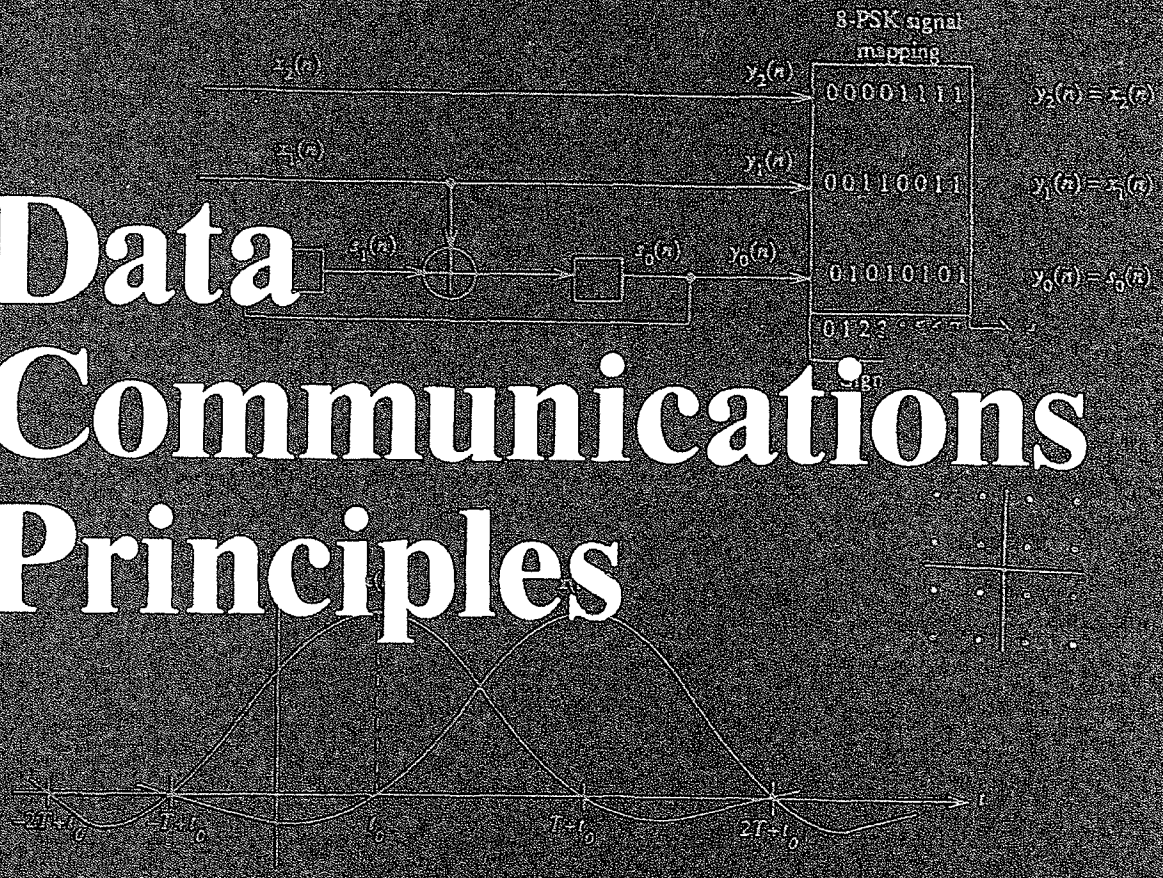
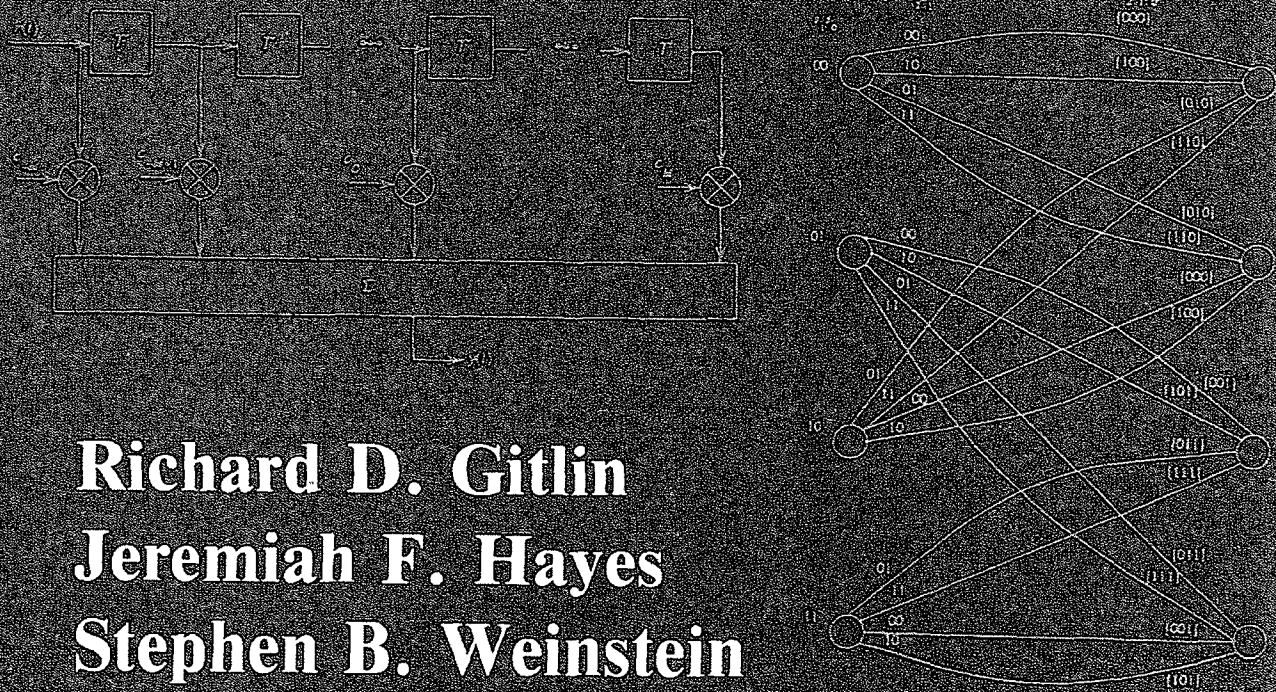
Applications of Communications Theory
Series Editor: R. W. Lucky

Data Communications Principles





Richard D. Gitlin
Jeremiah F. Hayes
Stephen B. Weinstein

Data Communications Principles

Applications of Communications Theory
Series Editor: R. W. Lucky, *AT&T Bell Laboratories*

Recent volumes in the series:

COMPUTER COMMUNICATIONS AND NETWORKS
John R. Freer

COMPUTER NETWORK ARCHITECTURES AND PROTOCOLS
Second Edition • Edited by Carl A. Sunshine

DATA COMMUNICATIONS PRINCIPLES
Richard D. Gitlin, Jeremiah F. Hayes, and Stephen B. Weinstein

DATA TRANSPORTATION AND PROTECTION
John E. Hershey and R. K. Rao Yarlagadda

DEEP SPACE TELECOMMUNICATIONS SYSTEMS ENGINEERING
Edited by Joseph H. Yuen

DIGITAL PHASE MODULATION
John B. Anderson, Tor Aulin, and Carl-Erik Sundberg

DIGITAL PICTURES: Representation and Compression
Arun N. Netravali and Barry G. Haskell

FIBER OPTICS: Technology and Applications
Stewart D. Personick

FUNDAMENTALS OF DIGITAL SWITCHING
Second Edition • Edited by John C. McDonald

MODELING AND ANALYSIS OF COMPUTER COMMUNICATIONS NETWORKS
Jeremiah F. Hayes

MODERN TELECOMMUNICATIONS
E. Bryan Carne

OPTICAL CHANNELS: Fibers, Clouds, Water, and the Atmosphere
Sherman Karp, Robert M. Gagliardi, Steven E. Moran, and Larry B. Stotts

PRACTICAL COMPUTER DATA COMMUNICATIONS
William J. Barksdale

SIMULATION OF COMMUNICATION SYSTEMS
Michel C. Jeruchim, Philip Balaban, and K. Sam Shanmugan

Data Communications Principles

Richard D. Gitlin
AT&T Bell Laboratories
Holmdel, New Jersey

Jeremiah F. Hayes
Concordia University
Montreal, Quebec

Stephen B. Weinstein
Bell Communications Research
Morristown, New Jersey

Plenum Press • New York and London

Library of Congress Cataloging-in-Publication Data

Gitlin, Richard D.
Data communications principles / Richard D. Gitlin, Jeremiah F. Hayes, and Stephen B. Weinstein.
p. cm. -- (Applications of communications theory)
Includes bibliographical references and index.
ISBN 0-306-43777-5
1. Data transmission systems. 2. Computer networks. I. Hayes, Jeremiah F., 1934- . II. Weinstein, Stephen B. III. Title.
IV. Series.
TK5105.G57 1992
621.382--dc20 92-19019
CIP

10 9 8 7 6 5 4

ISBN 0-306-43777-5

9

Echo Cancellation

9.0 INTRODUCTION

The simplest way to realize *full-duplex* data communication is by using completely separate transmission media for the two directions of transmission. But often, only a single, bilateral (simultaneous two-way) transmission medium is available rather than two one-way channels. For the ordinary telephone subscriber, served by a twisted-pair copper line from a central office, a dialed telephone circuit is a bilateral channel, usable for voice or data (at speeds up to about 14.4 kbps). It should be recognized that this circuit normally traverses interoffice facilities, in addition to the subscriber lines at each end, and that these facilities have a large influence on the characteristics of the end-to-end circuit.

There is a second type of bilateral channel becoming available to the telephone subscriber, and that is the channel offered by the twisted-pair copper line by itself, between the subscriber and the central office. For many subscriber locations, this subscriber line will, in the future, be dedicated to a full-duplex digital access channel to ISDN and other digital networks. The characteristics of the subscriber line alone are, of course, quite different from that of the dialed telephone line because of the absence of interoffice facilities and their restrictions, such as bandlimiting filters. Full-duplex signaling is carried out at 160 kbps (144 kbps information payload) for ISDN Basic access and faster for the high-speed digital subscriber line, as explained in Chapter 4.

There are several ways to use the available bandwidth to accomplish full-duplex data transmission over bilateral channels such as these, which can be summarized as *bandsplitting, time sharing, spectrum overlap,* and *same spectrum in both directions* (Figure 9.1). For bandsplitting, completely separate high-frequency and low-frequency subbands are used in each direction, with bandsplitting filters to avoid self-interference at each end of the circuit (Figure 9.2). The local transmitter can easily be 40 dB stronger in power than the

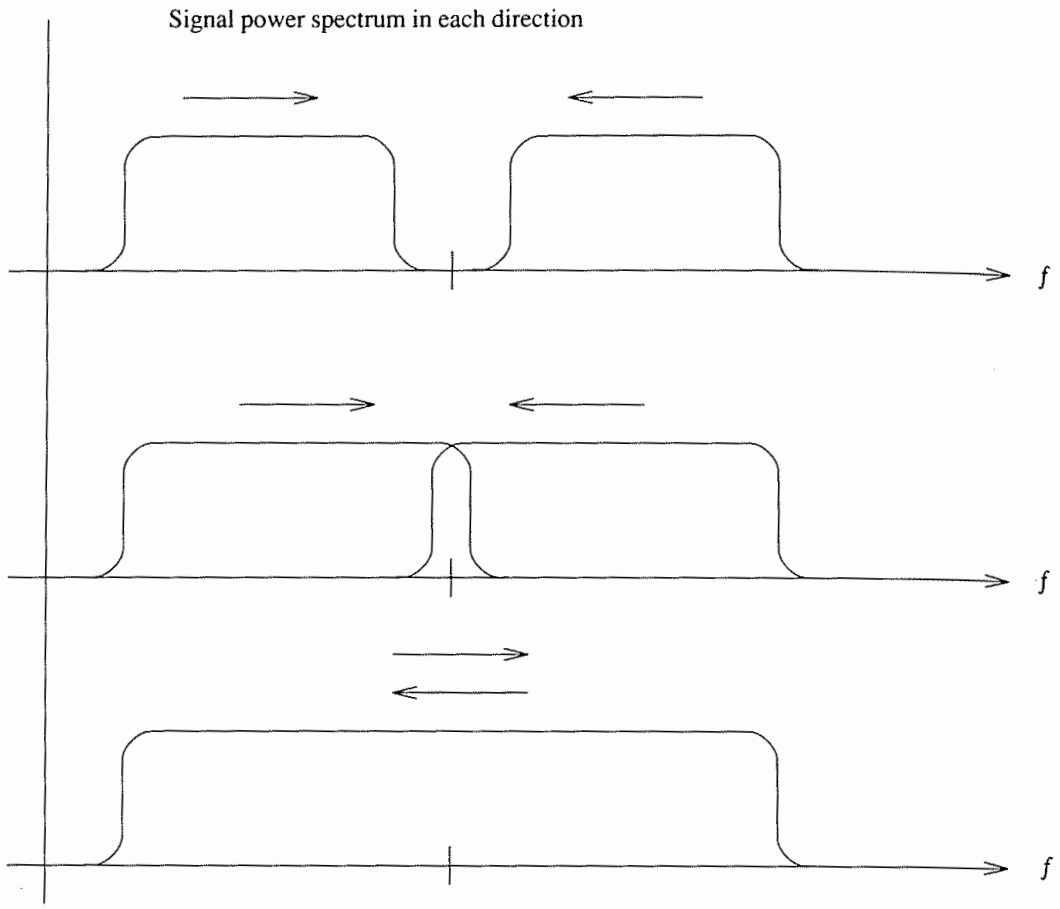


Fig. 9.1 Spectrum management for full-duplex communications: Bandsplitting, overlap, and same spectrum in both directions (used either simultaneously or alternately in time sharing).

desired signal arriving from the distant transmitter. A significant cost may be paid in waste of what is usually the least distorted part of the channel, at its center, and in degradations from band-edge distortions. Given a fixed total bandwidth and equal data rates in each direction as in the dialed telephone channel, the useful transmission rate in each direction is less than half of what the full

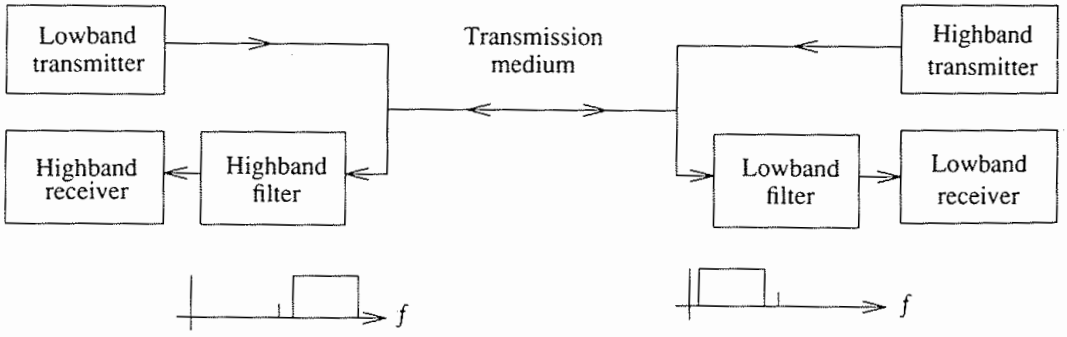


Fig. 9.2 Filtering to keep frequency bands from interfering in bandsplitting model.

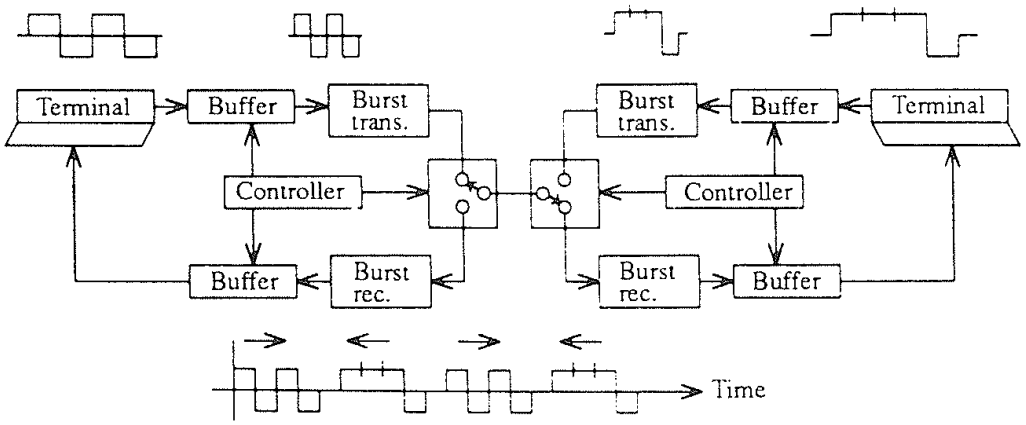


Fig. 9.3 Time-compression multiplexing (TCM), in which transmission is one way at a time in high-rate bursts.

channel would offer if it were used for one direction of transmission only. Bandsplitting is used in duplex voiceband modems up to 2.4 kbps and in two-way CATV systems (where the data rates in each direction are not equal).

Time compression multiplexing (TCM, Figure 9.3) is a time-domain sharing technique which is more or less equivalent to bandsplitting in its total bandwidth requirement, although each direction of transmission uses the same large spectrum. This technique alternates fast transmission bursts in each direction, saving up data (submitted to each transmitter at a lower rate) in buffers so that each of the end terminals has the illusion of a continuously available channel. If there were no propagation delay and hence no interval between bursts, the burst transmission rate required would be exactly twice the average data rate of each terminal and performance would be independent of buffer size (burst length). Realistically, there is propagation delay and the penalty must be paid in either buffer length (delay) or increased burst rate (bringing increased band-edge distortion), or in some reasonable combination (Exercise 9.1). TCM is practical for digital network access in most subscriber lines at rates up to about 80 kbps [1], but probably not for the 160 kbps required by ISDN.

For the spectrum overlap technique, filters suppress most of the self-interfering energy, and the remaining interfering energy, in the overlap region of the spectrum, is largely eliminated by the hybrid coupler/echo cancellation techniques explained below. Having partial rather than full overlap reduces the burden on the echo canceler, when bandwidth is available for this alternative.

When one is faced with the need to severely constrain bandwidth and achieve a high data rate, all with acceptable error rate, the best approach at the present time is the combination of hybrid coupler and echo canceler (Figure 9.4), which is the main subject of this chapter. The hybrid coupler, or "hybrid" for short, is a directional coupler realized as a passive 4-port [2]. Referring to Figure 9.4, if the impedances offered to the L and B ports are the same, and the impedances offered to the T and R ports are the same, then energy input to the T

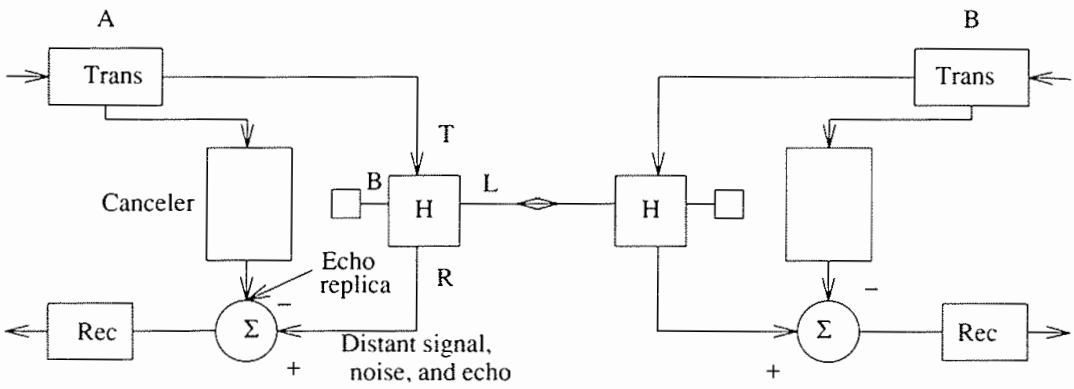


Fig. 9.4 Hybrid coupler and echo canceler for use of same spectrum in both directions.

port emerges from the L port (with 3 dB attenuation), and, ideally, none at all emerges from the R port. Similarly, energy arriving to the L port is totally directed (with 3 dB attenuation) to the R port. Thus the local transmitter does not interfere with the local receiver, and an ideal two-wire/four-wire interface is realized.

In practice, a compromise balancing impedance matches the typical two-wire line impedance. Also, the transmission circuit is not an ideal transmission line. Thus the reality is that considerable signal energy leaks from the T to the R port, and damaging "echoes" return from impedance mismatches at various points in the communication circuit. The hybrid coupler helps, but does not provide adequate separation of the transmitted and received signals.

Leakage through the hybrid coupler can be greatly reduced by isolating the line and balancing impedances (Figure 9.5) — a 50 dB echo suppression is possible [3] — but this presumes an ideal transmission line model for the two-wire line, with no returning echoes. There are also possibilities for adaptive hybrids that automatically generate a balancing impedance closely matched to the transmission line (to the extent that that line is an ideal transmission line).

In practice, it has been found more promising to use a compromise hybrid and depend on an echo canceler to attenuate local-transmitter energy that has

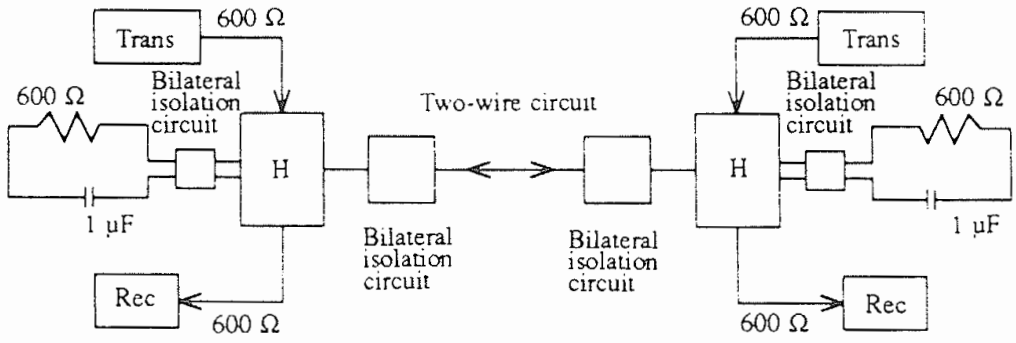


Fig. 9.5 Isolation of line and balancing impedances.

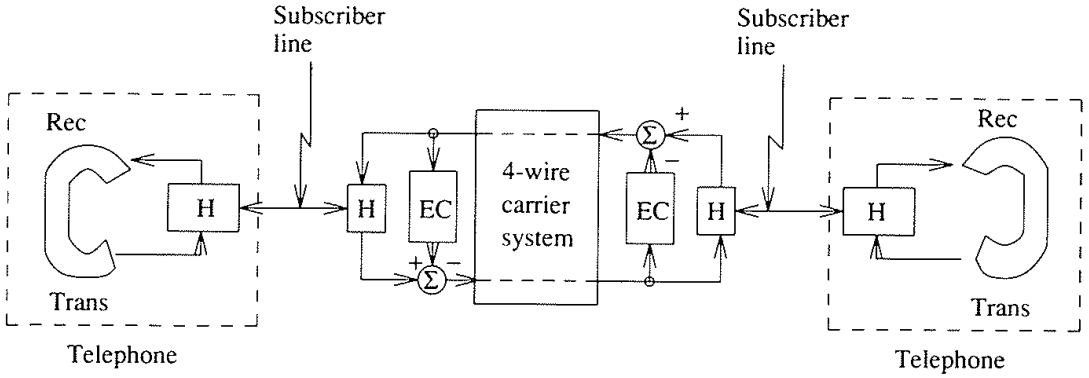


Fig. 9.6 Voice echo cancellation in the telephone network. Cancelers are provided only for "distant" echoes; nearer echoes are not disturbing to voice conversation.

leaked through the hybrid, to a level well below that of the desired distant signal. There is also a distant echo to be cancelled, as we shall see. Echo cancelers are also widely used within the telephone network (Figure 9.6) to attenuate distant echoes on voice connections with long delays, such as those including satellite links [4]. The original echo cancellation work was done in this voice echo cancellation context [5–7], and powerful integrated circuits were designed for deployment in the "tails" of long-distance telephone circuits [8]. But the differences between voice and data cancellation requirements, and in particular the fact that data may be transmitted in both directions simultaneously for long periods of time, unlike typical voice conversations, are such that voice echo cancelers in a long-distance telephone circuit (see Figure 9.6) are intentionally disabled for data communication sessions. A good overview is available in a paper by Messerschmitt [9].

A presumption is generally made here that the echo channel is linear and that the echo is free of timing, phase jitter, or frequency offset. In practice these impairments do occur, and later sections of this chapter suggest how these perturbations can be handled.

The bulk of this chapter describes structures and adaptation techniques for echo cancelers, seeking the best compromise among the generally conflicting goals of rapid adaptation, low residual cancellation error, and moderate implementation complexity. Some special considerations for operation of *digital* echo cancelers are described in Chapter 10. The next two sections of this chapter introduce system models for the two main applications of data echo cancellation in the telephone network, for full-duplex communication over dialed telephone circuits and via digital subscriber lines.

9.1 THE DIALED TELEPHONE CIRCUIT ECHO CANCELLATION MODEL

The application we are discussing is end-to-end full-duplex data communication over a dialed telephone circuit. Such connections are widely used for ter-

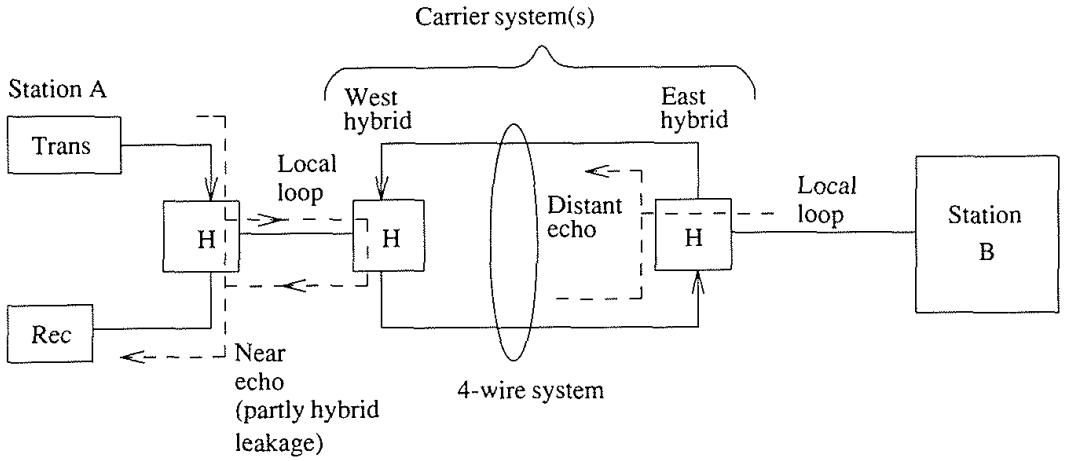


Fig. 9.7 Typical long-distance telephone circuit, showing the sources of most echoes returning to station A.

minal to host connections where digital data networks are unavailable, and as backups to four-wire private lines. They are also used for access to digital data networks when such access is not available at the nearest telephone office. We presume the standard 300–3300 Hz telephone channels described in Chapter 1.

As we shall see, the dialed telephone circuit typically consists of a concatenation of two-wire subscriber lines and four-wire circuits in carrier systems. In order to multiplex a number of voiceband circuits on a carrier system, the individual voiceband channels are, as described in Chapter 1, bandlimited to roughly 300–3300 Hz and either frequency-translated (FDM) or converted to digital format, usually with a nonlinear quantization characteristic, and time-division multiplexed (TDM). These operations introduce many of the limitations and distortions perturbing data communications in general and echo cancellation in particular. The model we develop here is applicable, with some minor changes, to other bandlimited channels, such as satellite circuits.

Figure 9.7 illustrates the origins of most echoes on dialed telephone channels. In general, a long-distance two-wire telephone connection, with a station at each end, will begin and end with two-wire facilities (i.e., the subscriber line) but will consist largely of four-wire carrier facilities. The two-wire–four-wire interfaces, in telephone offices, are provided by hybrid couplers with compromise balancing impedances. Depending on the characteristics of the two-wire transmission segment that the telephone switch has connected to the hybrid coupler, the hybrid may or may not be well balanced. Note that we are talking about a hybrid coupler inside the telephone network, using a compromise balancing impedance for the different two-wire subscriber lines it will be connected to. There may also be hybrid couplers within terminating station sets, as illustrated in Figure 9.7.

As a result of the mismatch, an echo returns through the carrier system, as illustrated by the "distant echo" dashed line in Figure 9.7. It returns with

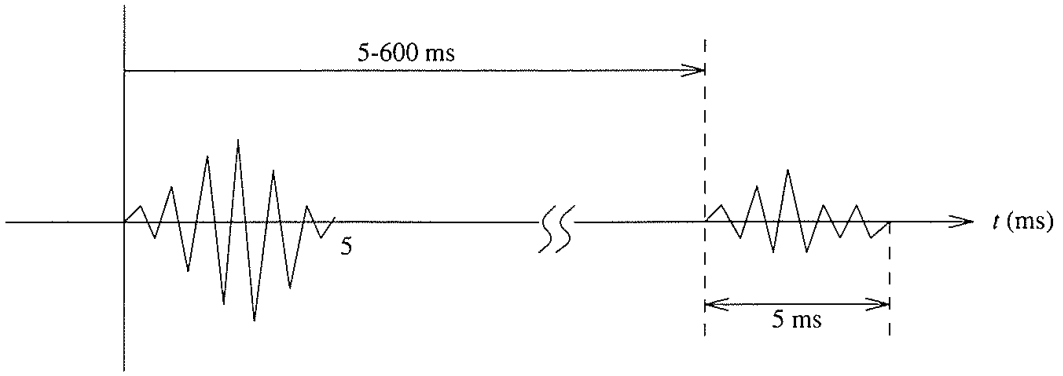


Fig. 9.8 Typical impulse response pattern of the baseband equivalent to a telephone echo channel. The echoes cluster in near and distant groups, each with a typical duration of perhaps 5 ms.

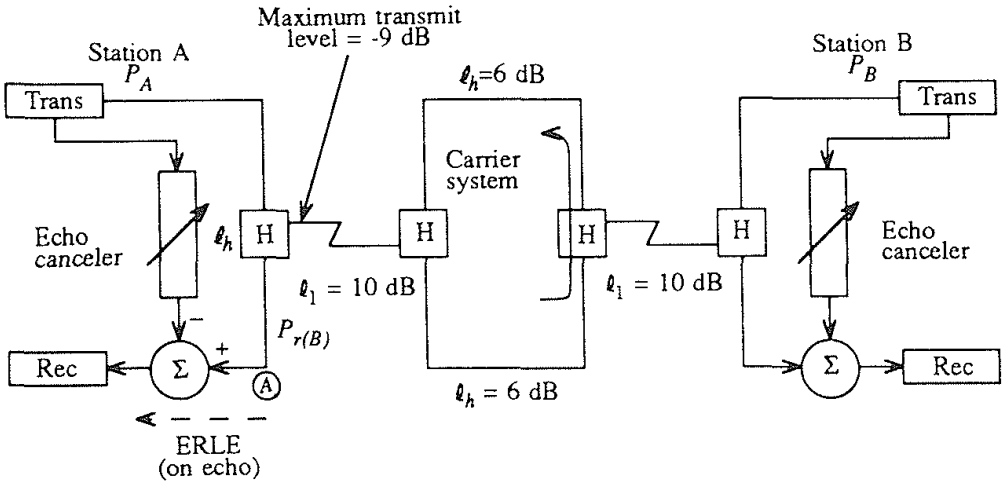
(mostly) linear distortion and at a delay ranging from 5 to 600 ms, depending on the electrical length of the circuit. The terrestrial fiber network has increased the maximum propagation delay in recent years. This round-trip signal is the "talker" distant echo. The distant echo returning to station A in Figure 9.7 results not only from leakage through the east hybrid coupler, but also from impedance mismatches all along the two-wire circuit between the east hybrid coupler and station B. It therefore exhibits dispersion, typically of the order of 5 ms (Figure 9.8).* It is possible for several distant echoes to return from several two-wire segments separated by four-wire transmission segments, but this is not usually the case for the public switched telephone network.

The distant echo may also return with phase perturbations contributed by the oscillators of FDM carrier systems. These are described, along with tracking methods for partially compensating them, in Section 9.5. Nonlinearities in the distant echo, contributed by companding devices (described in Chapter 1), cannot be canceled by the linear cancellation devices described in Section 9.3, but can be handled by the nonlinear "memory compensation" canceler defined in Section 9.4.

A near echo, composed of hybrid leakage at station A, reflections from impedance mismatches along the two-wire transmission path between station A and the two-wire–four-wire carrier system interface, and leakage through the west hybrid coupler, is also observed at the input to receiver A. This echo is ordinarily much larger, typically by 15 dB, than the distant echo, as suggested by Figure 9.8, although it, too, is dispersed over a relatively short interval. The separation between the near and distant echoes may, as already noted, be as large as 600 ms when there is a satellite in the four-wire system.

---

* It is important to note that the dispersion contributed by the transmitter filter, which constitutes part of the echo channel in the analysis of data driven echo cancelers, may contribute several additional milliseconds.



Definitions of terms:

ERLE	Echo return loss enhancement: the attenuation, in dB, of the echo, contributed by the echo canceler.
ℓ_h	Attenuation, in dB, of the hybrid leakage path from "transmit" to "receive" port. (Typical range equals 12 dB ± 4 dB.)
ℓ_1	Standard loss, in dB, of subscriber line. (10 dB would be close to the worst case.)
ℓ_2	Via net loss of carrier portion of transmission path.
P_A, P_B	Transmitter power, in dB.
P_{ea}	Power of echo, in dB. (-43 dBm is lowest level generally usable. Although unlikely to be that low on the public telephone network, it can happen on some private networks.)
$P_{r(B)}$	Power receiver from distant transmitter, in dB.
P_{near}	Power of near echo.
P_{dist}	Power of distant echo.

Fig. 9.9 Signal levels relevant to echo cancellation on a dialed telephone circuit. Noise is neglected. The pre-cancellation signal-to-echo ratio $SNR_{pre} = P_{r(b)} - P_{ea}$ is increased to $SNR_{post} = P_{r(B)} - (P_{ea} - ERLE)$ by the echo canceler.

The relative levels of received signal and total echo can vary widely. Nominally, the circuit losses can be roughly modeled as shown in Figure 9.9. The subscriber loops are often "padded out" to a standard loss ℓ_1 (dB) of around 10 dB, and the four-wire transmission plant has an intentional "via net loss" ℓ_2 of 4 dB on the average. The hybrid couplers both within the network and at station locations are presumed to have a leakage attenuation ℓ_h, between the incoming and outgoing lines of the four-wire circuit, in the range of 12 dB ± 4 dB. The attenuation is actually 3 dB greater, but in signal-to-echo calculations the 3 dB is canceled by the 3 dB loss in the normal signal path through the hybrid coupler, which is also neglected here.

With a local transmitted level P_A (in dB with respect to 1 mW into 600 ohms) and a distant-signal transmitted level of P_B, the received level of the desired signal at point A just before the canceler is (in dB)

$$P_{r(B)} = P_B - 2\,\ell_1 - 2\,\ell_2\,. \tag{9.1}$$

At the same point, the received level of the near echo signal is

$$P_{near} = P_A - \ell_h, \tag{9.2}$$

while the level of the distant echo signal is

$$P_{dist} = P_A - 2\,\ell_1 - 2\,\ell_2 - \ell_h\,. \tag{9.3}$$

From these formulas, the distant echo is seen to be roughly $2(\ell_1 + \ell_2)$ dB below the near echo, a number with a nominal range of 10 to 22 dB. The signal-to-near-echo and signal-to-distant-echo ratios are, respectively,

$$\begin{aligned} P_{r(B)} - P_{near} &= P_B - 2\,\ell_1 - \ell_2 - P_A + \ell_h \\ &= (P_B - P_A) - 2\,\ell_1 - \ell_2 + \ell_h \end{aligned} \tag{9.4a}$$

and

$$\begin{aligned} P_{r(B)} - P_{dist} &= P_B - 2\,\ell_1 - \ell_2 - P_A + 2\,\ell_1 + \ell_2 + \ell_h \\ &= (P_B - P_A) + \ell_2 + \ell_h. \end{aligned} \tag{9.4b}$$

The ranges of these ratios are, for $P_A = P_B = -4$ dBm and $\ell_1 = 10$ dB,

$$-18 < P_{r(B)} - P_{near} < -6, \qquad 10 < P_{r(B)} - P_{dist} < 22.$$

It is fortunate that it is the near echo rather than the distant echo which requires a great deal of attenuation, since the near echo is more stable and generally less dispersed than the distant echo.

The improvement which the canceler must contribute is called the echo return loss enhancement (ERLE). For a 9600-bps QAM modem requiring a 25 dB SNR for an error rate below 10^{-7}, the required ERLE for the near echo ranges from 31 to 41 dB, and for the distant echo from 7 to 15 dB. On some dialed lines the maximum required ERLEs are actually more than these numbers suggest — perhaps more than 50 dB on the near echo — because of noise and excessive transmission loss. As a severe example, if $P_A = 0$ dBm, $\ell_h = 6$ dB, and $P_{r(B)} = 43$ dBm, then a total ERLE of 62 dB is required to achieve a 25 dB SNR.

In order for the echo canceler to work most effectively, the transmitted signals from the two ends of the circuit should be uncorrelated. This will become clear from later analysis, where the distant data signal, if present, will be regarded by the local adapting echo canceler as a form of interfering noise. To make the signals uncorrelated, data scramblers, with different scrambling patterns, must be provided between terminal and modem at each end.

Furthermore, to use data echo cancelers at station locations, echo suppressors and echo cancelers in the network have to be disabled, as has already been suggested. The echo suppressor, a network element in use since the late 1920s now largely replaced by the echo canceler, allows only one direction of transmission at a time and is obviously incompatible with full-duplex transmission. Network echo cancelers are incompatible to a lesser degree, going out of adjustment with certain spectrally degenerate data signals and with some signal power level scenarios in full-duplex transmission. These elements can be (temporarily) disabled with appropriate in-band signals.

9.2 THE ECHO CANCELLATION MODEL FOR DIGITAL SUBSCRIBER LINES

The copper twisted-pair lines used for subscriber access to telephone service at a telephone central office are modified for access to a digital network (Figure 9.10) by installing echo-canceling transceivers at the two ends, just as the two-way transmission on a telephone line. The major differences between this circuit and the dialed telephone line are (as described in Chapter 1):

(1) The digital subscriber line is of limited distance, typically under 6000 meters.

(2) It is baseband (it passes low-frequency components).

(3) Transmission is at rates of 160 kbps full duplex and above, revealing entirely different channel characteristics and introducing new problems such as crosstalk.

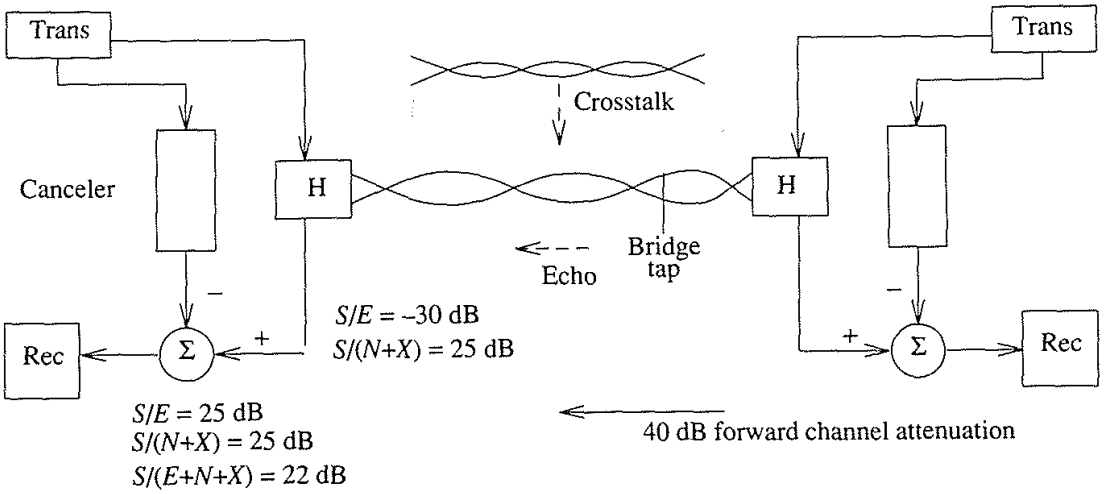


Fig. 9.10 Echo cancellation for full-duplex data transmission on a digital subscriber line. Crosstalk interference comes from other twisted pairs also carrying wideband data signals. Nominal ratios of signal to echo, signal to noise+crosstalk, and signal to echo+noise+crosstalk are shown for a channel requiring an echo cancellation (ERLE) of about 55 dB to achieve a final signal to "noise" ratio of about 22 dB.

(4) Transmission impairments characteristic of passband circuits in carrier systems (voice-band filtering, phase jitter, frequency offset) are absent.

The four impairments that must be traded off against one another are impulse noise, channel distortion, echo, and (mostly near-end) crosstalk (thermal noise is not an issue, usually). Echo comes from leakage of local transmitted signal through the hybrid to the local receiver, and from bridge taps and other impedance discontinuities in the subscriber line. Because of the limited distance, there is no distant echo; the echo is constrained to a total dispersion of perhaps a couple of milliseconds. The ratio of desired received signal to echo may be as low as –30 dB for a subscriber loop channel with 40 dB attenuation.

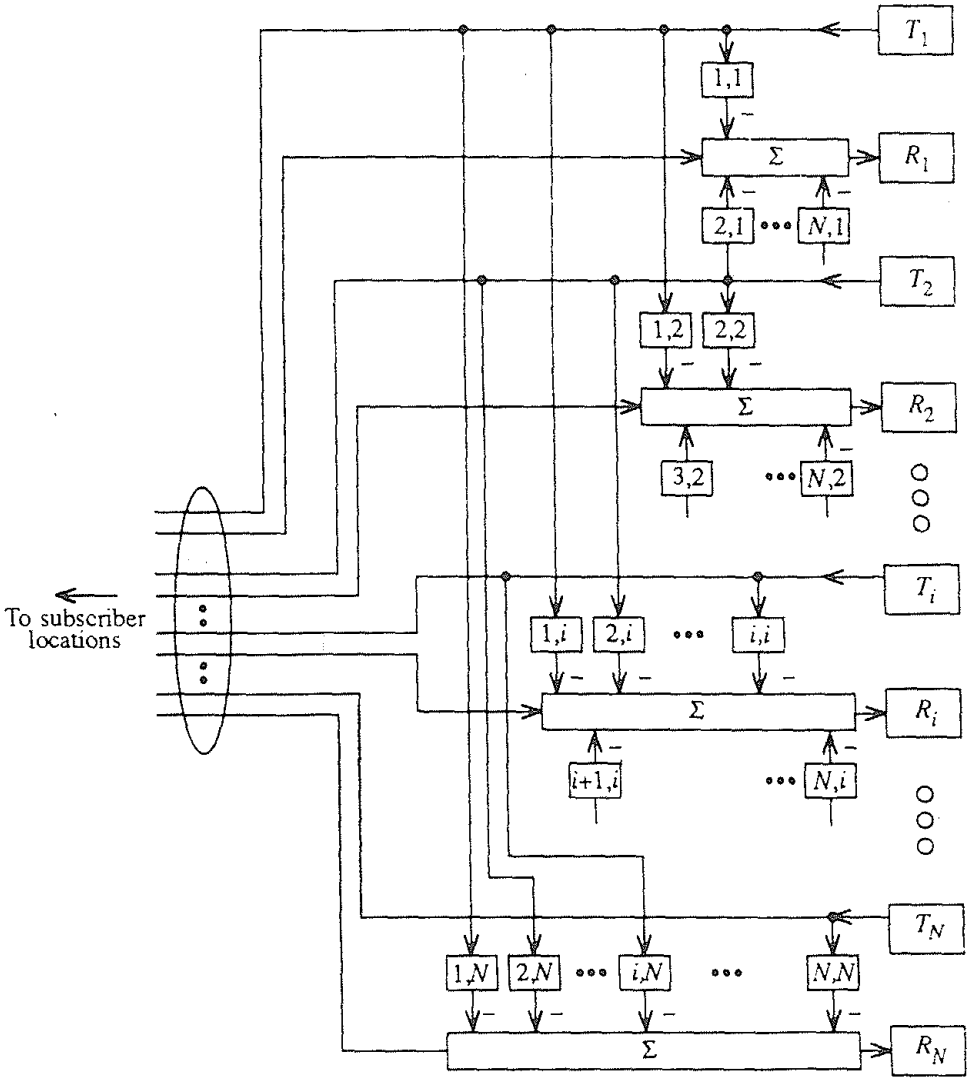


Fig. 9.11 Adaptive cancellation of near-end crosstalk (NEXT) at a serving location (e.g., a central office) for digital subscriber lines. The notation (i,j) within a box designates an adaptive filter replicating the crosstalk from transmitter i to receiver j.

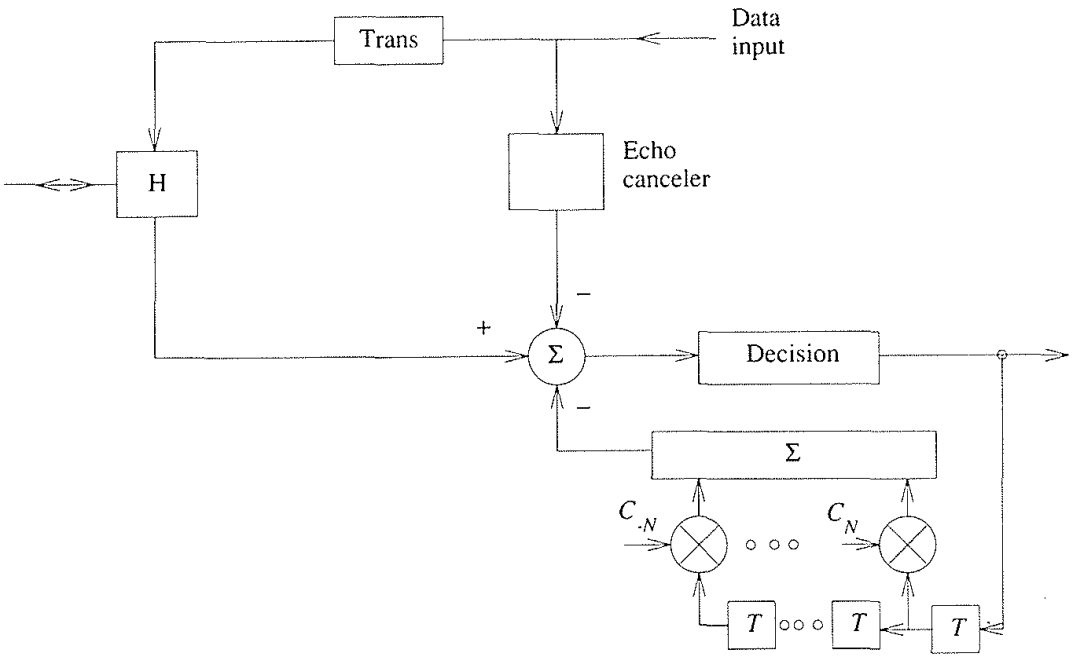


Fig. 9.12 The ISDN-U interface transceiver combining echo cancellation with decision-feedback equalization.

In order to realize a 25 dB signal-to-echo ratio, allowing a few dB margin for Gaussian noise and crosstalk, an ERLE of 55 dB is required [12].

Adaptive cancellation structures can, to a limited extent, be used to attenuate near-end crosstalk (NEXT) interference [11], an impairment characterized in Chapters 1 and 4. Consider the model of Figure 9.11 for the central office from which a bundle of digital subscriber lines (twisted pairs) radiate out to subscribers. The near-end crosstalk observed by one of the digital receivers comes from the other locally transmitted signals, which presumably are accessible and can be used to drive adaptive cancelers. Figure 9.11 shows a set of cancelers, one for each local transmitter contributing to the crosstalk interference, helping to clean up the signal supplied to the particular receiver. For N active line interfaces, N^2 adaptive cancelers would be required. Adaptation of this set of N^2 cancelers can be described as one large system adaptation [12]. There is also the possibility of NEXT cancellation at those subscriber locations, such as business offices, where subscriber terminations are clustered. In practice, it has proven difficult to access and utilize the locally transmitted signals to drive the N^2 cancelers.

A combination of techniques is required in the digital transceiver to bring all the distortions down to tolerable levels. For the ISDN Basic rate interface operating at 160 kbps, the preferred technique is to combine echo cancellation with decision-feedback equalization, as shown in Figure 9.12. As described in Chapter 7, the decision-feedback equalizer (DFE) can eliminate postcursor intersymbol interference but not precursor intersymbol interference. In the digital

subscriber loop, the intersymbol interference tends to be mostly postcursor due to the ringing of bridge taps and the loss of energy near DC in the hybrids [11]. A DFE of reasonable length, such as 40 symbol intervals, is ordinarily adequate.*

9.3 FIR (TAPPED DELAY LINE) CANCELER STRUCTURES

Just as for the channel equalizers described in the last two chapters, the tapped delay line offers an easily described and analyzed structure for realizing an echo canceler. For echo cancellation, the tapped delay line realizes a replica, within the constraint of its finite length, of the echo, which is subtracted from the arriving echo. But other structures, with better performance under certain conditions, are also possible, including lattice filters and lookup tables (memory compensation). These are introduced in Section 9.4. We will also show, in the present section, that fractionally spaced tapped delay line echo cancelers offer the same advantages as fractionally spaced channel equalizers, in particular an output that can be interpolated to any sampling epoch, freeing the system from complex synchronization requirements. One important difference between equalization (which is basically channel inversion) and echo cancellation (which is system identification) is that for the latter system the number of tap weights required can be determined by inspection of the echo channel. As described in Chapter 8, it is not a straightforward operation to determine the required number of equalizer taps.

Assume, for the moment, a real baseband model, shown in Figure 9.13, for the two-way data communication system. We will later generalize to a passband model, which helps the discussion of echo cancellation for the dialed telephone channel, but the important principles are easiest to show in the real baseband mode. We assume, for the present discussion, that the echo channel is a linear filter $h_e(t)$; that is, the received signal at the input to the canceler of Figure 9.13 is

$$r_A(t) = s_B(t) + s_A(t) * h_e(t) + n(t), \tag{9.5}$$

where $s_A(t)$ is the local transmitted signal, $s_B(t)$ is the desired distant transmitted signal, and $n(t)$ is a combination of noise and crosstalk interference. The echo is the convolution of $s_A(t)$ with the echo channel. The canceler output is an echo replica, $q(t)$, which is subtracted from $r_A(t)$ to produce the receiver input.

9.3.1 VOICE-TYPE ECHO CANCELER

Figure 9.13 also illustrates the tapped delay line echo canceler, consisting of a transversal (FIR) filter with $2N+1$ taps, spaced at T' sec intervals, where

* For the 2B1Q (4-level PAM) line signaling described in Chapter 4, the baud for ISDN Basic access is 80,000 symbols/sec. Forty symbol-interval taps corresponds to a maximum dispersion of 0.5 ms. This may not correspond to all of the postcursor dispersion, but does include the great bulk of postcursor energy.

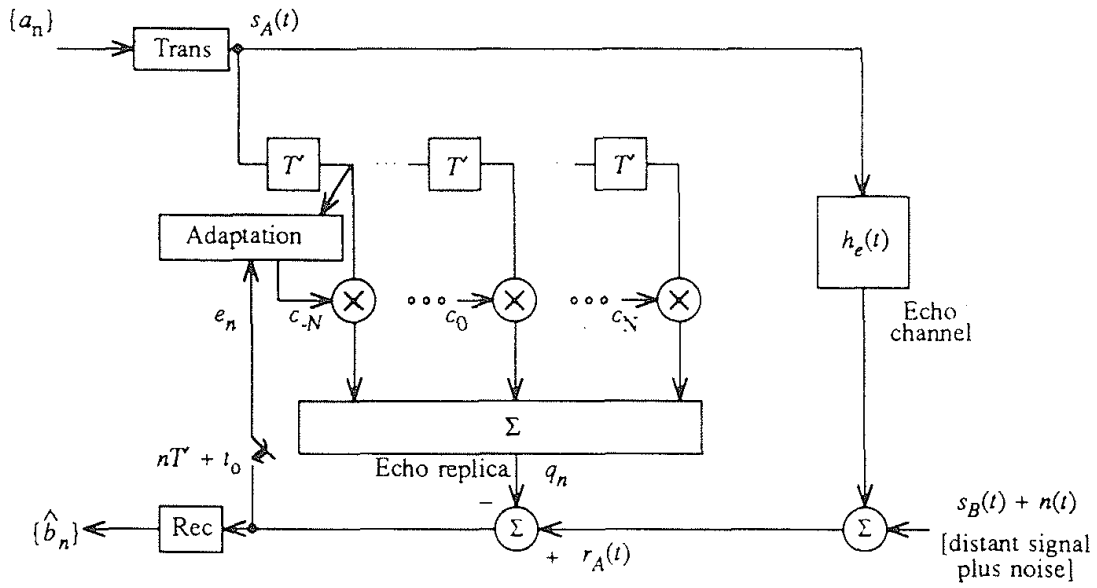


Fig. 9.13 Model for a "voice signal type" echo cancellation system. Adaptation is shown for one tap only.

$1/2T'$ is greater than the highest significant frequency in the received signal $r_A(t)$. The reader will recognize the maximum allowed value of T' as the Nyquist interval. The length $2NT'$ of the tapped delay line spans the maximum effective dispersion of the echo channel. The canceler as shown here is driven by the local transmitted line signal $s_A(t)$, and its tap weights, $c_{-N},\ldots,c_0,\ldots,c_N$ are adapted by a correlation between the tap voltages (at time nT') and the cancellation error e_n, as will be described.

The canceler of Figure 9.13 is a voice-type structure, identical to that widely used in the voice echo canceler of Figure 9.6. Although an echo canceler for data communication can be built this way [13], it has limitations which can be overcome by slightly different structures tailored to the data application. The major limitation is implementation complexity, requiring $2(2N+1)/T'$ high-resolution multiplications per second for filtering and tap updating in a digital implementation. Other problems concern convergence rate, stability, and residual error. *Data-driven* structures mitigating these problems are described in subsequent sections.

If the transversal filter's total length of $2NT'$ sec is greater than the dispersion of the echo channel, then any echo can theoretically be exactly replicated, and canceled. The fact that any echo can be replicated is seen in the transfer function

$$C_\infty(f) = \sum_{-\infty}^{\infty} c_m \, e^{-j2\pi mfT'} \tag{9.6}$$

of an infinite-tap transversal filter, which is a Fourier series which can generate an exact copy of any echo channel transfer function bandlimited to $1/2T'$ Hz.

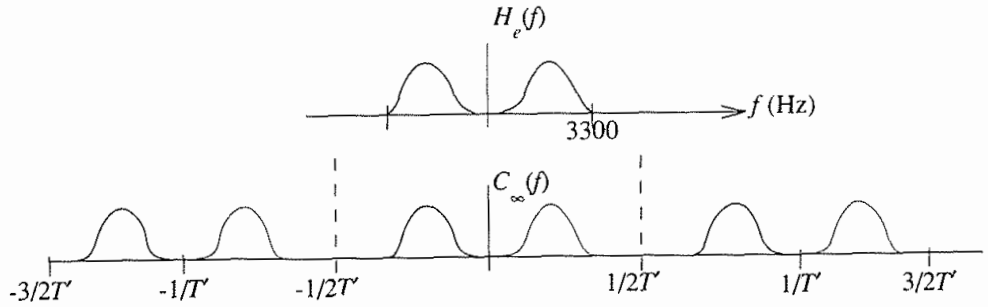


Fig. 9.14 Echo channel transfer function $H_e(f)$, and periodic transfer function generated by an ideal infinite-length transversal filter $C_\infty(f)$ with tap spacing T'.

As Figure 9.14 shows, the transfer function $C_\infty(f)$ is periodic with period T'. No aliasing distortion will result if there is no signal, echo, or noise energy at frequencies above $1/2T'$ Hz. The transversal filter limited to $2N+1$ taps, with transfer function

$$C_N(f) = \sum_{m=-N}^{N} c_m \, e^{-j2\pi mfT'}, \tag{9.7}$$

is all that is needed to generate any echo channel effectively bandlimited to $1/2T'$ Hz and with a dispersion effectively less than $2NT'$ sec. [Strictly speaking, $h_e(t)$ cannot be simultaneously bandlimited and time limited.] As mentioned above, the dispersion of an echo channel can be easily measured, allowing one to specify an appropriate number of taps. This is very different from the problem faced by a designer of a channel equalizer, where the delay spread determining the length of the transversal filter is not obvious from the transmission channel dispersion; it depends on the inverse of the channel characteristic in a complicated way, as derived in Chapter 8.

The active length of the echo canceler need not exceed the actual dispersion introduced by the echo channel. For the dialed telephone circuit, the distant echo may be delayed by as much as 600 ms, although the dispersion of the distant echo may be only a few milliseconds. Rather than make the tapped delay line 600 ms long, which could imply thousands of taps, the active taps can be provided in two shorter sections separated by a bulk delay, as shown in Figure 9.15. This matches the echo response suggested in Figure 9.8. The bulk delay must be determined by some separate means, such as an initial channel sounding (described below).

The voice-type echo canceler generates outputs not only at intervals T', but at any time instant, using the interpolation specified by the sampling theorem (Chapter 4). One could, in fact, construct a digital version of the canceler clocked to the sampling epoch τ of the local receiver A, so that clean symbol-interval samples at the canceler output would be available at the optimum times for sampling the desired distant signal $s_B(t)$. A slight mismatch in symbol rates between the transmitters at the two ends would be compensated by a continual

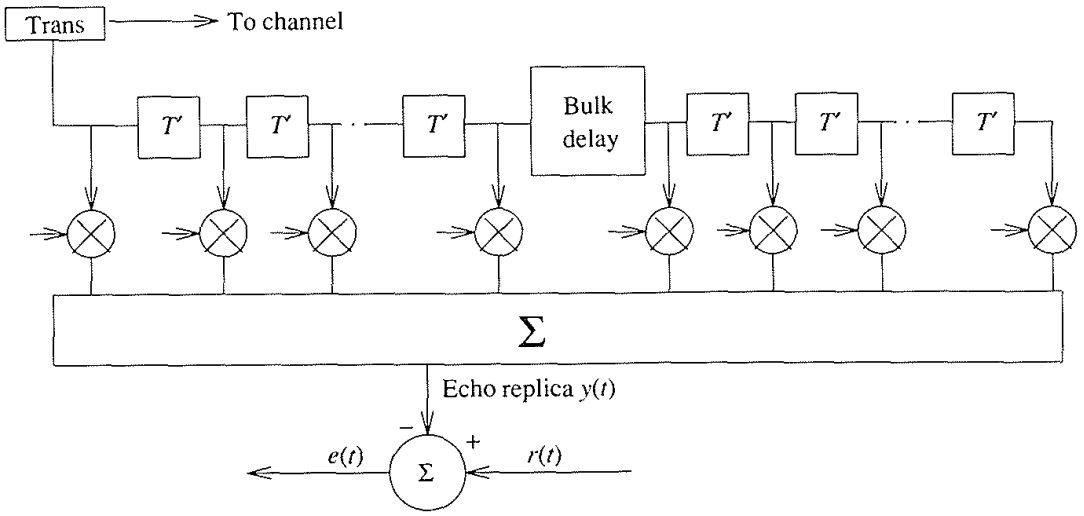


Fig. 9.15 Active echo canceler sections separated by a bulk delay.

tracking adaptation in the echo canceler. This synchronization mismatch is more difficult to handle in the data-driven cancelers described in subsequent sections.

By simple Fourier analysis (Exercise 9.2), the tap weights of the ideal echo canceler are the samples, at intervals T', of the impulse response $h_e(t)$ of the echo channel. One can easily show (Exercise 9.7) that this choice of tap weights minimizes the mean-square error at the canceler output provided the local and distant data trains are uncorrelated.

We have already noted that the distant signal $s_B(t)$ adds to the interfering noise at end A. This is equally true for the data-driven structures described in the rest of this chapter. For initial start-up, when the tap weights of the echo canceler may be far from their optimum settings, this "noise" can make convergence extremely slow. The common solution, which is specified, for example, in the V.32 standard for full-duplex modems on dialed telephone lines, is to initially adapt one end at a time. That is, transmitter A sends a training sequence and transmitter B is silent so that echo canceler A can adapt, and then transmitter B sends a training sequence and transmitter A is silent so that echo canceler B can adapt. The situation when both transmitters are simultaneously active (as in normal full-duplex data communication) is known as *double-talking*. Tracking of changes in the echo channels (one for each direction) is very slow during double talking, as implied above, but this is not a serious problem if the echo channels change very slowly and the cancelers are already adapted.

An alternative technique for startup is to measure the echo channel impulse response directly and "jam-set" the canceler tap weights as samples, at intervals T', of the measured impulse response. The impulse response of a fixed linear filter (which the echo channel is to a first approximation) is the expected value of the cross-correlation of a white-noise input with the filter output. Mathematically, if $n(t)$ is the white noise input with two-sided spectral power density equal

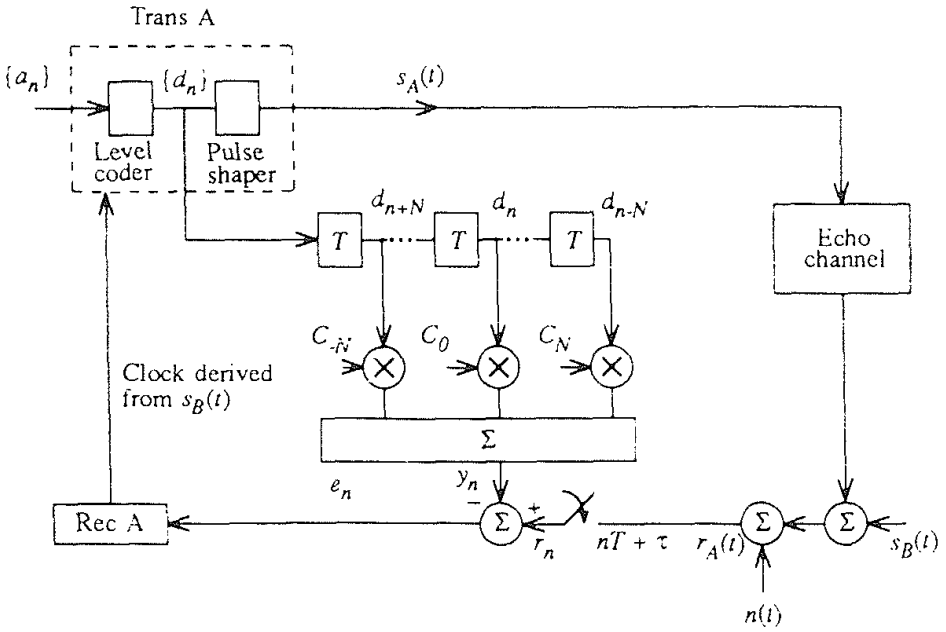


Fig. 9.16 Symbol-interval data-driven echo canceler. Adaptation, driven by top voltages and e_n, is not shown.

to 1 watt/Hz, $h_e(t)$ is the unknown echo channel impulse response, and $y(t)$ is the echo channel output, then the expected value of the cross-correlation of $n(t)$ with $y(t)$ is

$$E[n(t)\ y(t+\tau)] = \int_{-\infty}^{t+\tau} n(t)n(\beta)h_e(t+\tau-\beta)d\beta = h_e(\tau). \quad (9.8)$$

The expectation can be estimated by averaging a series of cross-correlations, and the initial tap values are simply these estimates at $\tau = mT'$, $-N \le m \le N$. An equivalent procedure is possible in a data-driven echo canceler [17].

In practice, a channel is sometimes used intermittently. In these cases, the tap weights from a previous period of activity can be saved and used as the initial settings for a new period of activity.

9.3.2 SYMBOL-INTERVAL DATA-DRIVEN ECHO CANCELER

We now begin consideration of *data-driven* echo cancelers [14], where the driving signal is the data stream at the input to the local transmitter rather than the line signal at its output. Our first model is the *symbol-interval data-driven canceler* of Figure 9.16. Here the transversal filter develops echo channel estimates at symbol intervals T, rather than the finer Nyquist intervals T', implying considerably fewer taps to span a given dispersion. For example, if the echo dispersion is 5 ms, a symbol-interval canceler with tap spacing $T = 1/2400$ sec requires 12 taps, while a canceler with Nyquist-interval taps of $T' = 125\ \mu s$

requires 40 taps. This substantially reduces the number of multiplications for tap updating. Also, each multiplication is much simpler, since the data value in the tapped delay line only takes on a few discrete values, e.g., four for four-level PAM.

Except for these important factors, the model remains the same as that of Figure 9.13. A linear echo channel $h_e(t)$ is estimated by the FIR filter, and the received signal $r_A(t)$ consists of the sum of echo, noise, and desired distant data signal.

It seems reasonable to build a canceler to cancel echoes at symbol intervals, since decisions are made by the local receiver only at symbol intervals, but this approach brings with it the difficult synchronization problems alluded to in the previous section. The times at which corrected samples must be generated for receiver A are the optimal sampling times for the data train modulating the received distant signal $s_B(t)$. The echo canceler serving receiver A is, on the other hand, driven by the data train modulating the local transmitted signal $s_A(t)$, and its outputs are necessarily synchronized with that data train, unlike the voice-type canceler, which has more flexibility. When the transmitters at the two ends of a dialed connection run on separate, unsynchronized clocks, as is normally the case, it is impossible to synchronize a symbol-interval canceler clocked by transmitter A with the local receiver clock derived from $s_B(t)$.

Even when transmitting at nominally the same rate, there is a drift between the data trains that requires special buffering and slippage correction. In contrast, the transmitters at the two ends of a subscriber line used for digital access (Section 9.2) are synchronized by a master clock arrangement, allowing symbol-interval cancellation, although even here there may be implementation conveniences from having canceler outputs at Nyquist intervals. Transmitter B provides the master clock, so that timing of transmitter A is controlled by the timing signal that receiver A derives from the data stream arriving from transmitter B.

Synchronization, when possible, offers a potentially large advantage in adaptive reference echo cancellation [21], described in Section 9.3.5. In this scheme, subtraction of an estimate of the distant data signal largely eliminates that signal as a component of the "noise" interfering with adaptation of the local echo canceler, leading to faster convergence and/or smaller residual cancellation errors.

9.3.3 GRADIENT ADAPTATION ALGORITHM

The "known gradient" (or steepest descent) algorithm, introduced in Chapter 8, gives insight into the behavior of a tapped delay line echo canceler. The stochastic least mean squares (LMS) adaptation algorithm is also applicable to echo cancelers. In this section, the known gradient algorithm for adaptation of the canceler tap weights is derived and analyzed. The algorithm could also be described for the voice-type canceler of Figure 9.13.

Referring to Figure 9.16, at time nT, define the tap-weight set as the column vector

$$\mathbf{c}_n = \left[c_{-N}(n),\ldots, c_N(n)\right]' , \tag{9.9}$$

and the tap voltage vector as

$$\mathbf{d}_n = \left[d_{n+N},\ldots, d_{n-N}\right]' . \tag{9.10}$$

The constant and (for convenience) noncausal echo channel impulse response is

$$\mathbf{h}_e = \left[h_e(-NT + \tau),\ldots, h_e(NT + \tau)\right]' , \tag{9.11}$$

where τ is whatever sampling epoch is chosen by the receiver. Note that for different sampling epochs, there is effectively a different echo channel as far as the echo canceler is concerned. Note also that the transmitter pulse shaping is now included in $h_e(t)$. The sample at time nT of the received signal is

$$r_n = \mathbf{d}_n'\mathbf{h}_e + \nu_n , \tag{9.12}$$

where ν_n is the desired distant signal plus Gaussian noise at time nT (and is noise in its entirety from the point of view of canceling the echo signal). In (9.12) and throughout this book, the prime denotes transpose, so that $\mathbf{d}_n'$ is a row vector.

Whatever technique is used to specify the initial tap weight vector $\mathbf{c}_0$, we will try to adapt the tap weights to arrive at an eventual value $\mathbf{c}_{opt}$ which minimizes the mean-square error

$$MSE = E\, e_n^2 , \tag{9.13}$$

where the cancellation error at time nT is (from Figure 9.16)

$$e_n = \mathbf{r}_n - \mathbf{c}_n'\,\mathbf{d}_n = (\mathbf{h}_e - \mathbf{c}_n)'\,\mathbf{d}_n + \nu_n . \tag{9.14}$$

From (9.14) we have

$$\begin{aligned} MSE &= E\left\{(\mathbf{c}_n - \mathbf{h}_e)'\,\mathbf{d}_n\mathbf{d}'\,(\mathbf{c}_n - \mathbf{h}_e) - 2\nu_n\,(\mathbf{c}_n - \mathbf{h}_e)'\,\mathbf{d}_n + \nu_n^2\right\} \\ &= (\mathbf{c}_n - \mathbf{h}_e)'\,R(\mathbf{c}_n - \mathbf{h}_e) - 2(\mathbf{c}_n - \mathbf{h}_e)'\,\mathbf{p} + \sigma_n^2 , \end{aligned} \tag{9.15}$$

where

$$R = E\,[\mathbf{d}_n\mathbf{d}_n'] \tag{9.16}$$

is the reference signal covariance matrix. Note (Exercise 9.5) that if the data input levels to the tapped delay line are zero mean and uncorrelated, R is, except for a scale factor, the identity matrix. We shall, however, allow the possibility of a correlated sequence in the analysis below. That makes this analysis also applicable to the voice-type echo canceler, driven by line signal samples.

We further define the noise-data correlation vector,

$$\mathbf{p} = E\nu_n \, \mathbf{d}_n \ , \tag{9.17}$$

and the noise variance $\sigma^2 = E\ \nu_n^2$. From Chapter 8, recall that the quadratic form (9.15) is minimized by selecting tap weights

$$\mathbf{c}_{opt} = \mathbf{h}_e + R^{-1}\mathbf{p} \ . \tag{9.18}$$

If the "noise" ν_n (including the distant desired signal) is uncorrelated with the signal giving rise to the echo, i.e., $\mathbf{p} = 0$, the optimum tap weight setting is the echo channel impulse response, as has already been suggested.

As with the adaptive equalization discussion in Chapter 8, the known gradient algorithm adjusts the tap weight vector in the direction of the negative gradient of the MSE. We begin with

$$MSE = E\ e_n^2 = E\left[r_n - \sum_{m=-N}^{N} c_m \ d_{n-m}\right]^2 \ , \tag{9.19}$$

where $c_m = c_m(n)$ and d_{n-m} are understood as elements of $\mathbf{c}_n$ and $\mathbf{d}_n$, and r_n is the sample at time nT of the received signal (including echo) at the subtractor input. The MSE can be shown to be a convex function of the tap weights just as was shown in Chapter 8 for channel equalization.

The partial derivative of MSE with respect to c_m is $-2\ E(e_n d_{n-m})$, so the gradient with respect to $\mathbf{c}_n$ is $-2\ E(e_n \mathbf{d}_n)$ and the adaptation algorithm, moving in the negative gradient direction, is

$$\mathbf{c}_{n+1} = \mathbf{c}_n + \beta\ E(e_n \mathbf{d}_n) \ , \tag{9.20}$$

where β is the step size, which is arbitrary at this point.

To examine convergence, define the tap-weight error vector as

$$\xi(n+1) = \mathbf{c}_{n+1} - \mathbf{c}_{opt} \ .$$

Then the expected tap-weight error vector is

$$\mathbf{E}_{n+1} = E(\xi_{n+1}) = E\{\mathbf{c}_n + \beta\ e_n \mathbf{d}_n - \mathbf{c}_{opt}\} = [I - \beta\ R]\ \mathbf{E}_n \ , \tag{9.21}$$

where we have assumed that the tap voltage vector is uncorrelated with the tap weight vector. Thus

$$\mathbf{E}_n = [I - \beta\ R]^n\ \mathbf{E}_0 \ , \tag{9.22}$$

and diagonalizing R,

$$R = \mathbf{P}\ \Lambda\ \mathbf{P}' \ , \tag{9.23}$$

where $\mathbf{P}$ is the (orthonormal) matrix of (column) eigenvectors of R and Λ is the (diagonal) eigenvalue matrix, we have (Exercise 9.6)

$$(I - \beta\ R)^n = \mathbf{P}(I - \beta\Lambda)^n \mathbf{P}' \ . \tag{9.24}$$

Because R is symmetric and positive definite, it is invertible and has positive real eigenvalues. Its inverse R^{-1} is also symmetric. Placing (9.24) into (9.22) gives an iterative equation for the tap error:

$$\mathbf{E}_n = \mathbf{P}(1 - \beta\Lambda)^n \mathbf{P}' \mathbf{E} . \tag{9.25}$$

For $\mathbf{E}_n$ to converge to zero, each element must converge to zero. The fastest converging (and least stable) element corresponds to λ_{max}; all elements will converge if this one does, i.e., if

$$\left|1 - \beta\lambda_{max}\right| < 1 ,$$

or

$$0 < \beta < 2/\lambda_{max} . \tag{9.26}$$

For $\beta = 1/\lambda_{max}$, an appropriate value for rapid convergence, the slowest-converging element of $\mathbf{E}_n$ is the factor

$$\left[1 - \beta\lambda_{min}\right]^n = \left[1 - \frac{\lambda_{min}}{\lambda_{max}}\right]^n . \tag{9.27}$$

The results (9.26) and (9.27) put a cap on the step size and show how the rate of convergence is decreased with increasing eigenvalue spread.

EXAMPLE 1 Convergence of ideal gradient algorithm for uncorrelated and correlated driving signals

1. Uncorrelated driving signal samples $\{d_m\}$, and a 2-tap canceler. Assume that the $\{d_m\}$ are uncorrelated random variables. Then if each element has variance A,

$$R = E\mathbf{d}_n\mathbf{d}_n' = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}, \qquad \Lambda = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} .$$

   Let $\beta = 1/\lambda_{max} = 1/A$. Then $I - \beta\Lambda = 0$, implying $\mathbf{E}_1 = 0$, independent of the starting value $\mathbf{c}_0$ of the tap-weight vector. Convergence, in this degenerate example, occurs in one iteration! This will not be the case for the stochastic adaptation algorithm, even when, as is frequently the case, the elements of $\mathbf{d}_n$ are uncorrelated and R is a diagonal matrix, with unity eigenvalue ratio.

2. Correlated signal samples $\{d_n\}$. Assume $\mathbf{h}_e' = [1, -1]$. Then $\mathbf{c}_{opt}' = [1 \ -1]$.

Assume further that the correlation of the signal samples $\{d_n\}$ is such that

$$R = E\mathbf{d}_n\mathbf{d}_n' = A\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

The eigenvalues of R are found from the determinant equation

$$0 = |R - \lambda I| = \begin{vmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + 3.$$

Thus the eigenvalue and eigenvector matrices are

$$\Lambda = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

Let $\beta = 1/\lambda_{\max} = 1/3$. Then

$$I - \beta\Lambda = \begin{bmatrix} 1 - \beta\lambda_1 & 0 \\ 0 & 1 - \beta\lambda_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1/3 \end{bmatrix}$$

and

$$P(1 - \beta\Lambda)^n P' = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = (1/3)^n \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Thus equation (9.25) implies that both elements of the tap-weight errors vector decrease as $(1/3)^n$.

For the idealized model we have been considering, in which the gradient of the mean-square error is assumed known, the expected tap-weight error converges to zero, as (9.25) indicates. We will see in the next section that for stochastic adaptation, which is a practical approximation to the gradient algorithm, a residual error remains for any finite step size.

9.3.4 STOCHASTIC (LMS) ADAPTATION ALGORITHM

When ensemble expectations are unknown, time averages must suffice. That is the essential motivation for the stochastic LMS adaptation algorithm here and in Chapter 8. This algorithm adapts in the general direction of the negative gradient of the mean-square error through a long series of small random movements whose time-averaged value is in the correct direction.

Explicitly, the least mean-square (LMS) algorithm utilizes the gradient of the *instantaneous* squared error, rather than the mean-square error. This gradient is

$$\partial e_n^2/\partial \mathbf{c}(n) = e_n[\partial e_n/\partial \mathbf{c}(n)] = -2e(n)\mathbf{d}(n) , \tag{9.28}$$

so that the stochastic adaptation algorithm is

$$\mathbf{c}(n+1) = \mathbf{c}(n) + \beta e(n)\ \mathbf{d}(n) . \tag{9.29}$$

The properties of (9.29) can be evaluated through examination of either the mean-square tap-weight error or the mean-square cancellation error at the nth iteration. We choose the latter, but the analysis begins with definition of the tap-weight error vector. This is, assuming the local data signal is uncorrelated with the distant data signal and the additive noise so that the ideal echo canceler tap weights are the echo channel samples,

$$\xi(n) = \mathbf{c}(n) - \mathbf{c}_{opt} = \mathbf{c}(n) - \mathbf{h}_e . \tag{9.30}$$

From (9.29) and (9.30)

$$\xi(n+1) = \xi(n) + \beta e(n)\mathbf{d}_n . \tag{9.31}$$

Since $\nu(n)$ denotes the combination of received distant signal and noise in Figure 9.16, the error at the subtractor output is

$$e(n) = [\mathbf{h}_e' - \mathbf{c}'(n)]\ \mathbf{d}(n) + \nu(n) = -\xi'(n)\ \mathbf{d}(n) + \nu(n) . \tag{9.32}$$

Then the MSE is

$$\begin{aligned} E\ e_n^2 &= E\left\{(-\xi_n'\mathbf{d}_n + \nu(n))(-\mathbf{d}_n'\ \xi_n + \nu(n))\right\} \\ &= E\left\{\xi_n'\mathbf{d}_n\mathbf{d}_n'\xi_n\right\} - 2\ E\left\{\nu(n)\xi_n'\mathbf{d}_n\right\} + E\nu^2(n) . \end{aligned} \tag{9.33}$$

With the assumption, nonrigorous but in practice, that the tap weights are essentially uncorrelated with the current data vector, the first expectation is

$$E\{\xi_n'\mathbf{d}_n\mathbf{d}_n'\xi_n\} = E\{\xi_n' E\ R\xi_n\} ,$$

where the matrix $R = \mathbf{d}_n\mathbf{d}_n'$ (prior to taking the expected value) has elements

$$r_{jk} = d_{n-j}d_{n-j} . \tag{9.34}$$

The expectation of R is the diagonal matrix AI, where $A = E\{d^2(n+i)\}$ is the power in any data symbol, and recalling that the data symbols are mutually uncorrelated.

The second term on the right hand side of (9.33) is zero, since the noise (including desired distant signal) is uncorrelated with the echo or its replica.

Thus (9.33) becomes

$$E\, e_n^2 = A\, E\{\xi_n'\, \xi_n\} + \sigma^2 , \tag{9.35}$$

where $\sigma^2 = E\{\nu_n^2\}$.

At the $(n + 1)$th iteration,

$$e_{n+1} = \xi_{n+1}' \mathbf{d}(n + 1) + \nu_{n+1} , \tag{9.36}$$

so that (9.35) becomes

$$\begin{aligned} E\{e_{n+1}^2\} &= A\, E\{\xi_{n+1}' \xi_{n+1}\} + \sigma^2 \\ &= \left[1 - 2\beta A + \beta^2(2N + 1)A^2\right] E\{e_n^2\} + 2\beta A\, \sigma^2 . \end{aligned} \tag{9.37}$$

Following Mueller [15] and Werner [16], the solution for this recurrence relationship is

$$\begin{aligned} E\{e_n^2\} = {} & \left[1 - 2\beta A + \beta^2(2N + 1)A^2\right]^n E\{e_0^2\} \\ & + \frac{1 - \left[1 - 2\beta A + \beta^2(2N + 1)A^2\right]^n}{1 - \left[1 - 2\beta A + \beta^2(2N + 1)A^2\right]} [2\beta A \sigma^2] . \end{aligned} \tag{9.38}$$

The expected square error decreases if

$$\left|1 - 2\beta A + \beta^2(2N + 1)A^2\right| < 1 ,$$

i.e.,

$$0 < \beta < 2A/(2N + 1)A^2 . \tag{9.39}$$

As with adaptive equalization, the step size is inversely proportional to the number of taps. Intuitively, the more taps contributing random fluctuations, the slower one must adapt the tap weights by the stochastic algorithm. It is this consideration, not the eigenvalue ratio (which was the most prominent factor in determining convergence rate for the ideal gradient algorithm), which limits the convergence rate for the stochastic algorithm.

The step size providing the fastest convergence rate may be derived by equating to zero the derivative of the bracketed term in (9.38), yielding

$$\beta = 1/\left[(2N + 1)A\right] . \tag{9.40}$$

Substituting (9.40) into the first term of (9.38) shows that convergence proceeds as

$$E\{e_n^2\} = \left[1 - 1/(2N + 1)\right]^n E\{e_0^2\} . \tag{9.41}$$

The residual error after an infinite number of iterations ($n \to \infty$) is found from

(9.38) to be

$$E\{e_\infty^2\} = \sigma^2 / \left[1 - (2N + 1)\beta A/2\right], \tag{9.42}$$

which is the unavoidable "noise" variance, σ^2, increased by a factor which is greater than one for any $\beta \neq 0$. Since the "noise" with variance σ^2 is largely the desired distant signal, a high received SNR (ratio of desired distant signal to echo plus other noise) requires that the excess in $E\{e_\infty^2\}$ above σ^2 be much smaller than σ^2.

It is important to note that convergence for the LMS algorithm *cannot* occur in one step, which *is* possible for the known-gradient algorithm. Even with a unity eigenvalue ratio, for independent data symbols, the structure of the LMS algorithm results in exponential convergence, as per (9.41).

EXAMPLE 2 Suppose that an echo canceler must provide a ratio of desired distant signal to echo + noise of 20 dB. There is a 33-tap FIR filter in the canceler, a symbol power $A = 1$, and additive Gaussian noise 26 dB below the desired distant signal. What is the maximum asymptotic step size β_∞?

ANSWER: $\sigma^2 = P_B + \sigma_n^2$, where P_B is the power in the desired distant signal and σ_n^2 is the power in the Gaussian noise. Since $10 \log_{10}(P_B/\sigma_n^2) = 26$, $P_B = 400\ \sigma_n^2$, and $\sigma^2 = 1.0025\ P_B$.

For a ratio of desired signal to echo + noise of 20 dB, $E\{e_\infty^2\} = 1.01\ P_B$, i.e., the excess MSE is 1/100 (20 dB below) P_B. Thus $1.01\ P_B = 1.0025\ P_B/[1 - 33\beta/2]$, or $\beta = 0.00045$.

Figure 9.17 illustrates some experimental results [16] for SNR vs. step size, with a 32-tap filter, an (uncancelled) echo-to-far-signal ratio of 20 dB, and with and without additive Gaussian noise 24 dB below the distant signal. Convergence characteristics are shown in later sections.

9.3.5 LEAST-SQUARES (KALMAN) ADAPTATION ALGORITHM

The rate of convergence of the stochastic adaptation (LMS) algorithm is such that even with optimum step sizes, the number of iterations required to converge can be an order of magnitude greater than the number of taps to be adapted. As described in Chapter 8, least squares algorithms, minimizing a weighted sum of a number of most recent errors, can, in contrast, be adapted in a number of iterations not much greater than the number of taps to be adapted. This advantage comes with a price, consisting partly of increased complexity and partly of instability under certain conditions. The complexity price can be considerably reduced by the (computationally) fast Kalman algorithm, which brings the complexity down to the order of the gradient algorithm.