/2h_1^2]_{\rho_0, \mathbf{E}_0} > 0$. From equation (84),

$$\text{Min}\left[\frac{F}{2h_1^2}\right]_{\rho_0,\mathbf{E}_0} = \text{Min}\left[\int_{-\infty}^{\infty}[\sigma(t) - q(t)]^2\,dt\right]_{\rho_0} + \left[\int_{-\infty}^{\infty} q(t)s(t)\,dt\right]_{\rho_0,\mathbf{E}_0} - \left[\int_{-\infty}^{\infty} q^2(t)\,dt\right]. \tag{87}$$

In writing the above equation, we have used the fact that $\sigma(t)$ is independent of $\mathbf{E}$, $s(t)$ is independent of $\mathbf{h}$, and $q(t)$ is independent of ρ, $\mathbf{E}$, and $\mathbf{h}$.

Consider the first term on the right side of equation (87). From equation (85), $\sigma(t)$ is a function of $\hat{\gamma}(t)$, where $\hat{}$ indicates Hilbert transform. It can be shown from equations (7) and (19) that

$$[\hat{\gamma}(t)]_{\rho_0} = [\gamma(t)]_{\rho_0+\pi/2}\,. \tag{88}$$

Now we have

$$\begin{aligned}
&\text{Min}\left[\int_{-\infty}^{\infty}[\sigma(t) - q(t)]^2\,dt\right]_{\rho_0}\\
&= \text{Min}\int_{-\infty}^{\infty}\left[\sum_{n=-N}^{N} e_n'[\hat{\gamma}(t - nT_0 - NT_0)]_{\rho_0} - q(t)\right]^2 dt,\\
&= \text{Min}\int_{-\infty}^{\infty}\left[\sum_{n=-N}^{N} e_n'[\gamma(t - nT_0 - NT_0)]_{\rho_0+\pi/2} - q(t)\right]^2 dt,\\
&= \left\{\text{Min}\int_{-\infty}^{\infty}\left[\sum_{n=-N}^{N} e_n'\gamma(t - nT_0 - NT_0) - q(t)\right]^2 dt\right\}_{\rho_0+\pi/2}.
\end{aligned} \tag{89}$$

Note that equation (88) is used in the second step above. Since the term $\sum_{n=-N}^{N} e_n'\gamma(t - nT_0 - NT_0)$ in the last expression above is similar to $s(t)$ in equation (8), the whole term in the { } in the last expression above is equal to $\mathcal{E}_0$ minimized with respect to $\mathbf{E}$. This proves that $\text{Min}\,[\int_{-\infty}^{\infty}[\sigma(t) - q(t)]^2\,dt]_{\rho_0}$ is equal to $\mathcal{E}_0$ minimized with respect to $\mathbf{E}$ and evaluated at $\rho = \rho_0 + \pi/2$. It can be easily shown from equations (21) and (22) that $\mathcal{E}_0$ minimized with respect to $\mathbf{E}$ is equal to $1/2f_0[\sum_{k=-\infty}^{\infty} q_k' - \mathbf{v}'\mathbf{y}^{-1}\mathbf{v}]$; hence,

$$\begin{aligned}
&\text{Min}\left[\int_{-\infty}^{\infty}[\sigma(t) - q(t)]^2\,dt\right]_{\rho_0}\\
&= \frac{1}{2f_0}\sum_{k=-\infty}^{\infty} q_k^2 - \frac{1}{2f_0}[\mathbf{v}'\mathbf{y}^{-1}\mathbf{v}]_{\rho_0+\pi/2}\,,\\
&= \frac{1}{2f_0}\sum_{k=-\infty}^{\infty} q_k^2 - \frac{1}{2f_0}[2\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\nu} - \exp(-j2\rho_0)\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\mu} - \exp(j2\rho_0)\boldsymbol{\nu}'\mathbf{y}^{-1}\boldsymbol{\nu}],
\end{aligned} \tag{90}$$

where equation (33) has been used in the last step.

We have evaluated the first term in equation (87). Consider now the second term in equation (87). From equations (20), (23), and (26),

$$\int_{-\infty}^{\infty} q(t)s(t)\,dt = \frac{1}{2f_0}\,\mathbf{v}'\mathbf{E}. \tag{91}$$

Substituting equations (33), (45), and (46) into equation (91) and simplifying, we obtain, after some steps,

$$\left[\int_{-\infty}^{\infty} q(t)s(t)\,dt\right]_{\rho_0,\mathbf{E}_0} = \frac{1}{f_0}\,\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\nu} + \frac{1}{2f_0}\exp(-j2\rho_0)\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\mu} + \frac{1}{2f_0}\exp(j2\rho_0)\boldsymbol{\nu}'\mathbf{y}^{-1}\boldsymbol{\nu}. \tag{92}$$

The last term in equation (87) is

$$\int_{-\infty}^{\infty} q^2(t)\,dt = \frac{1}{2f_0}\sum_{k=-\infty}^{\infty} q_k^2\,. \tag{93}$$

Substituting equations (90), (92), and (93) into (87) and canceling out the terms having opposite signs, we obtain

$$\operatorname{Min}\left[\frac{F}{2h_1^2}\right]_{\rho_0,\mathbf{E}_0} = \frac{1}{f_0}\exp(-j2\rho_0)\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\mu} + \frac{1}{f_0}\exp(j2\rho_0)\boldsymbol{\nu}'\mathbf{y}^{-1}\boldsymbol{\nu}. \tag{94}$$

Note that so far m_0 can be any integer. Let $|\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\mu}|$ denote the magnitude of $\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\mu}$. It can be shown from equation (44) that when m_0 is an even integer, the right side of equation (94) is equal to $2/f_0|\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\mu}|$. Hence, when m_0 is an even integer (including zero),

$$\operatorname{Min}\left[\frac{F}{2h_1^2}\right]_{\rho_0,\mathbf{E}_0} = \frac{2}{f_0}\,|\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\mu}|\,. \tag{95}$$

Since $|\boldsymbol{\mu}'\mathbf{y}^{-1}\boldsymbol{\mu}| \neq 0$, equation (95) shows that $\operatorname{Min}[F/2h_1^2]_{\rho_0,\mathbf{E}_0} > 0$. Thus, $[F]_{\rho_0,\mathbf{E}_0} > 0$ in Ω_1 when m_0 is an even integer.

Now consider Ω_2. Since $h_1 = 0$ in Ω_2, equation (53) is reduced to

$$F = 2\int_{-\infty}^{\infty}\left[\sum_{n=-N}^{N} h_{2+n+N}\gamma(t - nT_0 - NT_0)\right]^2 dt. \tag{96}$$

Since $\gamma(t - nT_0 - NT_0)$, $n = -N$ to N, are linearly independent and h_2 to h_{2N+2} cannot be all zero in Ω_2, the integrand in equation (96) cannot be zero for all t. Hence, $F > 0$ in Ω_2. This implies, of course, $[F]_{\rho_0,\mathbf{E}_0} > 0$ in Ω_2.

We have shown that when m_0 is an even integer, $[F]_{\rho_0,\mathbf{E}_0} > 0$ in Ω_1 and Ω_2. Hence, when m_0 is an even integer, $[F]_{\rho_0,\mathbf{E}_0} > 0$ for all $\mathbf{h}$ except $\mathbf{h} = 0$. This proves the first statement at the beginning of this Appendix.

Now we prove the second statement at the beginning of this Appendix. Note that the derivations from equations (84) through (94) hold for all integer m_0. It can be shown from equation (44) that when m_0 is an odd integer, the right side of equation (94) is equal to $-2/f_0 | \boldsymbol{\mu}' \mathbf{y}^{-1} \boldsymbol{\mu} |$. Hence, when m_0 is an odd integer, equation (94) becomes

$$\text{Min} \left[\frac{F}{2h_1^2} \right]_{\rho_0, \mathbf{E}_0} = -\frac{2}{f_0} | \boldsymbol{\mu}' \mathbf{y}^{-1} \boldsymbol{\mu} | . \tag{97}$$

Since $| \boldsymbol{\mu}' \mathbf{y}^{-1} \boldsymbol{\mu} | \neq 0$, $\text{Min} [F/2h_1^2]_{\rho_0, \mathbf{E}_0} < 0$. Thus, $[F]_{\rho_0, \mathbf{E}_0}$ can be negative in Ω_1 (and hence in Ω). It can also be positive in Ω (for instance, $[F]_{\rho_0, \mathbf{E}_0} > 0$ in Ω_2). Therefore, F is indefinite when ρ and $\mathbf{E}$ are given, respectively, by equations (45) and (46), and m_0 in these equations is an odd integer. This proves the second statement.

Finally, consider the third statement. We have shown that F is positive definite and $\mathcal{E}_0$ has a local minimum when ρ and $\mathbf{E}$ are given, respectively, by equations (45) and (46), and m_0 in these equations is an even integer. Substituting equations (45) and (46) into equation (22) and letting m_0 be an even integer, we see that the local minimum of $\mathcal{E}_0$ is

$$\mathcal{E}_0 = \frac{1}{2f_0} \sum_{k=-\infty}^{\infty} q_k^2 - \frac{1}{f_0} | \boldsymbol{\mu}' \mathbf{y}^{-1} \boldsymbol{\mu} | - \frac{1}{f_0} \boldsymbol{\mu}' \mathbf{y}^{-1} \boldsymbol{\nu}. \tag{98}$$

Since the right side of equation (98) is independent of m_0, all the local minima of $\mathcal{E}_0$ are equal and hence are all global minima. This proves the third statement at the beginning of this Appendix.

APPENDIX B

Discussion of Existing Method

A number of methods are described in Ref. 5 for determining the proportional constant in steepest descent adjustments. It has been pointed out that these methods can be used to solve certain nonlinear equations on computer.[12]

These methods require elaborate computations. For example, consider the possibility of determining α_k and β_k using the theorem on page 31 of Ref. 5. Since it is assumed in that theorem that a single proportional constant is used for all parameters, we change the scale of ρ or $\mathbf{E}$ such that it is also appropriate to use a single proportional constant in our case. After this is accomplished, we can use a single constant α_k in equations (56) and (57). To determine α_k, define

$$g(\rho_k, \mathbf{E}_k, \alpha_k) = \frac{\Delta \mathcal{E}_k}{\alpha_k \left\{ \left[\frac{\partial}{\partial \rho} \mathcal{E}(\rho_k, \mathbf{E}_k) \right]^2 + \sum_{n=-N}^{N} \left[\frac{\partial}{\partial e_n} \mathcal{E}(\rho_k, \mathbf{E}_k) \right]^2 \right\}} . \tag{99}$$

Choose a constant δ in the range $0 < \delta \leqq \frac{1}{2}$. Compute $g(\rho_k, \mathbf{E}_k, 1)$. If $g(\rho_k, \mathbf{E}_k, 1) < \delta$, choose $\alpha_k < 1$ so that $\delta \leqq g(\rho_k, \mathbf{E}_k, \alpha_k) \leqq 1 - \delta$ (it has been shown that this choice is always possible). If $g(\rho_k, \mathbf{E}_k, 1) \geqq \delta$, choose $\alpha_k = 1$.

It can be seen from this description that the method requires elaborate computations and α_k must be determined on a trial and error basis when $g(\rho_k, \mathbf{E}_k, 1) < \delta$. It is therefore difficult to use this method in a receiver.

APPENDIX C

Proof of Proposition 2

In this Appendix, we prove Proposition 2. Consider first $k \in \mathcal{K}_1$. Since $\alpha_k = 0$ when $k \in \mathcal{K}_1$, we have $\rho_{k+1} = \rho_k$ and

$$\begin{aligned} \Delta \mathcal{E}_k &= \mathcal{E}(\rho_k, \mathbf{E}_k) - \mathcal{E}(\rho_k, \mathbf{E}_{k+1}), \\ &= \mathbf{E}_k'\mathbf{y}\mathbf{E}_k - 2\mathbf{E}_k'\mathbf{v} + \sum_{h=-\infty}^{\infty} q_h^2 \\ &\quad - \left[\mathbf{E}_{k+1}'\mathbf{y}\mathbf{E}_{k+1} - 2\mathbf{E}_{k+1}'\mathbf{v} + \sum_{h=-\infty}^{\infty} q_h^2 \right]. \end{aligned} \tag{100}$$

From equations (57) and (38)

$$\mathbf{E}_{k+1} = \mathbf{E}_k - \beta_k(2\mathbf{y}\mathbf{E}_k - 2\mathbf{v}). \tag{101}$$

Substituting equation (101) into equation (100), we rearrange the resulting equation into the form

$$\Delta\mathcal{E}_k = 4\beta_k(\mathbf{y}\mathbf{E}_k - \mathbf{v})'[\mathbf{I} - \beta_k\mathbf{y}](\mathbf{y}\mathbf{E}_k - \mathbf{v}) \tag{102}$$

where $\mathbf{I}$ is the identity matrix. Let the eigenvalues of $\mathbf{y}$ be denoted, in the order of decreasing magnitude, by λ_i, $i = 1$ to $2N + 1$, so that

$$\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_{2N+1} . \tag{103}$$

It can be shown that $\mathbf{y}$ is positive definite; hence,

$$\lambda_i > 0, \qquad i = 1 \text{ to } 2N + 1. \tag{104}$$

Let $\mathbf{u}_i$, $i = 1$ to $2N + 1$, be a set of orthonormal eigenvectors of $\mathbf{y}$. Let

$$\mathbf{Q} = [\mathbf{u}_1\mathbf{u}_2 \cdots \mathbf{u}_{2N+1}]. \tag{105}$$

That is, the ith column of $\mathbf{Q}$ is $\mathbf{u}_i$. From well-known matrix properties,

$$\mathbf{y} = \mathbf{QDQ}' \tag{106}$$

where $\mathbf{D}$ is a diagonal matrix whose ith diagonal element is λ_i. Furthermore,

$$\mathbf{QQ}' = \mathbf{I}. \tag{107}$$

Substituting equations (106) and (107) into equation (102), one can write

$$\Delta \mathcal{E}_k = 4\beta_k \sum_{i=1}^{2N+1} (1 - \beta_k\lambda_i)[(\mathbf{yE}_k - \mathbf{v})'\mathbf{u}_i]^2. \tag{108}$$

If β_k satisfies equation (62), we can write

$$1 - \beta_k\lambda_1 \geqq P_1 > 0. \tag{109}$$

Then, from equations (103), (108), and (109), we have

$$\Delta \mathcal{E}_k \geqq 4\beta_k P_1 \sum_{i=1}^{2N+1} [(\mathbf{yE}_k - \mathbf{v})'\mathbf{u}_i]^2. \tag{110}$$

It can be shown from equations (38), (107), and (105) that

$$\left[\frac{\partial}{\partial \mathbf{E}}\, \mathcal{E}(\rho_k , \mathbf{E}_k)\right]' \left[\frac{\partial}{\partial \mathbf{E}}\, \mathcal{E}(\rho_k , \mathbf{E}_k)\right] = 4 \sum_{i=1}^{2N+1} [(\mathbf{yE}_k - \mathbf{v})'\mathbf{u}_i]^2. \tag{111}$$

Comparing equations (110) and (111) gives

$$\Delta \mathcal{E}_k \geqq \beta_k P_1 \left[\frac{\partial}{\partial \mathbf{E}}\, \mathcal{E}(\rho_k , \mathbf{E}_k)\right]' \left[\frac{\partial}{\partial \mathbf{E}}\, \mathcal{E}(\rho_k , \mathbf{E}_k)\right]. \tag{112}$$

Now we have the conclusion:

Conclusion 1. If equation (62) holds, $\Delta\mathcal{E}_k$ is bounded below by equation (112) which shows that $\Delta\mathcal{E}_k$ is positive and approaches zero only when the partial derivatives $(\partial/\partial\mathbf{E})\mathcal{E}(\rho_k , \mathbf{E}_k)$ approach zero.

The above is for $k \in \mathcal{K}_1$. Next we consider $k \in \mathcal{K}_2$. Since $\beta_k = 0$ when $k \in \mathcal{K}_2$, we have $\mathbf{E}_{k+1} = \mathbf{E}_k$. From equations (58) and (22)

$$\begin{aligned} \Delta \mathcal{E}_k &= \mathcal{E}(\rho_k , \mathbf{E}_k) - \mathcal{E}(\rho_{k+1} , \mathbf{E}_k), \\ &= -2\mathbf{E}_k'(\mathbf{v}_k - \mathbf{v}_{k+1}), \end{aligned} \tag{113}$$

where

$$\mathbf{v}_k = \exp(-j\rho_k)\boldsymbol{\mu} + \exp(j\rho_k)\boldsymbol{\nu}, \tag{114}$$

$$\mathbf{v}_{k+1} = \exp(-j\rho_{k+1})\boldsymbol{\mu} + \exp(j\rho_{k+1})\boldsymbol{\nu}. \tag{115}$$

From equation (56) and the two equations above,

$$\mathbf{v}_k = \mathbf{v}_{k+1} = \left\{1 - \cos\left[\alpha_k \frac{\partial}{\partial\rho}\, \mathcal{E}(\rho_k\,, \mathbf{E}_k)\right]\right\}[\exp(-j\rho_k)\boldsymbol{\mu}_u + \exp(j\rho_k)\mathbf{v}]$$
$$- j\sin\left[\alpha_k \frac{\partial}{\partial\rho}\, \mathcal{E}(\rho_k\,, \mathbf{E}_k)\right][\exp(-j\rho_k)\boldsymbol{\mu} - \exp(j\rho_k)\mathbf{v}]. \quad (116)$$

The vectors $\boldsymbol{\mu}$ and $\mathbf{v}$ are defined in equation (37) and their elements are defined in equations (34) and (35). It can be seen from equations (34) and (35) that the elements of $\boldsymbol{\mu}$ and $\mathbf{v}$ are complex numbers. Let $\boldsymbol{\mu}$ be written as

$$\boldsymbol{\mu} = \boldsymbol{\xi} + j\boldsymbol{\zeta} \quad (117)$$

where the elements of $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$ are real numbers that can be determined from equation (34). Comparing equation (35) with equation (34), we see that the elements of $\mathbf{v}$ are complex conjugates of those of $\boldsymbol{\mu}$. Hence $\mathbf{v}$ can be written as

$$\mathbf{v} = \boldsymbol{\xi} - j\boldsymbol{\zeta}. \quad (118)$$

Substituting equations (117) and (118) into equation (116) and then substituting the resulting equation into equation (113), we obtain after some steps,

$$\Delta\,\mathcal{E}_k = 4\left\{\cos\left[\alpha_k \frac{\partial}{\partial\rho}\, \mathcal{E}(\rho_k\,, \mathbf{E}_k)\right] - 1\right\}[\mathbf{E}_k'\boldsymbol{\xi}\cos\rho_k + \mathbf{E}_k'\boldsymbol{\zeta}\sin\rho_k]$$
$$- 4\sin\left[\alpha_k \frac{\partial}{\partial\rho}\, \mathcal{E}(\rho_k\,, \mathbf{E}_k)\right][\mathbf{E}_k'\boldsymbol{\zeta}\cos\rho_k - \mathbf{E}_k'\boldsymbol{\xi}\sin\rho_k]. \quad (119)$$

From equations (36), (117), and (118),

$$\frac{\partial}{\partial\rho}\, \mathcal{E}(\rho_k\,, \mathbf{E}_k) = 2j\mathbf{E}_k'(\exp(-j\rho_k)\boldsymbol{\mu} - \exp(j\rho_k)\mathbf{v}),$$
$$= 2j\mathbf{E}_k'(-2j\sin\rho_k\boldsymbol{\xi} + 2j\cos\rho_k\boldsymbol{\zeta}),$$
$$= 4[\mathbf{E}_k'\boldsymbol{\xi}\sin\rho_k - \mathbf{E}_k'\boldsymbol{\zeta}\cos\rho_k]. \quad (120)$$

It is clear from equation (56) that, if $(\partial/\partial\rho)\mathcal{E}(\rho_k\,, \mathbf{E}_k) = 0$, then $\rho_{k+1} = \rho_k$ and from equation (115), $\Delta\mathcal{E}_k = 0$. So we need to evaluate $\Delta\mathcal{E}_k$ only for $(\partial/\partial\rho)\mathcal{E}(\rho_k\,, \mathbf{E}_k) \neq 0$. It can be seen from equation (120) that, when $(\partial/\partial\rho)\mathcal{E}(\rho_k\,, \mathbf{E}_k) \neq 0$, $\mathbf{E}_k'\boldsymbol{\xi}$ and $\mathbf{E}_k'\boldsymbol{\zeta}$ cannot be simultaneously zero. Hence, we can define a quantity

$$\phi = \tan^{-1} \frac{\mathbf{E}_k' \zeta}{\mathbf{E}_k' \xi} \tag{121}$$

and a quantity

$$C = [(\mathbf{E}_k' \xi)^2 + (\mathbf{E}_k' \zeta)^2]^{\frac{1}{2}} > 0. \tag{122}$$

From equations (121) and (122), equation (120) can be rewritten as

$$\frac{\partial}{\partial \rho} \mathcal{E}(\rho_k, \mathbf{E}_k) = 4C \sin (\rho_k - \phi). \tag{123}$$

Substituting equation (123) into equation (119), using equations (121) and (122), and rearranging, we obtain, after a number of steps,

$$\Delta \mathcal{E}_k = 4C\{\cos [\rho_k - \phi - 4C\alpha_k \sin (\rho_k - \phi)] - \cos (\rho_k - \phi)\} \tag{124}$$

or

$$\Delta \mathcal{E}_k = 8C \sin [\rho_k - \phi - 2C\alpha_k \sin (\rho_k - \phi)] \sin [2C\alpha_k \sin (\rho_k - \phi)]. \tag{125}$$

It may be assumed that $0 \leqq \rho_k - \phi \leqq 2\pi$ (a factor of 2π or its multiple can be dropped). Since $(\partial/\partial\rho)\mathcal{E}(\rho_k, \mathbf{E}_k) \neq 0$, $\rho_k - \phi$ cannot be 0, π, or 2π [see equation (123)]. Hence, we have either

$$0 < \rho_k - \phi < \pi \tag{126}$$

or

$$\pi < \rho_k - \phi < 2\pi. \tag{127}$$

We shall consider equation (126) first.

As can be seen from equation (56), the constant α_k determines the size of each adjustment (note that $\alpha_k > 0$ is required so that the adjustments will be made in the negative gradient direction). We wish to determine the permissible range of α_k, that is, to determine a number δ such that $\Delta\mathcal{E}_k > 0$ for every α_k in the range $0 < \alpha_k < \delta$. It can be shown from equations (124) and (126) that δ can only be as large as $1/2C$; hence, the permissible range of α_k is

$$0 < \alpha_k < \frac{1}{2C}. \tag{128}$$

From equations (128) and (126), we obtain

$$0 < 2C\alpha_k \sin (\rho_k - \phi) < 1. \tag{129}$$

It can be easily shown from equation (129) that

$$\sin [2C\alpha_k \sin (\rho_k - \phi)] > \frac{1}{\pi} 4C\alpha_k \sin (\rho_k - \phi) > 0. \tag{130}$$

It can be shown from equations (126) and (128) that $[\rho_k - \phi - 2C\alpha_k \sin(\rho_k - \phi)]$ must be in the range

$$0 < [\rho_k - \phi - 2C\alpha_k \sin(\rho_k - \phi)] \leq \frac{\pi}{2} \tag{131}$$

or in the range

$$\frac{\pi}{2} < [\rho_k - \phi - 2C\alpha_k \sin(\rho_k - \phi)] < \pi. \tag{132}$$

When equation (131) holds,

$$\sin[\rho_k - \phi - 2C\alpha_k \sin(\rho_k - \phi)] \geqq \frac{2}{\pi}[\rho_k - \phi - 2C\alpha_k \sin(\rho_k - \phi)]. \tag{133}$$

Combining (133) and (130) and using equations (125) and (123), we obtain, after some simplification,

$$\Delta \mathcal{E}_k > \frac{2\alpha_k}{\pi^2}[2 - 4C\alpha_k]\left[\frac{\partial}{\partial\rho}\mathcal{E}(\rho_k, \mathbf{E}_k)\right]^2. \tag{134}$$

It can be shown in a similar manner that when equation (132) holds,

$$\Delta \mathcal{E}_k > \frac{8C\alpha_k^2}{\pi^2}\left[\frac{\partial}{\partial\rho}\mathcal{E}(\rho_k, \mathbf{E}_k)\right]^2. \tag{135}$$

It has been shown from equations (126) and (128) that either equation (134) or equation (135) holds. In a similar manner, it can be shown from equations (127) and (128) that either equation (134) or equation (135) holds. Thus, we have the conclusion:

Conclusion 2. When equation (128) holds, either equation (134) or equation (135) holds. Physically this means that, when equation (128) holds, $\Delta\mathcal{E}_k$ is positive and approaches zero only when $(\partial/\partial\rho)\mathcal{E}(\rho_k, \mathbf{E}_k)$ approaches zero.

It is not easy to determine whether equation (128) is satisfied in practice because the constant C depends on $\mathbf{E}_k$, $\boldsymbol{\xi}$, and $\boldsymbol{\zeta}$ and is somewhat difficult to compute. Hence, we wish to replace equation (128) with inequality (63) in Proposition 2.

From the definition of $\mathbf{v}_k$ in equation (114) and from equations (117) and (118), we can write

$$2\mathbf{E}_k'\mathbf{v}_k = 4[\cos\rho_k]\mathbf{E}_k'\boldsymbol{\xi} + 4[\sin\rho_k]\mathbf{E}_k'\boldsymbol{\zeta}. \tag{136}$$

From equations (136) and (22), one can show that

$$L_k - \mathcal{E}(\rho_k, \mathbf{E}_k) = 4\mathbf{E}_k'\boldsymbol{\xi} \cos \rho_k + 4\mathbf{E}_k'\boldsymbol{\zeta} \sin \rho_k \tag{137}$$

where

$$L_k = \mathbf{E}_k'\mathbf{y}\mathbf{E}_k + \sum_{h=-\infty}^{\infty} q_h^2 .$$

Note that the terms $\mathbf{E}_k'\mathbf{y}\mathbf{E}_k$, $\sum_{h=-\infty}^{\infty} q_h^2$, $\mathbf{E}_k'\boldsymbol{\xi}$, and $\mathbf{E}_k'\boldsymbol{\zeta}$, in equation (137) are independent of ρ. Furthermore, equation (137) holds for all ρ_k. Letting $\rho_k = 0$ in equation (137), we obtain

$$L_k - [\mathcal{E}(\rho_k, \mathbf{E}_k)]_{\rho_k=0} = 4\mathbf{E}_k'\boldsymbol{\xi} \tag{138}$$

where $[\mathcal{E}(\rho_k, \mathbf{E}_k)]_{\rho_k=0}$ is $\mathcal{E}(\rho_k, \mathbf{E}_k)$ evaluated at $\rho_k = 0$. Since $\mathcal{E}(\rho_k, \mathbf{E}_k) \geqq 0$ for all ρ_k, we have from equation (138)

$$L_k \geqq 4\mathbf{E}_k'\boldsymbol{\xi}. \tag{139}$$

It can be shown from equation (22) that

$$[\mathcal{E}(\rho_k, \mathbf{E}_k)]_{\rho_k=0} < 2L_k . \tag{140}$$

From equations (140) and (138),

$$-L_k < 4\mathbf{E}_k'\boldsymbol{\xi}. \tag{141}$$

From equations (139) and (141),

$$(\mathbf{E}_k'\boldsymbol{\xi})^2 \leqq \tfrac{1}{16}L_k^2 . \tag{142}$$

We have obtained equation (142) by letting $\rho_k = 0$ in equation (137). If we let $\rho_k = \pi/2$ in equation (137), we obtain, in a similar manner,

$$(\mathbf{E}_k'\boldsymbol{\zeta})^2 \leqq \tfrac{1}{16}L_k^2 . \tag{143}$$

From equations (142), (143), and (122),

$$C \leqq \frac{1}{2(2)^{\frac{1}{2}}} L_k . \tag{144}$$

Using sampling theory, we can verify that

$$L_k = 2f_0\left[\int_{-\infty}^{\infty} s^2(t)\, dt\right]_{\mathbf{E}_k} + 2f_0 \int_{-\infty}^{\infty} q^2(t)\, dt \tag{145}$$

where $[\int_{-\infty}^{\infty} s^2(t)\, dt]_{\mathbf{E}_k}$ is $\int_{-\infty}^{\infty} s^2(t)\, dt$ evaluated at $\mathbf{E} = \mathbf{E}_k$.

It can be seen from equations (144) and (145) that, if equation (63) holds, equation (128) is satisfied. Thus, we have the conclusion:

Conclusion 3. When equation (63) holds, equation (128) holds and, from Conclusion 2, either equation (134) or equation (135) holds. Hence, when equation (63) holds, $\Delta \mathcal{E}_k$ is positive and approaches zero only when $(\partial/\partial\rho)\mathcal{E}(\rho_k, \mathbf{E}_k)$ approaches zero.

Now, let us summarize the results presented in this appendix. It has been shown that when equation (62) holds in $\mathcal{K}_1$, $\Delta \mathcal{E}_k$ is bounded below by equation (112) (see Conclusion 1). It has also been shown that when equation (63) holds in $\mathcal{K}_2$, $\Delta \mathcal{E}_k$ is bounded below by either equation (134) or equation (135) (see Conclusion 3). It is seen from the lower bounds in equations (112), (134), and (135) that $\Delta \mathcal{E}_k$ is positive and approaches zero only when the partial derivatives approach zero. Since $\mathcal{K}_1$ and $\mathcal{K}_2$ each contain an infinite number of k's and since the mean-square error $\mathcal{E}$ cannot reduce without bound, $\mathcal{E}$ must converge to a stationary point. This proves Proposition 2 in Section VI.

REFERENCES

1. Lucky, R. W., Salz, J., and Weldon, E. J., Jr., *Principles of Data Communication*, New York: McGraw-Hill, 1968, pp. 128–165, 170.
2. Bennett, W. R., and Davey, J. R., *Data Transmission*, New York: McGraw-Hill, 1965, pp. 269–273.
3. 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.
4. Wilde, D. J., and Beightler, C. S., *Foundations of Optimization*, Englewood Cliffs, N. J.: Prentice-Hall, 1967, pp. 24–27, 296.
5. Goldstein, A. A., *Constructive Real Analysis*, New York: Harper and Row, 1967, pp. 26–39.
6. Marcus, M., and Minc, H., *A Survey of Matrix Theory and Matrix Inequalities*, Boston: Allyn and Bacon, 1964, pp. 139–148.
7. Gersho, A., "Adaptive Equalization of Highly Dispersive Channels for Data Transmission," B.S.T.J., *48*, No. 1 (January 1969), pp. 55–70.
8. Becker, F. K., Kretzmer, E. R., and Sheehan, J. R., "A New Signal Format for Efficient Data Transmission," B.S.T.J., *45*, No. 5 (May–June 1966), pp. 755–758.
9. Kretzmer, E. R., "Generalization of a Technique for Binary Data Communication," IEEE Trans. on Comm. Tech., *COM-14*, No. 1 (February 1966), pp. 67–68.
10. Alexander, A. A., Gryb, R. M., and Nast, D. W., "Capabilities of the Telephone Network for Data Transmission," B.S.T.J., *39*, No. 3 (May 1960), pp. 431–476.
11. Gibby, R. A., "An Evaluation of AM Data System Performance by Computer Simulation," B.S.T.J., *39*, No. 3 (May 1960), pp. 675–704.
12. Gersho, A., "Solving Nonlinear Network Equations Using Optimization Techniques," B.S.T.J., *48*, No. 9 (November 1969), pp. 3135–3138.