

An Adaptive Prefilter for Timing Recovery

by

Amani Sabri Amin, B. Eng.

Cairo University, Egypt

Department of Electrical Engineering

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment of the requirements
for the degree of
Master of Engineering

Department of Electrical Engineering

McGill University

Montréal, Canada

© October, 1985

To the memory of my father

to my mother

and

to my son, Ahmed

Abstract

This thesis presents a new technique for improving the performance of the timing recovery scheme for baseband synchronous pulse amplitude modulation (PAM) data signals. This technique uses adaptive prefiltering to adaptively shape the pulses entering the timing path. A review of the timing recovery problem and the timing jitter of the PAM system is first introduced. Then, a discussion of the properties of the timing wave, including the effects of the prefilter, is presented. The rest of the thesis describes the design, implementation, and performance of the adaptive prefilter with tap spacing of one-quarter of the symbol time interval T . An analysis of an adaptive algorithm for adjusting the tap weights of a tapped-delay line to minimize the mean square distortion is given. The intention is to see the effects of the number of taps and the step size on the speed of convergence. Attention is focused on the convergence of the mean coefficient vector. The thesis concludes with an implementation for a computer simulation examining the technique. Comparison with results obtained from some specific examples shows that the convergence of the mean coefficient vector also leads to fast convergence of the output mean square error.

Résumé

Cette thèse présente une nouvelle technique pour améliorer la performance de la méthode de récupération du rythme appliquée aux signaux de données de bande de base synchrones obtenus par modulation d'impulsions en amplitude (MIA). Cette technique utilise un préfiltrage afin de façonner, adaptativement, les impulsions apparaissant à l'entrée du circuit de synchronisation. Une révision du problème de la récupération du rythme et de la fluctuation de la synchronisation du système de MIA sera tout d'abord présentée. Ensuite, une discussion des propriétés de la séquence de réglage, incluant les effets du préfiltre est donnée. Le reste de la thèse décrit le design, la réalisation et le rendement d'un préfiltre adaptatif dont l'espacement des prises est un sous-multiple du temps d'intervalle T des symboles. L'analyse d'un algorithme adaptatif pour le réglage de la pondération des prises sur une ligne de retard à prises, afin de minimiser la moyenne du carré de la distorsion est incluse. L'effet du nombre de prises ainsi que celui de la dimension de l'échelon sur la vitesse de convergence est aussi analysé. L'attention sera portée sur la convergence du vecteur de coefficients moyens. La thèse conclue par une simulation sur ordinateur de la technique décrite. Les résultats obtenus avec des exemples spécifiques, démontrent que la convergence du vecteur de coefficients moyens, est accompagnée d'une convergence rapide de l'erreur du signal de sortie.

Acknowledgements

The author would like to sincerely thank Professor Peter Kabal for his numerous suggestions, helpful advice and guidance in the realization of this work.

I also owe Professor J. F. Hayes many thanks for his encouragement without which this work would not have been done.

Special thanks to Dr. Youssef Ghoneim for his support and to Mr. Mario Khayat for his kind assistance with the preparation of this thesis.

I would also like to mention the kind assistance of the Computer Vision and Robotics group, and especially Yvan Leclerc, for their help with the computer simulations.

I am grateful to Mr. Mike Parker for the great work he did in typing this thesis.

TABLE OF CONTENTS

Abstract	i
Résumé	ii
Acknowledgements	iii
Chapter 1 Introduction	1
1.1 Review of the timing recovery problem	1
1.2 General outline of the thesis	3
Chapter 2 Timing jitter in PAM timing recovery	5
2.1 PAM timing recovery	5
2.2 Generation of timing jitter	8
2.2.1 Pulse shape	10
2.2.2 Imperfections of a timing circuit	12
Chapter 3 Optimizing pulse shaping	
for a PAM timing recovery scheme	15
3.1 Introduction	15
3.2 Mean timing wave	17
3.3 Variance of timing wave	19
3.3.1 The band limiting condition on the band-pass filter $H(f)$	20
3.3.2 The band limiting condition on overall response $G(f)$	20
3.4 The rms timing deviations	26
3.5 Prefiltering to minimize jitter	28
Chapter 4 Analysis of a self-adjusting prefilter	30
4.1 Properties of the prefilter characteristics	30
4.2 The adaptive prefilter implementation	32

4.2.1	The prefilter tap spacing	32
4.2.2	The tap gain coefficients of the equalizer	39
4.3	The resulting minimum mean square error (MMSE)	44
4.4	Uniqueness of solution for the fractionally-spaced prefilter as the noise vanishes	45
Chapter 5 An iterative algorithm		
	for computing the tap gain coefficients	50
5.1	The MMSE gradient algorithm	50
5.2	The LMS stochastic gradient algorithm	53
5.3	Convergence behavior of the SG transversal algorithm	54
5.3.1	Mean coefficient vector	55
5.3.2	The time constant	58
5.3.3	Normalization of step size	59
5.3.4	Output mean square error	61
Chapter 6 Computer Simulation Study		
6.1	Computer simulation results	64
6.1.1	Simulation results for a data-quality telephone channel	64
6.1.2	Simulation results for a channel with severe intersymbol interference .	77
Chapter 7 Summary and Conclusions		
Appendix A Derivation of the variance of the timing wave		
Appendix B Proof of Equation (4.23)		
Appendix C Proof of Equation (4.24)		
Appendix D Computer Implementation		
D.1	The optimal tap-gain coefficients calculations	96
D.2	Generate the prefilter's input signal at each time interval $T/4$	96

D.3	Generate the prefilter's output signal	97
D.4	Updating the prefilter's coefficients	98
References	99

LIST OF FIGURES

2.1	A simplified model for a baseband PAM data transmission system.	6
2.2	A general timing extraction model [12]	9
2.3	An interpretation of the waveform function [8]	13
3.1	Signal path and timing path in a PAM receiver [9]	16
3.2	Graphical interpretation for the band-limiting condition on the band-pass filter $H(f)$	21
3.3	Variation in the mean square value of the timing wave process [9]	27
4.1	Graphical representation for the overall response in the time and frequency domains.	31
4.2	A model for the timing circuit including the fractional tap-delay prefilter ..	35
4.3	Fractional tap delay-line prefilter model	36
4.4	Spectral overlap at the input to a T equalizer	38
4.5	Conditions for solution uniqueness for a fractionally-spaced prefilter	49
6.1	The equivalent discrete-time characteristic for a data-quality telephone channel	67
6.2	Amplitude spectrum for the channel shown in Figure 6.1	67
6.3	The last 100 samples of the input prefilter waveform from 5000 samples, using the channel of Figure 6.1	68
6.4	The result of the last 100 iterations of the output signal of the prefilter for 5000 iterations for $\Delta_s = 0.02$ using the channel of Figure 6.1	69
6.5	The tap MSE convergence rate for $\Delta = 0.02$ and 41 taps using the channel of Figure 6.1	70
6.6	The result of the last 100 iterations of the output waveform for 5000 iterations	

for $\Delta = 0.004$ and 41 taps for the channel of Figure 6.1	71
6.7 The tap MSE convergence rate for $\Delta = 0.004$ and 41 taps, using the channel of Figure 6.1	72
6.8 The tap MSE convergence rate for $\Delta = 0.01$ using the channel given in Figure 6.1	75
6.9 The resulting last 100 iterations of the output waveform for 5000 iterations for $\Delta = 0.01$ using the channel from Figure 6.1	76
6.10 The equivalent discrete-time characteristic for a channel with severe intersymbol interference	79
6.11 Amplitude spectrum for the channel of Figure 6.10	79
6.12 The last 100 samples of the input prefilter waveform, from 5000 samples using the channel given in Figure 6.10	80
6.13 The result of the last 100 iterations of the output prefilter waveform, from 5000 iterations, for $\Delta = 0.115$, using the channel of Figure 6.10	81
6.14 The convergence characteristic of MSE of the prefilter with 9 taps for $\Delta = 0.115$, using the channel of Figure 6.10	82
6.15 The result of the last 100 iterations of the output waveform, from 5000 iterations, for $\Delta = 0.012$ and the channel of Figure 6.10	83
6.16 The convergence characteristics of the MSE for $\Delta = 0.012$ and the channel of Figure 6.10	84

LIST OF TABLES

1	The eigenvalues of the autocorrelation matrix corresponding to a data-quality telephone channel (Figure 6.1)	66
2	The eigenvalues of the autocorrelation matrix corresponding to a channel with severe intersymbol interference (Figure 6.7)	78

Chapter 1

Introduction

1.1. Review of the timing recovery problem

Synchronization is the process of aligning the time scales between two or more periodic processes that are occurring at spatially separated points. This is one of the most critical receiver functions in synchronous communication systems. The receiver synchronization problem is to obtain accurate timing information indicating the optimal sampling instants of the received data signal. This is called clock synchronization, timing recovery or clock recovery. In early systems, the timing information was transmitted on a separate channel or by means of a discrete spectral line at an integer multiple of the clock frequency imposed on the data signal itself. However, these systems had many disadvantages, including inefficient utilization of bandwidth and poor jitter performance.

In digital communication systems that are efficient in power requirements and bandwidth occupancy, the timing information must be derived from the data signal itself and based on some meaningful optimization criterion which determines the steady-state location of the timing instants. For this self-timed mode of operation, the received data signal is fed into a timing circuit which produces a “timing wave”

which ideally has some periodic attribute, such as uniformly spaced zero crossings. Bandwidth- and power-efficient signal design usually dictates that any discrete component of the clock be suppressed and transmitted power be devoted exclusively to data. In this case, the timing circuit must regenerate a clock from a signal that does not contain timing information in explicit form. Nonlinear devices are necessary to regenerate a discrete reference from a signal in which the clock has been suppressed. For binary or multilevel pulse amplitude modulation (PAM) data signals, several timing recovery methods are known [1-8]. A broad classification into analog and digital schemes can be made.

Some analog schemes use the threshold crossings (at zero if the signal is two-level, or halfway between the reference levels if the signal is multilevel) or the derivative of the received baseband data signal to correct or update the phase of the receiver clock [3-5]. These schemes can be used with a variety of algorithms within the control loop and different parameters can be used during the initial training mode and during the subsequent tracking mode. These types of systems operate on the baseband signal. Other analog systems use pre-filtering and nonlinear processing of the received waveform to generate a spectral line at the clock frequency (or an integer multiple thereof) [6-9]. An advantage of these systems is their ability to work with either the baseband or the passband signal. However, the performance of analog systems depends significantly on the availability of signal bandwidth in excess of the Nyquist bandwidth (i.e., one half the frequency corresponding to the bit rate) at the receiver.

The current trend in timing recovery systems is towards a fully digital implementation using the presently available medium- and large-scale integration (MSI and

LSI) technology [10]. In such a system, the received data signal is sampled by the clock signal, A/D converted and processed. Then timing recovery is usually achieved with a feedback configuration.

Previous work has indicated that the main cause of the fluctuation in the timing wave is the random nature of the data sequence [6, 8, 14, 15]. Control of this randomness by adaptively controlling the pulse shape is the solution proposed in this thesis. Originally, prefiltering the data sequence (with a fixed prefilter) before the nonlinear device was proposed by Takasaki [8]. Franks and Bubrouski [9] give a condition on the overall response of this prefilter for jitter-free timing recovery.

In the following work, this overall response is used to design an adaptive fractionally-spaced prefilter. Some of the analyses of fractionally-spaced equalizers [Ungerboeck, [18], Gitlin and Weinstein [19]] are readily adaptable to the case of a fractionally-spaced adaptive prefilter. The prefilter problem differs from the equalizer problem in that it requires normalization of the tap-gain coefficients. This leads to a new adaptive algorithm. Simulation results are presented which indicate that the adaptive prefilter leads to improvements in the timing wave regularity.

The added complexity of an adaptive prefilter can be justified by the better timing recovery performance, and with digital signal processing hardware costs continually decreasing, the cost of a hardware implementation will be minimal.

1.2. General outline of the thesis

In chapter 2 of this thesis, a review of the PAM timing recovery problem is presented in Section 2.1. Investigation of the sources of timing jitter in the timing wave makes up Section 2.2.

In chapter 3, we study the data pulse shape for a baseband PAM signal in Section 3.1. The statistical properties of the lowest-order moments of the cyclostationary timing wave are examined: the mean timing wave in Section 3.2 and the variance in Section 3.3. In Section 3.4 an expression for rms timing jitter is obtained. From this expression the effects of pulse shape, bandwidth and mistuning of the circuit are deduced. The result of the investigation, which is a condition on the prefilter and the passband filter for complete elimination of timing jitter, is presented in Section 3.5.

Chapters 4 and 5 contain the details of the design and implementation of the prefilter as an adaptive equalizer. Starting from the properties of its characteristics in Section 4.1, The implementation of the adaptive prefilter by a fractional tap-spaced equalizer is covered in Section 4.2. Minimizing the mean-squared error using an iterative algorithm for computing the tap gain coefficients is presented in chapter 5. The least mean square algorithm and its convergence properties are studied, as an approximation for examining the convergence rate of the MSE, in Sections 5.2 and 5.3, respectively.

In chapter 6, the results of the computer simulations are presented, with a comparison between a data-quality telephone channel and a channel with severe inter-symbol interference for different values of the step size.

Chapter 2

Timing jitter in PAM timing recovery

2.1. PAM timing recovery

A simplified model of a baseband PAM data transmission system, with an overall impulse response $g(t)$, is shown in Figure 2.1. The received PAM signal can be described as

$$u(t) = \sum_{k=-\infty}^{\infty} a_k g(t - kT) + n(t) \quad (2.1)$$

where $\{a_k\}$ is the discrete-valued data sequence and $n(t)$ is the additive noise process. For the most part in the signal, the data sequence will be assumed to consist of independent, identically distributed zero mean values. In many data transmission systems scramblers are used to obtain an approximately independent sequence. As will be indicated later, the noise term $n(t)$ can often be neglected in the timing path.

Now, the receiver synchronization problem is to find the correct sampling instants for extracting the data sequence. Assuming a normalization of $g(0) = 1$, and that $g(t)$ is defined so that the best sampling instants are at $t = iT, i = 0, \pm 1, \pm 2, \dots$,

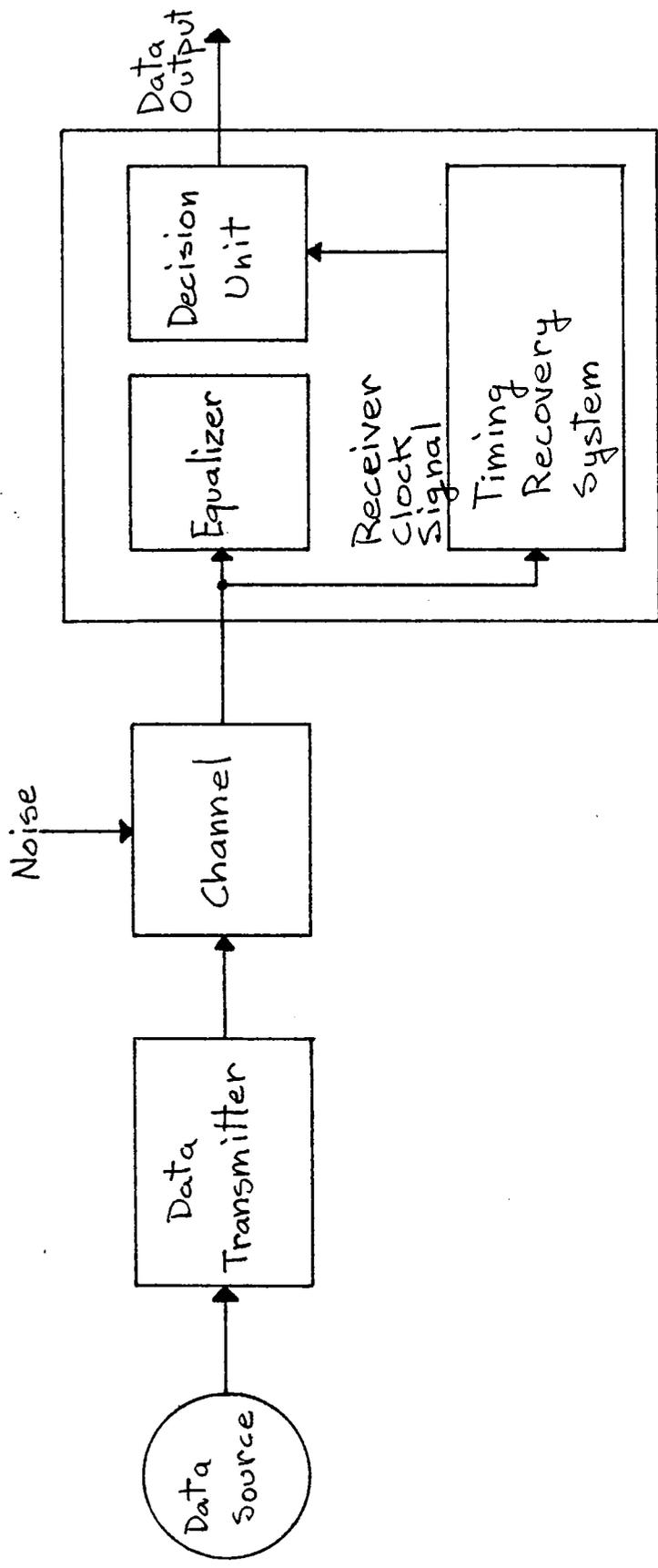


Figure 2.1. A simplified model for a baseband PAM data transmission system.

i.e.,

$$\begin{aligned}\hat{a}_i &= u(iT) \\ &= a_i + \sum_{\substack{k \\ k \neq i}} a_k g(iT - kT) + n(iT)\end{aligned}$$

The first term in the right-hand side of this equation is the desired value. The second and third terms represent intersymbol interference and noise, respectively. The objective is to recover a close replica of the message sequence $\{a_i\}$ in terms of the sequence $\{\hat{a}_i\}$. In the noise-free case, the difference between a_i and \hat{a}_i is due to intersymbol interference, which can be minimized by proper shaping of the data pulse $g(t)$. With perfect timing, the intersymbol interference is

$$\hat{a}_i - a_i = \sum_{k \neq i} a_k g(iT - kT) \quad (2.2)$$

and this term can be made to vanish for pulses satisfying the Nyquist criterion, i.e., $g(iT) = 0$ for $i \neq 0$. For band-limited Nyquist pulses, the intersymbol interference will not be zero for imperfect timing, and if the bandwidth is not close to the Nyquist bandwidth ($1/(2T)$), the intersymbol interference can be quite severe even when the timing error is small.

In some cases, the timing circuit is simply a narrow-band filter tuned to a harmonic of the pulse repetition frequency ($1/T$). This scheme works in situations where the PAM signal has discrete spectral components. Assuming a random data sequence modulating the signal, the existence of discrete spectral components requires both that the data sequence have a nonzero mean value and that the Fourier transform of the data pulse not vanish at some multiple of the pulse repetition frequency [11]. In the interest of meeting power and bandwidth limitations, it is desirable to design a

system where neither of these conditions holds. It has been recognized, however, that the same timing recovery circuit will work if a nonlinear element, such as a square law device or a full wave rectifier is inserted before the narrow-band filter [7–9]. A timing circuit involving a square law device followed by a narrow-band filter will give satisfactory performance even when the PAM signal is band limited to less than the pulse repetition frequency.

This widely used timing recovery system is illustrated in Figure 2.2 [12]. The incoming pulse train $u(t)$ first undergoes a nonlinear processing (NLC), which introduces a spectral line at the pulse rate or a multiple thereof. The NLC is assumed to be memoryless, with an input-output relationship specified by a characteristic $f(\cdot)$, so that its output is $e(t) = f(\{x(t)\})$. Then, the signal is filtered by a narrow-band resonant tank circuit, tuned as close as possible to the pulse rate, resulting in the sinusoidal timing wave. Narrow pulses are finally generated at the negative-going zero crossing of the timing wave. In modern receivers, narrow-band resonant circuits are often implemented using phase-locked loop circuits.

2.2. Generation of timing jitter

In the ideal situation the timing wave would consist of regularly spaced pulses, as mentioned before. However, this never occurs exactly because a constant-amplitude, correct-frequency timing wave is not realizable in practice. The timing wave has random modulations in both amplitude and phase and consequently pulses generated at the zero-crossings contain fluctuations which represent an error, time jitter, in the extracted timing information.

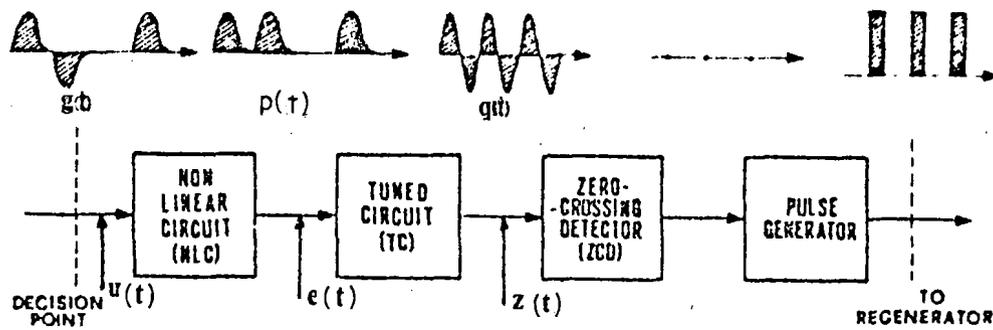


Figure 2.2. A general timing extraction model [12]

Fluctuations can arise from additive or multiplicative disturbances in the channel or from self-noise, which is called systematic jitter. The main causes of this noise are the random nature of the data itself and of the action of the nonlinear device. Another name for this kind of disturbance is pattern-dependent jitter, because the fluctuations depend on the data sequence.

Jitter investigations have shown that jitter due to random noise is not a serious problem in practice and clearly indicate that the predominant jitter is systematic and is caused by the signal pulse pattern [7–8, 13–14]. The sources of self-noise in the timing wave are (1) pulse shape and (2) imperfections of a timing circuit, i.e., (a) amplitude to phase conversion and (b) mistuning of a timing circuit. These are discussed separately below.

2.2.1. Pulse shape

As pointed out by J. M. Manley [14], the spectrum of the pulse train contributes very little timing noise itself if the pulses are narrow. This spectrum consists of discrete components at harmonics of the pulse rate and broadband noise which is divided into small, evenly spaced bands. Taking these side frequencies in pairs about the pulse rate, we may think of the pulse train in the vicinity of the pulse rate as a carrier wave, amplitude modulated by a number of small components. This amplitude modulation at the output of the pass-band filter may cause further phase modulation. Here the pulse train itself is now a source of timing jitter components, having a spectrum with nonzero value at zero frequency.

The dependence of static pattern jitter on pulse overlaps has been investigated in [8]. It was shown that the complex weight of the timing wave $Z^2(f_0)$ can be

expressed as follows [8, Eqs. (5)–(12)]:

$$Z^{(2)}(f_0) = 2 \int_0^{f_0/2} |W(f)|^2 \hat{H}(f_0, f) df \quad (2.3)$$

where

$$W(f) = \sum_{i=0}^{M-1} a_i \exp \left[-2\pi j \left(\frac{f}{f_0} \right)^i \right] \quad (2.4)$$

$$\hat{H}(f_0, f) = \sum_{n=-\infty}^{\infty} H(f_0 - \overline{f - nf_0}) H(f + nf_0) \quad (2.5)$$

and where

f_0 : pulse repetition frequency

M : number of time slots in a pattern period

a_i : amplitude of the pulse in the i^{th} time slot

$H(f)$: spectrum of an isolated pulse

$W(f)$: spectrum of a data sequence of impulses

$\langle 2 \rangle$: indicates the use of square-law nonlinearity

The amplitude and phase of the timing wave are given by $|Z^{(2)}(f_0)|$ and angle $\angle Z^{(2)}(f_0)$, respectively. The function $|W(f)|^2$ is called the “pattern function”, and $\hat{H}(f_0, f)$, the “waveform function”. This nomenclature is used because they depend only on pulse patterns and pulse waveforms, respectively. The fact that the inner product of these two functions is expressed by the complex timing wave function $X^{(2)}(f_0)$, together with the fact that the pattern function is a non-negative real function, make timing jitter analyses simpler than they might otherwise be. Let us

confine our attention to $\hat{H}(f_0, f)$. The relation between the spectrum of a pulse waveform and its waveform function can be interpreted as illustrated in Figure 2.3. As seen from the Figure and Equation (2.5), a waveform function is constant when the function

$$\tilde{H}(f_0, f) = \left[\frac{|H(f_0/2)|}{H(f_0/2)} \right]^2 H(f_0/2 + f)H(f_0/2 - f) \quad (2.6)$$

satisfies Nyquist's first criterion. For instance, $\hat{H}(f, f_0)$ is constant for an ideal low-pass function

$$H(f) = \begin{cases} 1 & |f| \leq f_0 \\ 0 & \text{otherwise} \end{cases} \quad (2.7)$$

In this case pulse overlaps have no influence on the timing wave. The amplitude of the waveform function, however, need not be constant. On the other hand, the angle of the waveform function must be constant, since otherwise the phase of the timing wave fluctuates with the variation of digital patterns. One can easily verify this by using the fact that a symmetrical waveform makes the angle of $\hat{H}(f_0, f)$ constant. It is seen from Equation (2.3) that the worst jitter case corresponds to the maximum phase difference in $\hat{H}(f_0, f)$.

2.2.2. Imperfections of a timing circuit

a) Amplitude to phase conversion

Amplitude to phase conversion occurs in detecting zero crossings of the timing wave in order to form spikes for ideal retiming. If the threshold is not exactly at the zero level of the timing wave, a phase shift will be introduced. This depends on the amplitude of the timing wave, which varies with the density of the pattern. The amplitude variation chiefly depends on the use of coding schemes such as constrained bipolar or PST [8]. Pulse overlaps, however, often exacerbate the amplitude variation.

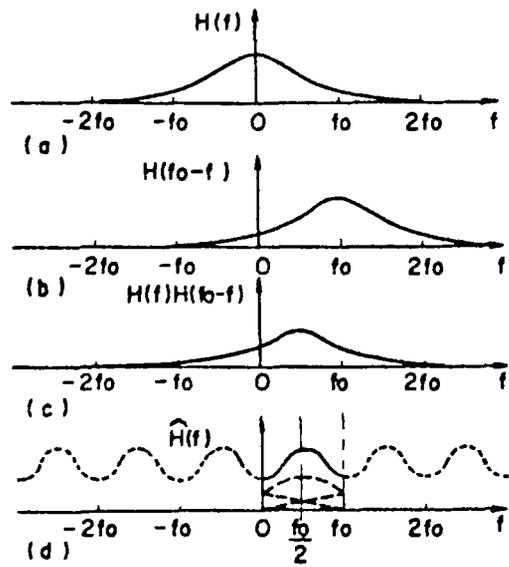


Figure 2.3. An interpretation of the waveform function [8]

The magnitude of a timing wave can be represented as

$$\begin{aligned} Z_k^{(2)}(f_0) &= W_k \cdot \hat{H} \\ &\triangleq |W_k| \cdot |\hat{H}| \cdot \cos \theta_k \end{aligned} \quad (2.8)$$

We want to minimize the variation ratio

$$R = \frac{\max_k \{|W_k| \cos \theta_k\}}{\min_\ell \{|W_\ell| \cos \theta_\ell\}} \quad (2.9)$$

Since the optimum waveform functions that reduce the amplitude variation ratio R are sometimes not very suitable for pulse detection, we must use prefiltering in a timing path which can properly shape the amplitude-frequency characteristics of a waveform function.

b) Mistuning of a timing circuit

The tuned circuit which selects the pulse rate fundamental from the spectrum of the incoming signal to provide a timing wave also admits some of the self-noise side frequencies, which are still symmetrical if the tuned circuit is centered exactly on the pulse rate. When it is detuned from the pulse rate, the side-frequency pairs in the response are no longer symmetrical. The asymmetry in amplitude, phase, or both, is equivalent to phase modulation of the timing wave. But the most important characteristic of the spectra of this kind of timing noise, caused by mistuning only, is that there is no energy at zero frequency [14], which means this source of jitter is not serious compared to the other sources.

Chapter 3

Optimizing pulse shaping for a PAM timing recovery scheme

3.1. Introduction

The timing circuit under consideration is shown in Figure 3.1. The received PAM signal is subjected to additional filtering by the prefilter in the timing path (as explained in Section 2.3.2). The individual data pulse shape at the input to the square-law device is denoted by $g(t)$. The model for the baseband PAM signal is repeated here for convenience,

$$u(t) = \sum_{k=-\infty}^{\infty} a_k g(t - kT) \quad (3.1)$$

Before examining the circuit, it is helpful to ask what properties the signal $u(t)$ must possess so that operations on $u(t)$ to produce a good estimate of the sampling instants. A general answer to this question lies in the cyclostationary nature of the $u(t)$ process. A cyclostationary process has statistical moments which are periodic in time rather than constant as in the case of stationary processes. To a large extent, synchronization capability can be characterized by the lowest-order moments of the process, such as the mean and autocorrelation. The $u(t)$ process

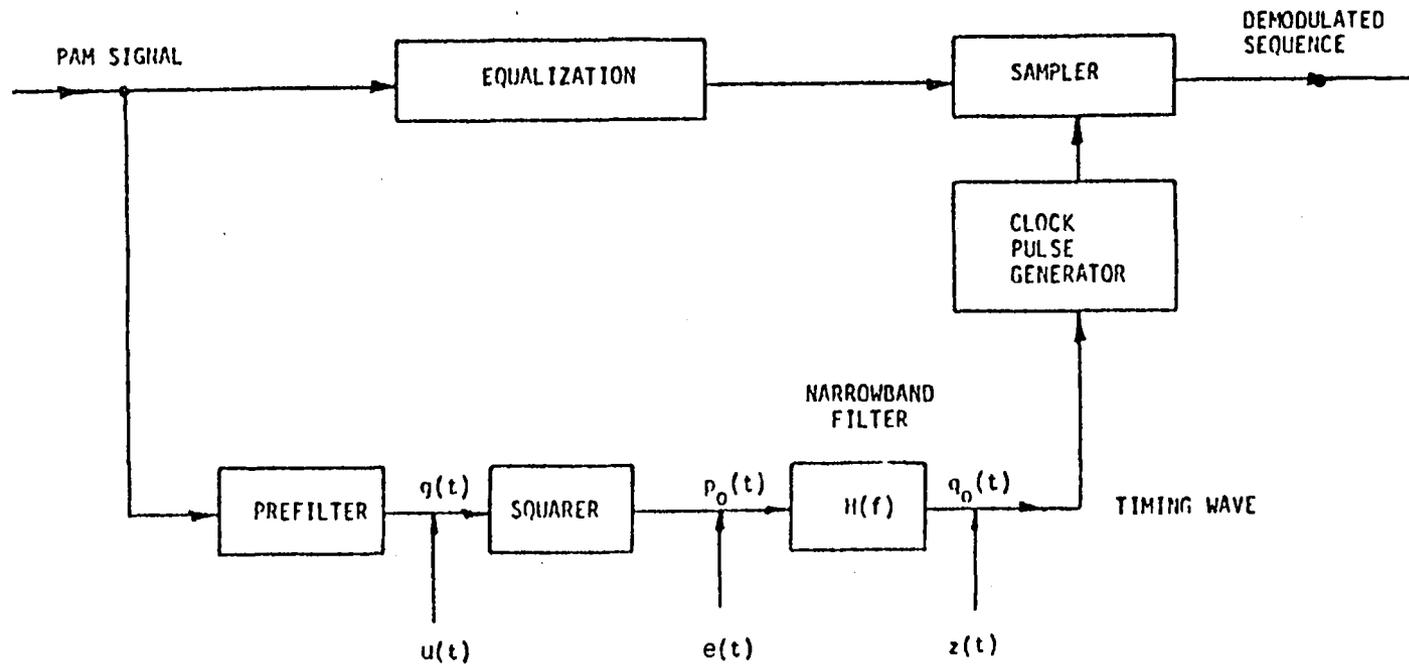


Figure 3.1. Signal path and timing path in a PAM receiver [9]

is said to be cyclostationary in the wide sense if $E[u(t)]$ and $R_{uu}(t + \tau, t)$ are both periodic functions of t . A process modeled by Equation (3.1) is cyclostationary with period of T [15].

We assume that $\{a_k\}$ is a zero-mean stationary sequence with independent elements. The resulting PAM signal is a zero-mean cyclostationary process, although there may not be any periodic components present.

3.2. Mean timing wave

The square of the PAM signal does, however, possess a periodic mean value [9]

$$E[u^2(t)] = \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} R_m p_m(t - kT) \quad (3.2)$$

where

$$R_m = E[a_k a_{k+m}]$$

$$p_m(t) = g(t)g(t - mT)$$

It is convenient to express the periodic mean function in terms of its Fourier series. Using the Poisson Sum Formula [11],

$$E[u^2(t)] = \frac{1}{T} \sum_{m=-\infty}^{\infty} \sum_{\ell=-\infty}^{\infty} R_m P_m \left(\frac{\ell}{T} \right) \exp \left(\frac{j2\pi \ell t}{T} \right) \quad (3.3)$$

where

$$P_0 \left(\frac{\ell}{T} \right) \triangleq \int_{-\infty}^{\infty} G \left(\frac{\ell}{T} - f \right) G(f) df$$

For high bandwidth efficiency, we are often concerned with data pulses whose bandwidth is at most equal to twice the Nyquist bandwidth. Then $|G(f)| = 0$ for $|f| > 1/T$ and there are only three nonzero terms ($\ell = 0, \pm 1$) in Equation (3.3).

Then,

$$E[u^2(t)] = \frac{1}{T} \sum_{m=-\infty}^{\infty} R_m \left[P_m(0) + 2P_m \left(\frac{1}{T} \right) \cos \left(\frac{2\pi t}{T} \right) \right] \quad (3.4)$$

The band-pass filter $H(f)$ (Figure 3.1) is tuned to the symbol rate, $1/T$ [8, 15–16] with band-limiting condition

$$H \left(\frac{\ell}{T} \right) = 0 \quad \text{for } |\ell| \neq 1 \quad (3.5)$$

and with the assumption that the data sequence is stationary independent and has unit variance and zero mean, i.e., $\overline{a^2} = 1$ and $E[a_k a_{k+m}] \triangleq R_{aa}(m) = 0$ for $m \neq 0$.

Then, the mean timing wave is a sinusoid with a phase of ϕ or $-2\pi\epsilon/T$, for a real $G(f)$, i.e.,

$$\begin{aligned} E(z) &= 2|U_1| \cos \left(\frac{2\pi t}{T} + \phi \right) \\ &= 2|U_1| \cos \left(\frac{2\pi t}{T} - \frac{2\pi\epsilon}{T} \right) \end{aligned} \quad (3.6)$$

where

$$\begin{aligned} U_1 &= \frac{1}{T} H \left(\frac{1}{T} \right) \int_{-\infty}^{\infty} C(f) df \\ C(f) &= G(f)G \left(\frac{1}{T} - f \right) \end{aligned}$$

$$z(t) \triangleq \text{the timing wave}$$

We see that the zero-crossings of the mean timing wave are at a fixed time offset ($T/4$) relative to the desired sampling instants.

3.3. Variance of timing wave

The actual zero crossings of $z(t)$ fluctuate about the desired sampling instants because the timing wave depends on the actual realization of the entire data sequence, so different zero crossings result for different data sequences. To evaluate the statistical nature of this jitter, we need to obtain the variance of the timing wave as given by Equation (3.7) [Franks, 9]. The details of the derivations are carried out in Appendix A.

$$\text{Var } z(t) = \sum_{r=-\infty}^{\infty} V_r \exp \left[\frac{j2\pi r t}{T} \right] \quad (3.7)$$

where

$$V_r = \frac{2R(0)}{T} A \left(\frac{r}{T} \right) + \frac{\overline{a^4} - 3R(0)}{T} B \left(\frac{r}{T} \right) \quad (3.7a)$$

where

$$A(f) = \frac{1}{T} \sum_{\ell=-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(\nu) H(f - \nu) G \left(\nu + \eta - \frac{\ell}{T} \right) G(\eta) G \left(\frac{\ell}{T} - \eta \right) G(f - \eta - \nu) d\nu d\eta \quad (3.7b)$$

$$B(f) = \iiint_{-\infty}^{\infty} H(\nu) H(f - \nu) G(\nu - \eta) G(\eta) G(\lambda) G(f - \lambda - \nu) d\lambda d\eta d\nu \quad (3.7c)$$

Finally, we can restrict the bands for $G(f)$ and $H(f)$ in order to simplify the evaluation of the V_r . These are reasonable assumptions for bandwidth-efficient systems.

3.3.1. The band limiting condition on the band-pass filter $H(f)$

$$H(f) = 0 \quad \text{for } \left| |f| - \frac{1}{T} \right| > \frac{1}{2T} \quad (3.8)$$

This condition limits the bandwidth of $H(f)$ to $1/T$ (centered at $1/T$).

From Equation (3.7c),

$$B(f) = \iiint_{-\infty}^{\infty} H(\nu)G(\nu - \eta)G(\eta)G(\lambda)H(f - \nu)G(f - \lambda - \nu) d\eta d\nu d\lambda \quad (3.9)$$

Then,

$$B\left(\frac{r}{T}\right) = \iiint_{-\infty}^{\infty} H(\nu)G(\nu - \eta)G(\eta)G(\lambda)H\left(\frac{r}{T} - \nu\right)G\left(\frac{r}{T} - \lambda - \nu\right) d\eta d\nu d\lambda \quad (3.10)$$

We can see from Figure 3.2 that the condition for $H(f)$ causes the Fourier coefficients V_r in Equation (3.7) to vanish except for $r = 0$ and ± 2 . This means that $\text{Var } z(t)$ is simply a constant plus a sinusoidal term and that only two Fourier coefficients need to be evaluated.

3.3.2. The band limiting condition on overall response $G(f)$

$$G(f) = 0 \quad \text{for } |f| > \frac{1}{T} \quad (3.11)$$

From Equation (3.7b),

$$A(f) = \frac{1}{T} \sum_{\ell=-\infty}^{\infty} \int_{\eta=-\infty}^{\infty} \int_{\nu=-\infty}^{\infty} H(f - \nu)H(\nu)G(f - \nu - \eta)G(\eta)G\left(\nu + \eta - \frac{\ell}{T}\right)G\left(\frac{\ell}{T} - \eta\right) d\nu d\eta \quad (3.12)$$

But since $G(f)$ is real and symmetrical,

$$G\left(\nu + \eta - \frac{\ell}{T}\right) = G\left(\frac{\ell}{T} - \nu - \eta\right)$$

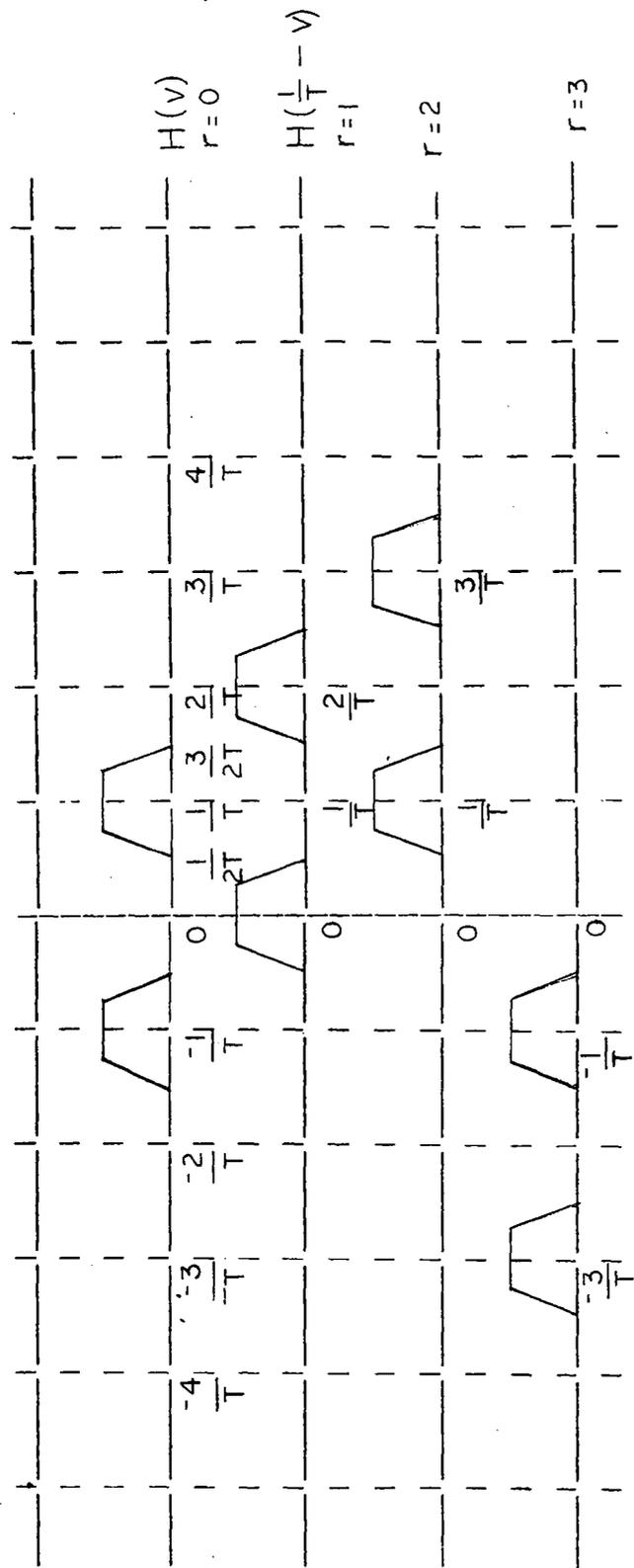


Figure 3.2. Graphical interpretation for the band-limiting condition on the band-pass filter $H(f)$

From the results in 3.3.1, we need to calculate $A(f)$ and the Fourier coefficients V_r for $r = 0, \pm 2$ only.

(1) For $r = 0$

$A(r/T)$ will vanish except for $l = 0$ and ± 1 because of the term $G(\frac{\ell}{T} - \nu - \eta) \times G(\frac{r}{T} - \nu - \eta)$ in Equation 3.12. For the particular case of $(+1, -1)$ binary data, $\bar{a}^4 = R_0^2 = 1$ and under conditions (3.8) and (3.11), we have

$$V_0 = \frac{2}{T}A(0) - \frac{2}{T}B(0) \quad (3.13)$$

$$A(0) = \frac{1}{T} \int_{-\infty}^{\infty} |H(f)|^2 \left\{ |G(f+v)|^2 |G(v)|^2 dv + 2\text{Re} \int_{-\infty}^{\infty} C^*(f+v)C(v) dv \right\} df \quad (3.14)$$

$$B(0) = \int_{-\infty}^{\infty} |H(f)|^2 \left| \int_{-\infty}^{\infty} G(f-v)G(v) dv \right|^2 df \quad (3.15)$$

Then,

$$V_0 = \frac{2}{T^2} \int_{-\infty}^{\infty} |H(f)|^2 \left\{ \int_{-\infty}^{\infty} |G(f_v)|^2 |G(v)|^2 dv - T \left| \int_{-\infty}^{\infty} G(f-v)G(v) dv \right|^2 + 2\text{Re} \int_{-\infty}^{\infty} C^*(f+v)C(v) dv \right\} df \quad (3.16)$$

(2) For $r = \pm 2$

$$A(f) = A\left(\frac{r}{T}\right) = A\left(\frac{2}{T}\right) = \frac{1}{T} \sum_{\ell=-1}^1 \iint_{-\infty}^{\infty} H\left(\frac{2}{T} - \nu\right) H(\nu) G\left(\frac{2}{T} - \nu - \eta\right) G(\eta) G\left(\nu + \eta - \frac{\ell}{T}\right) G\left(\frac{\ell}{T} - \eta\right) d\nu d\eta \quad (3.17)$$

From Equation (3.11), the bandwidth of $G(f) = 0$ for $|f| > 1/T$. Then from the terms $G\left(\frac{\ell}{T} - \eta\right) G(\eta)$ we can see that $\ell = 0, \pm 1$ will give a response, but from the terms $G\left(\frac{2}{T} - \nu - \eta\right) G\left(\nu + \eta - \frac{\ell}{T}\right)$ we see that $A(2/T)$ will vanish except for $\ell = 1$, so then

$$A\left(\frac{2}{T}\right) = \frac{1}{T} \iint_{-\infty}^{\infty} H\left(\frac{2}{T} - \nu\right) H(\nu) G\left(\frac{2}{T} - \nu - \eta\right) G(\eta) G\left(\nu + \eta - \frac{1}{T}\right) G\left(\frac{1}{T} - \eta\right) d\nu d\eta$$

From Equation (3.6),

$$C\left(f - \frac{1}{T} + \nu\right) = G\left(f - \frac{1}{T} + \nu\right) G\left(\frac{2}{T} - \nu - f\right)$$

which leads to

$$A\left(\frac{2}{T}\right) = \frac{1}{T} \int_{-\infty}^{\infty} H\left(\frac{2}{T} - f\right) H(f) \left\{ \int_{-\infty}^{\infty} C\left(f - \frac{1}{T} + \nu\right) C(\nu) d\nu \right\} df \quad (3.18)$$

Then,

$$\begin{aligned} B\left(\frac{2}{T}\right) &= \iiint_{-\infty}^{\infty} H(\nu) H\left(\frac{2}{T} - \nu\right) G(\nu - \eta) G\left(\frac{2}{T} - \lambda - \nu\right) G(\eta) G(\lambda) d\nu d\eta d\lambda \\ &= \int_{-\infty}^{\infty} H(\nu) H\left(\frac{2}{T} - \nu\right) \\ &\quad \int_{-\infty}^{\infty} G(\eta) G(\nu - \eta) d\eta \int_{-\infty}^{\infty} G(\lambda) G\left(\frac{2}{T} - \lambda - \nu\right) d\lambda d\nu \end{aligned}$$

But from Equation (3.2),

$$P_0(f) = G(f) * G(f) = \int_{-\infty}^{\infty} G(\nu) G(f - \nu) d\nu$$

and

$$P_0\left(\frac{2}{T} - f\right) = \int_{-\infty}^{\infty} G(\nu)G\left(\frac{2}{T} - f - \nu\right) d\nu$$

Then

$$B\left(\frac{2}{T}\right) = \int_{-\infty}^{\infty} H(f)H\left(\frac{2}{T} - f\right) P_0\left(\frac{2}{T} - f\right) P_0(f) df \quad (3.19)$$

and then

$$V_2 = \frac{2}{T^2} \int_{-\infty}^{\infty} H\left(\frac{2}{T} - f\right) H(f) \left\{ \int_{-\infty}^{\infty} C\left(f - \frac{1}{T} + \nu\right) C(\nu) d\nu - TP_0\left(\frac{2}{T} - f\right) P_0(f) \right\} df \quad (3.20)$$

Now we can rewrite the variance of the time wave (Equation (3.7)) in the form

$$\text{Var } z(t) = V_0 + |V_2| \cos\left(\frac{4\pi}{T}t + \theta\right) \quad (3.21)$$

From Equations (3.16) and (3.20), we can see that the variance coefficients are constants, depending on $G(f)$ and $H(f)$, with $V_0 \geq V_2 > 0$. The cyclostationary nature of the timing wave is apparent from this expression. Letting $H(1/T) = 1$, the mean timing wave in Equation (3.6) will be independent of the bandwidth of $H(f)$ but as we can see from the expression for the variance, as the bandwidth of $H(f)$ approaches zero, the value of V_2 approaches V_1 . Therefore, there is a considerable time variation in the variance of $z(t)$. Note that the minimum variance occurs just at the instant of the mean zero crossings, hence the fluctuations in zero crossings are much less than would be expected from a consideration of the average variation of the timing wave over a symbol period. This again points out the error in disregarding the

cyclostationary nature of the timing wave process as, for example, in using the power spectral density of the squarer output to analyze the jitter phenomenon. The mean timing wave can be regarded as a kind of discriminator characteristic or S-curve for measuring the parameter θ . For the band-limited case we are discussing here, this S-curve is just a sinusoid, with a zero crossing at the true value of the parameter. Discrimination is enhanced by increasing the slope at the zero crossing, which means less sensitivity to noise perturbation. As this slope is proportional to U_1 , we can see how the shape of the data pulse $g(t)$ affects timing recovery. From Equation (3.6) we see that the value of U_1 depends on the amount of overlap of the functions $G(f)$ and $G(\frac{1}{T} - f)$, and hence it depends on the amount by which the bandwidth of $G(f)$ exceeds the $1/2T$ Nyquist bandwidth. With no excess bandwidth, $U_1 = 0$ and this method of timing recovery fails. The situation improves rapidly as the excess bandwidth factor increases from 0 to 100 percent. With very large increases in bandwidth there are more harmonic components in the mean timing wave, and its zero-crossing slope can be further increased without increasing signal level by proper phasing of these components.

3.4. The rms timing deviations

Calculation of the statistical properties of the actual zero crossings of the timing wave is difficult. A useful approximation can be obtained by locating the zero crossing by linear interpolation using the mean slope at the mean zero crossing.

A typical realization of the timing wave process appears as a nearly sinusoidal waveform with slowly varying amplitude and small fluctuations in phase-shift. The mean-squared value of the timing wave is periodic in t and will dip to a minimum

value in the vicinity of the zero crossings of $z(t)$ as indicated in Figure 3.3. This minimum value of $E[z^2(t)]$ is an effective indication of the amount of fluctuation in the position Δt of zero crossings. When this quantity is very small relative to the total power in the timing wave, then it is clear that zero crossings will be tightly clustered about the mean zero crossing t_0 . The rms deviations from the t_0 relative to the pulse interval T can be expressed by an approximate formula involving only the first and second moments of $z(t)$ [9].

$$\left(\frac{\Delta t}{T}\right)_{\text{rms}} = \frac{\frac{1}{T} \{E^2[z(t_0)]\}^{1/2}}{\text{mean slope at } t_0} \quad (3.22)$$

where t_0 is the mean zero crossing given approximately by $E[z(t_0)] = 0$.

Using Equation (3.6) we can determine that

$$\frac{t_0}{T} = \frac{n}{4} - \frac{\phi}{2\pi} \quad (3.23)$$

where n is any odd integer.

So, the expression for rms relative time jitter, Equation (3.22), can be rewritten in terms of the Fourier coefficients for the mean and variance of the timing wave

$$\left(\frac{\Delta t}{T}\right)_{\text{rms}} = \frac{1}{4\pi |U_1|} \{V_0 - 2|V_2| \cos(\theta - 2\phi)\}^{1/2} \quad (3.24)$$

where θ and ϕ are the phase angles for V_2 and U_1 respectively. For $\theta \neq 2\phi$ the maximum variance does not occur at t_0 , but at the point t_1 given by

$$\frac{t_1 - t_0}{T} = \frac{2\phi - \theta}{4\pi} \quad (3.25)$$

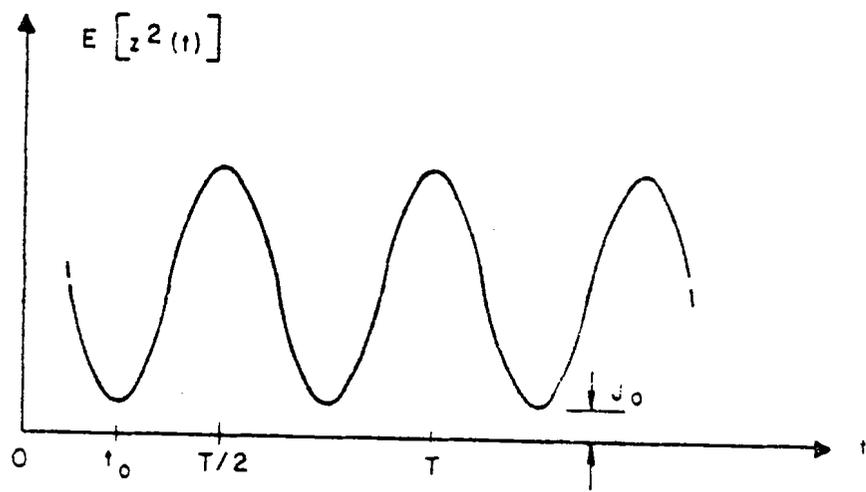


Figure 3.3. Variation in the mean square value of the timing wave process [9]

If the relative time difference between t_1 and t_0 is small, we could take advantage of the smaller variance by implementing an device which observes axis-crossings when the axis is set at a level difference from zero. Hence, if we observe crossings of the axis $E[z(t_1)]$, then the rms jitter is characterized by the smaller value

$$\left(\frac{\Delta t}{T}\right)_{\text{rms}} = \frac{1}{4\pi|U_1|} \{V_0 - 2|V_2|\}^{1/2} \quad (3.26)$$

3.5. Prefiltering to minimize jitter

A careful examination of Equations (3.8) and (3.11) indicates that certain symmetries in $G(f)$ and $H(f)$ will tend to reduce jitter. In particular, it is seen that if $G(f)$ has even symmetry about $1/(2T)$ and is band limited to the interval $1/(4T) < |f| < 3/(4T)$, and if $H(f)$ also has even symmetry about $1/T$, then $\text{Var } z(t)$ vanishes and we have error-free timing recovery. This suggests that bandwidths greater than 50 percent in excess of the Nyquist bandwidth are unnecessary. The symmetry condition is approached by proper design of the prefilter in the timing path.

In the usual situation the received PAM signal will be such that the prefilter will be required to emphasize signal components in the region above $1/(2T)$ and attenuate components below $1/(2T)$. This operation has the additional advantage of rejecting low-frequency additive noise which would otherwise enter the timing path.

The fact that all zero crossings occur at the same point under the symmetry conditions on $G(f)$ and $H(f)$ suggests a certain independence of the statistical properties of the data sequence. This is indeed true and it is shown that under the symmetry conditions stated, the timing wave $z(t)$ has zero crossings at all odd multiples of $T/4$ for an arbitrary data sequence $\{a_k\}$.

Chapter 4

Analysis of a self-adjusting prefilter

From the previous chapter, we see that a overall prefiltering response $g(t)$ which is symmetric about $1/(2T)$ and band-limited to the interval $1/(4T) < |f| < 3/(4T)$ will tend to reduce jitter and that with certain symmetry in the band-pass filter $H(f)$, we can have jitter-free timing recovery.

4.1. Properties of the prefilter characteristics

Consider a baseband signal $S(f)$ which is symmetric about zero and band-limited to the interval $-1/(4T) < |f| < 1/(4T)$ as shown in Figure 4.1a. The band-pass signal $g(t)$ is given by

$$g(t) = S(t) \cos\left(2\pi \frac{1}{2T} t\right) \quad (4.1)$$

The Fourier transform $G(f)$ of $g(f)$ is symmetric and band-limited to the interval $1/(4T) < |f| < 3/(4T)$, as indicated in Figure 4.1b. It is clear that $g(t)$ represents the desired overall impulse response (Figure 4.1e), which has a zero crossing at each odd multiple of $T/2$.

Implementing an accurate analog filter to achieve these characteristics is not an easy job, especially because the timing wave is strongly dependent on the pulse

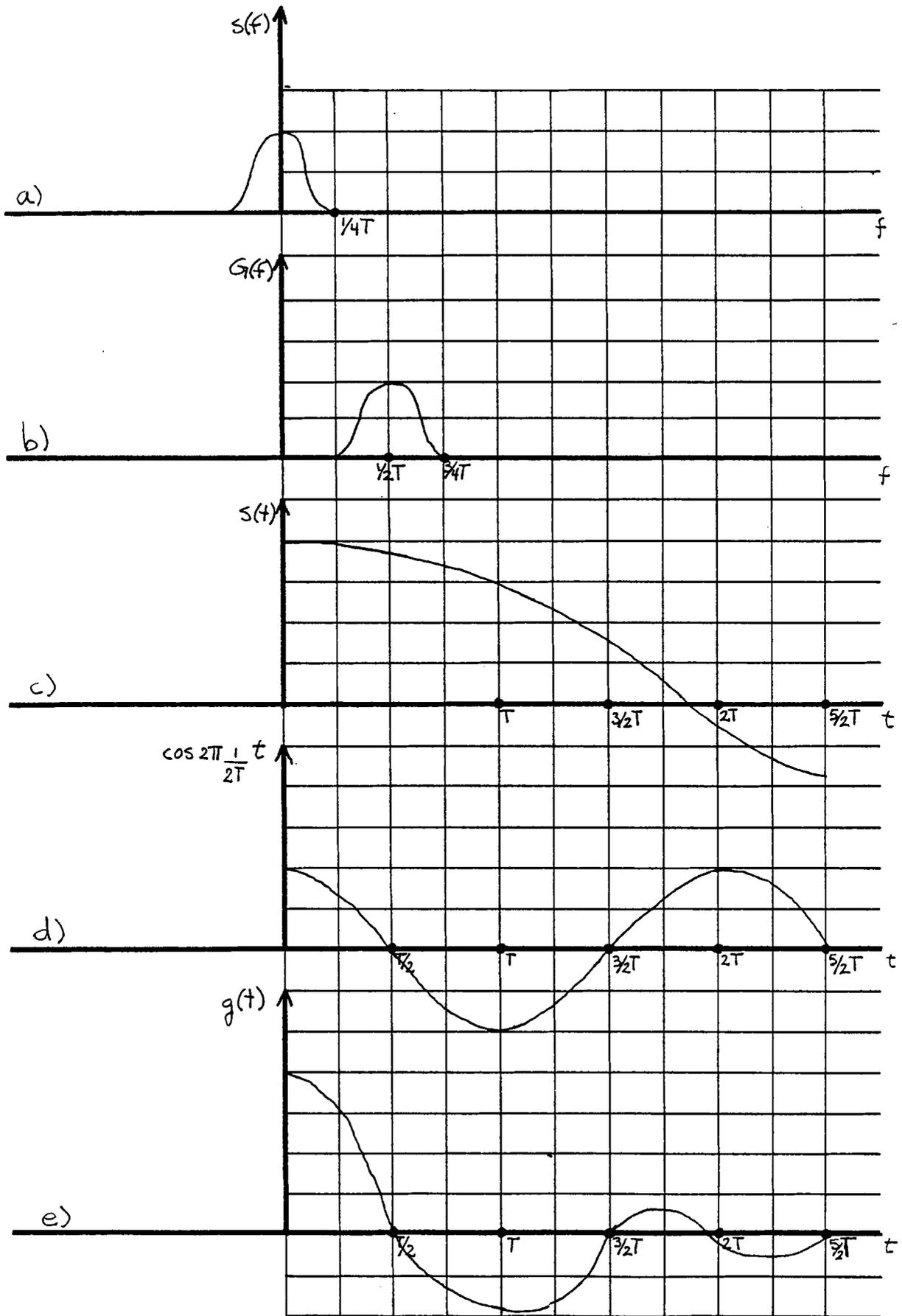


Figure 4.1. Graphical representation for the overall response in the time and frequency domains.

shape and consequently on the (slowly time-varying) unknown channel characteristics. Adaptive equalizers can be used to compensate and track slow changes in such channels [17]. An analog realization of the equalizer is, usually, a low-pass filter (to band-limit the noise) followed by an equalizer and a sampler. However, because of the delay-line structure, the sampler can just as well precede the equalizer. The equalizer then becomes a discrete time filter, which is usually to be preferred from an implementational point of view. So, a low-pass filter, a sampler, and a tapped-delay-line filter (with an automatically adjustable gain at each tap) will be used instead of a time-invariant prefilter to improve performance.

4.2. The adaptive prefilter implementation

The performance of the timing circuit depends on a proper choice of

- 1) the prefilter tap spacing and
- 2) the gain coefficients of the equalizer.

4.2.1. The prefilter tap spacing

In this section, first we examine the sampling rate required for the input signal of the prefilter to ensure complete reconstruction of the signal without aliasing.

In the previously mentioned timing recovery circuit several signal processing operations have taken place: the timing wave is prefiltered, squared and filtered again. For these operations analog signal processing has been used, thus the signal needed to derive timing information is continuous in both time and amplitude. In this thesis, we are confronted with a somewhat different situation. The suggested adaptive prefilter is implemented with digital signal processing. The analog timing recovery

scheme could then be “digitized” and would still perform in a functionally equivalent way. The timing wave is sampled at the input of the timing circuit to be available only at discrete time intervals for further processing. To combat the effect of aliasing, the sampling interval is typically selected such that the bandwidth occupied by the timing wave satisfies the Nyquist condition, i.e., $T_s < 1/(2f_{\max})$ where T_s is the sampling interval and f_{\max} is the maximum frequency occupied by the signal. This maximum frequency in the output of the timing recovery circuit ($f_{2\max}$) is double the one in the input of this circuit ($f_{1\max}$) because of the square law device. This means that band sampling will not permit an exact signal reconstruction by interpolation techniques, except in the case of a baseband signal which is strictly limited to the Nyquist frequency (i.e., half the sampling rate), which is not the case here. Sampling must be performed at a high enough rate to allow a complete reconstruction. In general, the desired sampling interval can be represented as

$$T_p = \xi T \quad (4.2)$$

where

$$\xi \triangleq \frac{K}{L} \leq 1 \quad K, L \text{ are integers,}$$

and T is the band sampling interval, which satisfies the Nyquist interval, i.e.,

$$T < \frac{1}{2f_{1\max}}$$

and T_p also satisfies it, i.e.,

$$T_p < \frac{1}{2f_{2\max}}$$

but $f_{2\max} = 2f_{1\max}$, so

$$T_p < \frac{1}{2 \cdot 2f_{1\max}}$$

or

$$T_p < \frac{T}{2} \quad (4.3)$$

Note that the output impulse response of the prefilter, which ideally has zero-crossings at each odd multiple of $T/2$, can be represented by a transfer function with zero-crossings at nT for $n = 0, \pm 1, \pm 2, \dots$ followed by a shifter with a constant group delay equal to $T/2$. This means that a sampling interval equal to $T/3$ is not adequate because it does not give samples at every T interval. However, choosing the sampling interval as

$$T_p = \frac{T}{4} \quad (4.5)$$

will satisfy the Nyquist interval and will also give the desired samples, i.e., it is the suitable choice as a sampling interval.

On the other hand, the time span between one tap of the prefilter equalizer and the next has to be equal to or a fraction of the sampling intervals of the input signal. Otherwise, some of the signal information will be lost. So, if the delay line taps are spaced at the interval T_p (the sampling interval) then the tap spacing is smaller than T (the band sampling interval). This type of equalizer is called a fractional tap spacing equalizer (FSE) [18, 19]. The timing circuit and the FSE prefilter are shown in Figures 4.2 and 4.3 respectively. One important property of this type of equalizer is the insensitivity of its performance to the choice of its input signal's sampler phase. This distinction between the conventional T -spaced and FSE for an appropriately (i.e., not exceeding twice the Nyquist bandwidth) band-limited input can be heuristically explained as follows.

(i) Symbol-rate sampling at the input to a T equalizer causes spectral overlap or aliasing. When the phases of the overlapping components match they add con-

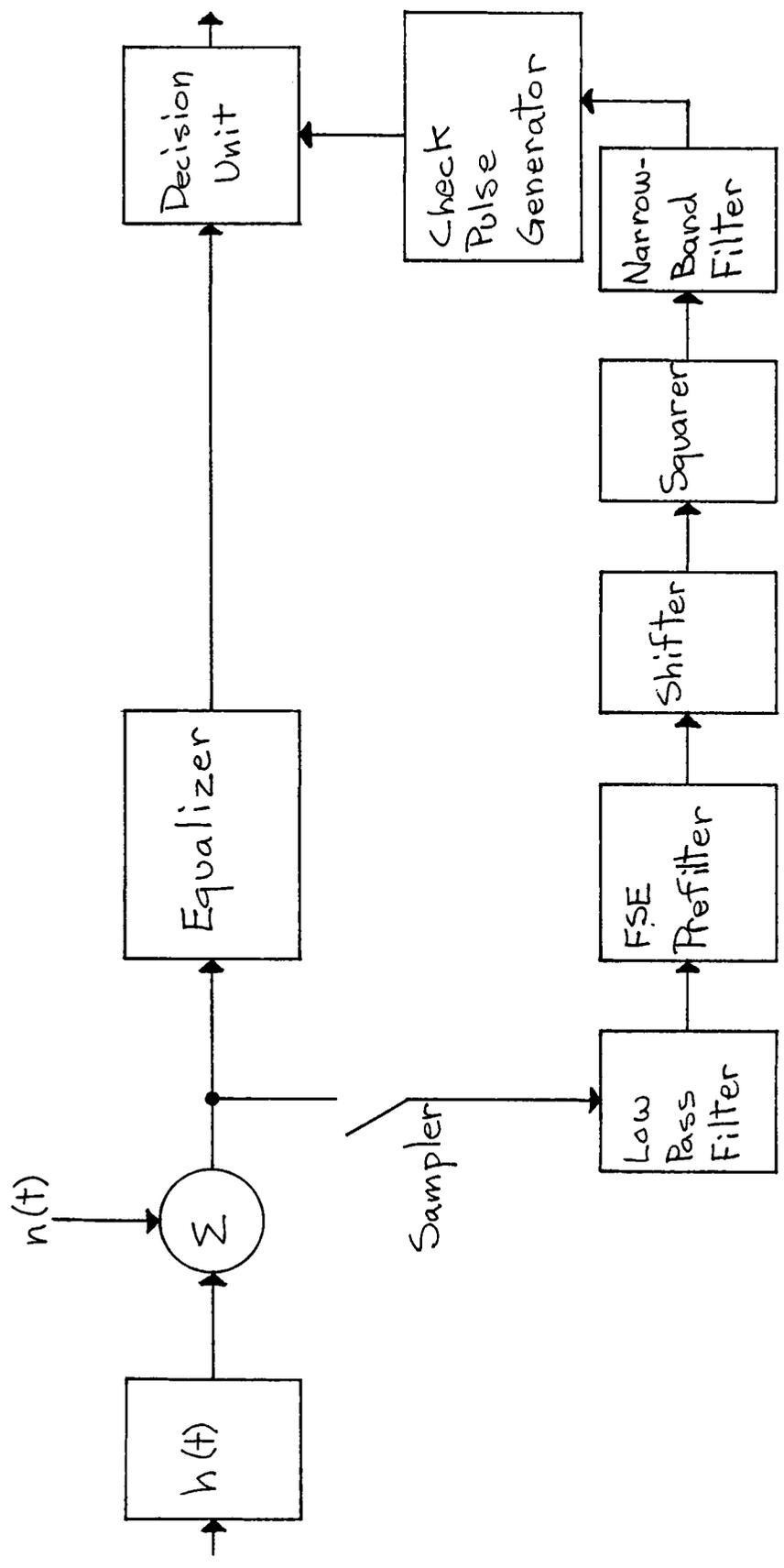


Figure 4.2. A model for the timing circuit including the fractional tap-delay prefilter

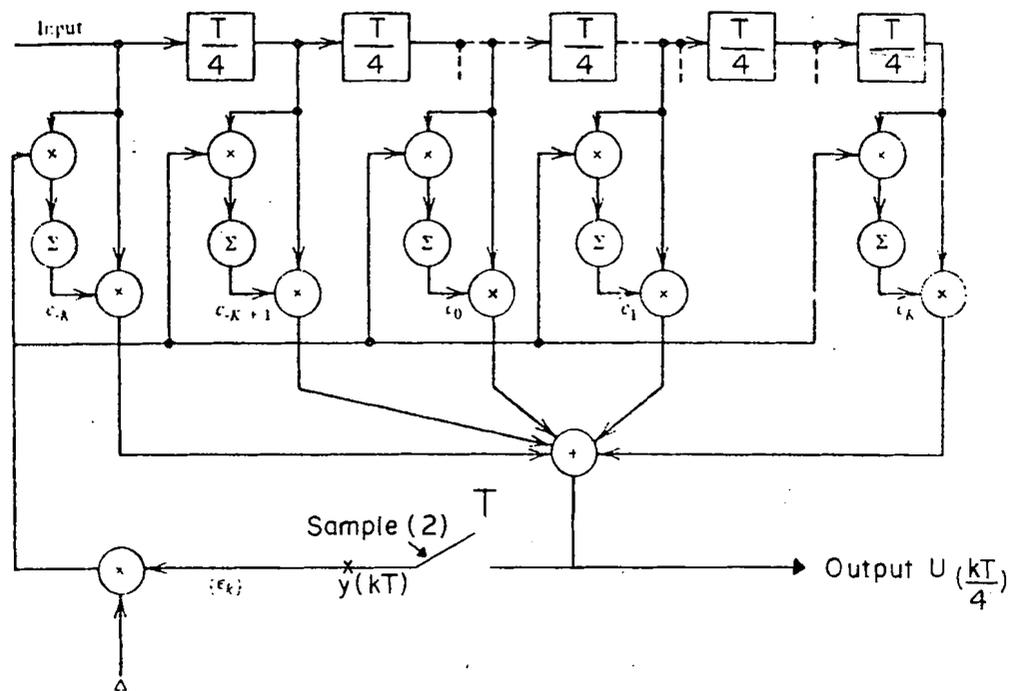


Figure 4.3. Fractional tap delay-line prefilter model

structively, when the phases are 180° apart they add destructively, which results in the cancellation or reduction of amplitude as shown in Figure 4.4. Variation in the sampler phase or timing instant corresponds to a variable slope being added to the signal spectrum. Thus, changes in the sampler phase strongly influence the effect of aliasing, i.e., they influence the amplitude and delay characteristics in the spectral overlap region of the sampled equalizer output.

In contrast, there is no spectral overlap at the input to an FSE because the longer period (in the frequency domain) of its transfer function allows the equalizer to control the phases of the aliasing roll-off components independently of one another.

(ii) The T-spacing equalizer cannot suppress noise at frequencies outside the Nyquist band, $|f| \leq 1/(2T)$, so the burden of rejecting such noise components is placed entirely on the filters preceding the equalizer. The FSE, by virtue of its sampling rate, can synthesize the best combination of the characteristics of an adaptive matched filter and a T-spaced equalizer, within the constraints of its length and delay. An FSE can effectively compensate for more severe delay distortion and deal with amplitude distortion with less noise enhancement than a T-spaced equalizer.

The cost of fractional tap spacing would then primarily be an L times as high analog/digital (A/D) conversion rate (Equation (4.2)), L times as many memory stages for the equalizer delay line, and L/K times as many multiplications for the same length of the equalizer delay line in terms of total delay.

4.2.2. The tap gain coefficients of the equalizer

The prefilter equalizer is shown in Figure 4.3, in which the spacing between the

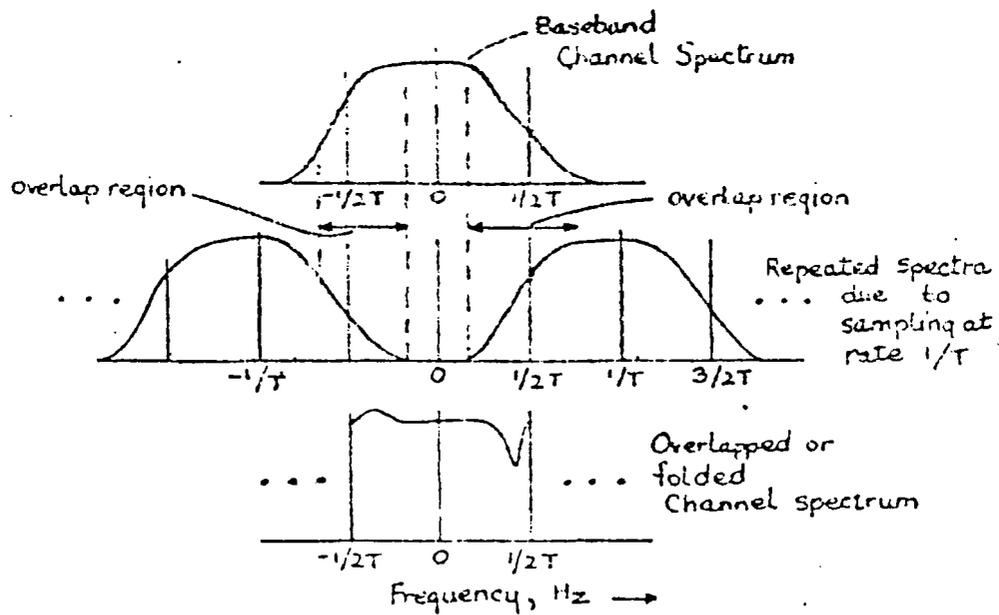


Figure 4.4. Spectral overlap at the input to a T equalizer

taps is $T/4$. The signal at its input is given by

$$x\left(\frac{qT}{4} + \tau_s\right) = \sum_i a_i h\left(\frac{qT}{4} - iT - \tau_s\right) + n\left(\frac{qT}{4} + \tau_s\right) \quad (4.6)$$

where

$\{a_n\}$ is the sequence of data symbols transmitted

$h(t)$ denotes the elementary pulse waveform

τ_s is the constant time offset of the sampler with respect to the data source clock, and

$n(t)$ represents stationary noise, characterized by the autocorrelation function

$$R_n(\tau) = E[n(t)n(t + \zeta)] \approx 2N_0\delta(\tau) \quad (4.7)$$

For simplicity, the assumption is made that the noise is white over the entire frequency range of $h(t)$, with constant spectral power density N_0 . Outside this range the noise is band-limited by the low-pass filter.

Then, equalizer output

$$U\left(q\frac{T}{4} + \tau_s\right) = \sum_j c_j x\left[q\frac{T}{4} - j\frac{T}{4} + \tau_s\right] \quad -N \leq j \leq N \quad (4.8)$$

is observed, where $2N + 1$ is the number of taps.

According to Figure 4.3, at sampling instant $kT + \tau_s$, the output signal of the equalizer is given by

$$\hat{y}(kT + \tau_s) = \sum_j c_j x\left(kT - j\frac{T}{4} + \tau_s\right) \quad (4.9)$$

Let $C : \{c_j, -N \leq j \leq N\}$ be the vector of the adjustable tap gains. The equalizer output therewith becomes

$$\hat{y}_k = C^T \cdot X_{k-j} \quad (4.10)$$

and

$$X_{k-j} : \left\{ x_{\text{in}} = x \left[kT + \tau_s - j \frac{T}{4} \right] \right\} \quad -N \leq j \leq N \quad (4.11)$$

If \hat{y}_k is defined as the estimate of the desired value y_k , and from the transfer function of the prefilter (Section 4.2.1), the desired value is equal to zero at each time interval T . Then the mean square error (MSE) between the desired value and the estimated one is

$$J = E |e|^2 = E |y_k - \hat{y}_k|^2 = E |\hat{y}_k|^2 \quad (4.12)$$

Using Equations (4.1) and (4.9) in (4.12) yields

$$\begin{aligned} E |\hat{y}_k|^2 &= E \left| \sum_j c_j x \left(kT - j \frac{T}{4} - \tau_s \right) \right|^2 \\ &= \sum_j \sum_i c_j c_i (\psi_{ji} + \phi_{ji}) \end{aligned} \quad (4.13)$$

where

$$\begin{aligned} R_{nm} &= E[a_n a_m] \\ \psi_{ji} &= \sum_n \sum_m R_{nm} h \left(kT - nT - j \frac{T}{4} + \tau_s \right) h \left(kT - mT - i \frac{T}{4} + \tau_s \right) \\ \phi_{ji} &= E n \left(j \frac{T}{4} \right) n \left(i \frac{T}{4} \right) \end{aligned}$$

In matrix notation, Equation (4.13) becomes

$$E |\hat{y}_k|^2 = C^T A C \quad (4.14)$$

where C is a column vector of the $2N + 1$ optimum tap coefficients and A is the $(2N + 1)$ by $(2N + 1)$ covariance matrix of input samples $x(kT)$, $-N \leq k \leq N$ with elements given by

$$A_{ij} = E \left[x \left(kT - i \frac{T}{4} + \tau_s \right) x \left(kT - j \frac{T}{4} + \tau_s \right) \right] \quad (4.15)$$

Equation (4.14) represents the MSE as a function of the tap coefficients $\{c_j\}$. One approach to minimizing this function with respect to $\{c_j\}$ is invoking the orthogonality principle in the mean square estimation, selecting the coefficients $\{c_j\}$ to render the error e_k orthogonal to the signal sequence $\{x_{k-\ell}\}$, for $-N \leq \ell \leq N$, i.e.,

$$\langle y_k - \hat{y}_k, x_{k-\ell} \rangle = 0 \quad (4.16)$$

But, for this specific case, $y_k = 0$, which means that we are projecting a null vector. Thus the orthogonality principle is not applicable.

Before discussing the proper approach to minimizing the MSE, we have to study the covariance matrix A more closely. From Equation (4.3),

$$x \left(q \frac{T}{4} \right) = \sum_i a_i h \left(q \frac{T}{4} - iT + \tau_s \right) + n \left(q \frac{T}{4} + \tau_s \right)$$

Applying this equation to Equation (4.15), which represents the elements of the matrix A , yields

$$A_{ij} = \sum_p \sum_r E[a_p a_r] h\left(kT - pT - i\frac{T}{4} + \tau_s\right) h\left(kT - rT - j\frac{T}{4} + \tau_s\right) + E\left[n\left(kT - i\frac{T}{4} + \tau_s\right) n\left(kT - j\frac{T}{4} + \tau_s\right)\right] \quad (4.17)$$

By defining $R_{aa}(\cdot)$ as the data source autocovariance function, and assuming that the data source is stationary with zero mean, we find that

$$R_{aa}(p - r) \triangleq E[a_p a_r]$$

Defining

$$m \triangleq p - r \quad \text{and} \quad n = k - r$$

$$\phi_{nn}\left[(i - j)\frac{T}{4}\right] \triangleq E\left[n\left(kT - i\frac{T}{4} + \tau_s\right) n\left(kT - j\frac{T}{4} + \tau_s\right)\right]$$

leads to

$$A_{ij} = \sum_m \sum_n R(m) h\left((n - m)T - i\frac{T}{4} + \tau_s\right) h\left(nT - j\frac{T}{4} + \tau_s\right) + \phi_{nn}\left[(i - j)\frac{T}{4}\right] \quad (4.18)$$

For a white data source, white noise with powers σ_a^2 and σ_n^2 respectively, we get

$$A_{ij} = \sigma_a^2 \sum_n h\left(nT - i\frac{T}{4} + \tau_s\right) h\left(nT - j\frac{T}{4} + \tau_s\right) + \sigma_n^2 \delta_{i,j} \quad (4.19)$$

The A matrix has several important properties. First, even though it is symmetric like the autocorrelation matrix for a T equalizer, it is unlike them in that its (i, j) th element is not a function of $i - j$, i.e., it is not a Toeplitz matrix.

From Equation (4.19), the A matrix has the form

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ a_{31} & a_{32} & a_{33} & a_{34} & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ a_{41} & a_{42} & a_{43} & a_{44} & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ a_{51} & \cdot & \cdot & \cdot & a_{11} & a_{12} & a_{13} & a_{14} & \cdots & \cdot \\ a_{61} & \cdot & \cdot & \cdot & a_{21} & a_{22} & a_{23} & a_{24} & \cdots & \cdot \\ a_{71} & \cdot & \cdot & \cdot & a_{31} & a_{32} & a_{33} & a_{34} & \cdots & \cdot \\ a_{81} & \cdot & \cdot & \cdot & a_{41} & a_{42} & a_{43} & a_{44} & \cdots & \cdot \\ \vdots & \ddots & \vdots \\ a_{w1} & a_{w2} & a_{w3} & a_{w4} & \cdot & \cdot & \cdot & \cdot & \cdots & a_{ww} \end{bmatrix} \quad (4.20)$$

where $w = 2N + 1$.

To explicitly see the non-Toeplitz nature of the A matrix we can rewrite Equation (4.19) in the frequency domain as

$$\begin{aligned} \frac{1}{\sigma_a^2} (A_{i-j} - \sigma^2 \delta_{i,j}) &= \int_{-\pi/T}^{\pi/T} \left\{ \sum_n H \left(\omega + n \frac{2\pi}{T} \right) \exp \left[-j \left(\omega + n \frac{2\pi}{T} \right) i \frac{T}{4} \right] \right\} \\ &\quad \left\{ \sum_m H \left(\omega + m \frac{2\pi}{T} \right) \exp \left[-j \left(\omega + m \frac{2\pi}{T} \right) j \frac{T}{4} \right] \right\} \frac{d\omega}{2\pi} \\ &= \int_{-\pi/T}^{\pi/T} e^{-j(i-j)T/4} \left\{ \sum_n H \left(\omega + n \frac{2\pi}{T} \right) \exp \left[-j \frac{ni2\pi T/4}{T} \right] \right\} \\ &\quad \left\{ \sum_m H \left(\omega + n \frac{2\pi}{T} \right) \exp \left[-j \frac{mj2 - piT/4}{T} \right] \right\} \frac{d\omega}{2\pi} \quad (4.21) \end{aligned}$$

and for systems with nonzero excess bandwidth the bracketed terms depend on i and j individually, rather than on $i - j$.

Second, since the matrix is an autocorrelation matrix, it is positive semidefinite. In most but not all applications, it can be assumed that this matrix is positive definite, and hence nonsingular. Cases in which this is not true will be discussed in Section 4.4, but as long as it is true then the autocorrelation matrix has positive real eigenvalues.

4.3. The resulting minimum mean square error (MMSE)

The MSE is defined in Equation (4.14) as follows:

$$E|\hat{y}_k|^2 = C^T A C$$

This quadratic form is minimized with respect to $\{c_j\}$ by the trivial solution $\{c_j\} = 0$. Thus, a restriction $C \neq 0$ must be imposed on the coefficients to prevent this situation from occurring. One way this can be done is to normalize the tap coefficients so that

$$C^T C = 1. \tag{4.22}$$

so as to avoid the trivial solution. Two different forms of this constraint lead to the same MMSE solution as demonstrated by the theorem given in Appendix B, which implies that [20]

$$\text{Minimizing } \frac{C^T A C}{C^T C} \text{ for any } C \neq 0$$

is equivalent to (4.23)

$$\text{Minimizing } C^T A C \text{ subject to } C^T C = 1$$

From the previous section, the matrix A is a special case of a Hermetian matrix in a sense, as it is a real symmetric matrix. To minimize this quadratic form with A

as a Hermetian matrix (Equation (4.23)) we need to apply the theorem explained in Appendix C, which proves that [20]

$$\lambda_{\min} = \frac{\min C^T A C}{C^T C} \quad \text{for any } C \neq 0 \quad (4.20)$$

which is equivalent to

$$\lambda_{\min} = \min_{C^T C=1} C^T A C$$

This means that the resulting minimum mean square error is

$$\mathcal{E}_{\min} = \lambda_{\min} \quad (4.21)$$

where λ_{\min} is the minimum eigenvalue for the matrix A . Then, the optimum tap coefficients vector (C_{opt}) is the eigenvector corresponding to this eigenvalue (λ_{\min}).

Note that, analytically, the normalized tap coefficients are preferable because they tend to preserve the signal's energy, which means that the output signal carries the same amount of energy as the input signal (at least for white inputs). So, we choose the normalization for the tap-spacing coefficients as a constraint.

4.4. Uniqueness of solution for the fractionally-spaced prefilter as the noise vanishes

Referring to Equation (4.19), the ij th element of the channel correlation matrix can be written as

$$A_{ij} = \sigma_a^2 \sum_{n=-\infty}^{\infty} h(nT - iTp)h(nT - jTp) + \sigma_n^2 \delta_{i,j} \quad (4.26)$$

As can be seen from this equation, the matrix A is the sum of two matrices. The channel-dependent component of A is always positive semidefinite. Since the other component, $\sigma^2 I$, is positive definite for nonzero σ^2 , A will also be positive definite, and we can conclude that when there is noise present, the optimum tap setting is unique. Now, consider the situation as the noise becomes vanishingly small, i.e., A is positive semidefinite. This means that the matrix has one or more zero eigenvalues (λ_{\min}) and, clearly, the corresponding eigenvector is not unique. In this case, from Equation (4.25), the resulting minimum mean square error is optimal because it is zero. But the corresponding solution for the optimal tap coefficient vector C_{opt} is finite but not unique due to the eigenvectors. The optimum tap setting will be unique if, and only if, A is nonsingular. A sufficient condition for A to be nonsingular is the nonvanishing of the quadratic form $C^T A C$, for any nonzero vector C , (which implies a non-zero mean square error). In the following discussion, which considers the quadratic form in detail, the work by Gitlin *et. al.* [21] on T or $T/2$ spaced equalizers is extended to our case, a $T/4$ spaced prefilter. We start by rewriting the quadratic form as

$$\begin{aligned}
C^T A C &= \sum_{m=-N}^N \sum_{n=-N}^N c_m A_{mn} c_n \\
&= R_{aa}^2(0) \sum_{m=-N}^N \sum_{n=-N}^N c_m c_n \sum_{\ell=-\infty}^{\infty} h\left(\ell T - n \frac{T}{4}\right) h\left(\ell T - m \frac{T}{4}\right) \\
&= R_{aa}^2(0) \sum_{\ell=-\infty}^{\infty} \left[\sum_{m=-N}^N c_m h\left(\ell T - m \frac{T}{4}\right) \right]^2 \geq 0
\end{aligned} \tag{4.27}$$

$C^T AC$ can vanish only if

$$\sum_m c_m h\left(\ell T - m\frac{T}{4}\right) = 0 \quad \ell = 0, \pm 1, \pm 2, \dots \quad (4.28)$$

If we define the periodic Fourier transform

$$c_{T/4}(\omega) = \sum_m c_m e^{j\omega m T/4} \quad |\omega| \leq \frac{4\pi}{T} \quad (4.29)$$

then we can proceed further by noting that

$$\begin{aligned} \sum_m c_m h\left(\ell T - m\frac{T}{4}\right) &= \sum_m c_m \int_{-\infty}^{\infty} h(\omega) e^{-j\omega(\ell T - m T/4)} \frac{d\omega}{2\pi} \\ &= \int_{-\infty}^{\infty} \left[\sum_m c_m e^{j\omega m T/4} \right] h(\omega) e^{-j\omega \ell T} \frac{d\omega}{2\pi} \\ &= \int_{-\infty}^{\infty} c_{T/4}(\omega) h(\omega) e^{-j\omega \ell T} \frac{d\omega}{2\pi} \\ &= \sum_k \int_{(2k-1)\pi/T}^{(2k+1)\pi/T} c_{T/4}(\omega) h(\omega) e^{-j\omega \ell T} \frac{d\omega}{2\pi} \\ &= \int_{-\pi/4}^{\pi/4} \left[\sum_k c_{T/4}\left(\omega + \frac{k2\pi}{T}\right) h\left(\omega + \frac{k2\pi}{T}\right) \right] e^{-j\omega \ell T} \frac{d\omega}{2\pi} \end{aligned} \quad (4.30)$$

The right-hand side of Equation (4.30) is recognized as the sample, at $t = \ell T$, of a function whose transform, $y_{\text{eq}}(\omega)$, is contained in the brackets. If Equation (4.30) is to be zero for every value of ℓ , then it must be that the Fourier transform inside the integral vanishes completely, i.e.,

$$y_{\text{eq}}(\omega) = \sum_k c_{T/4}\left(\omega + \frac{k2\pi}{T}\right) h\left(\omega + \frac{k2\pi}{T}\right) = 0 \quad |\omega| \leq \frac{\pi}{4} \quad (4.31)$$

In Figure 4.5, we show the situation when there is no excess bandwidth. Since the sum in Equation (4.31) reduces to one term, the only way for $y_{\text{eq}}(\omega) \equiv 0$ is for either $h(\omega)$ or $c_{T/4}(\omega)$ to be zero. Since this implies that $c_{T/4}(\omega) \equiv 0$, which would violate the nonzero requirement on C , we must conclude that for this case, A is positive definite. A similar sketch for the less than 100 percent excess bandwidth case is shown in Figure 4.5b, which it is noted that only the $k = 0, \pm 1$ terms contribute to the sum.

However, since in the non-rolloff region, $|\omega| \leq (1 - \alpha)\pi/T$, only the $k = 0$ term influences the sum. For channels that do not vanish over the entire non-rolloff region, it is clear that for $y_{\text{eq}}(\omega)$ to vanish it is required that $c_{T/4}(\omega)$ vanish at least over the entire non-rolloff region. Since $c_{T/4}(\omega)$ is a finite term Fourier series, it cannot vanish over an interval without vanishing everywhere, which in turn would again make $C = 0$. Note that if the channel vanished over a portion of the non-rolloff region, then since $y_{\text{eq}}(\omega)$ is a finite term Fourier series, its energy could not be totally concentrated in the region where there was no channel energy. Thus, the solution still would be unique. It is worth noting that in the extreme case of 100 percent excess bandwidth, $y_{\text{eq}}(\omega)$ can vanish. For example, in Figure 4.5c we sketch the situation for a constant $h(\omega)$, and with $c_{T/4}(\omega) = \cos(\omega T/4)$ it is apparent that $Z_{\text{eq}}(\omega) = 0$. Thus for a finite-length fractional spaced prefilter with less than 100 percent excess bandwidth, we can conclude that even as the noise becomes vanishingly small, the A matrix is nonsingular and there is a unique optimum tap setting.

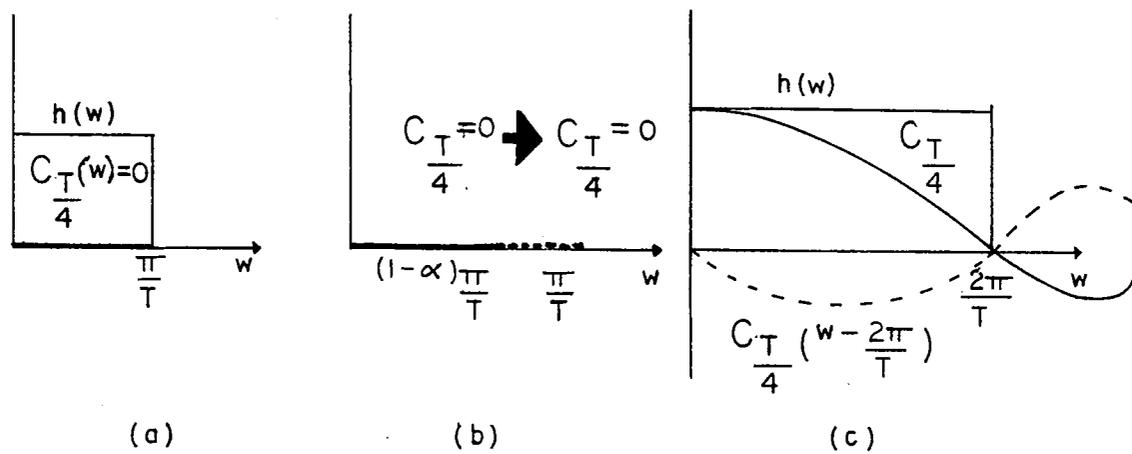


Figure 4.5. Conditions for solution uniqueness for a fractionally-spaced prefilter

Chapter 5

An iterative algorithm for computing the tap gain coefficients

The previous results indicate that solving for the optimal prefilter coefficients C_{opt} involves calculating the eigenvector for the minimum eigenvalue λ_{min} of the autocovariance matrix A . Alternatively, an iterative procedure may be used to determine C_{opt} . In the related equalizer problem, one of the most efficient methods is the use of a stochastic gradient (SG) algorithm. This adaptive algorithm is also sometimes called the LMS (least-mean-square) adaptive transversal filter.

In the following section, an outline of the MMSE gradient algorithm will be reviewed. Although this algorithm is of little practical interest in itself, it will lead directly to the widely used stochastic gradient (SG) algorithm. Furthermore, an understanding of the properties of the gradient algorithm will be very helpful in understanding the SG algorithm.

5.1. The MMSE gradient algorithm

The approach in this algorithm is to define a sequence of coefficient vectors which is guaranteed to converge to the optimum coefficient vector. As a starting point for

the derivation, one defines a vector $\{x_n\}$ as the tap-input vector whose elements consist of the $(2N + 1)$ tap inputs of the filter:

$$x_n = [x(nT + Tp), \dots, x(nT), \dots, x(nT - Tp)] \quad -N \leq n \leq N \quad (5.1)$$

x_n is assumed to be a discrete time random process with a known autocorrelation matrix* defined by Equation (4.19).

The output mean square error (MSE) given by Equation (4.14) is a quadratic form in the coefficient vector which has a unique global minimum. The approach is to adjust the weights iteratively to minimize the MSE by descending along the performance surface towards the minimum. Since the algorithm is iterative in nature, a notation for the coefficient vector which reflects this is needed. So, call the j th iteration of the coefficient vector C_j , which corresponds to some point on the quadratic MSE surface in the $(2N + 1)$ -dimensional space of coefficients. The gradient (error) vector G_j , having the $2N + 1$ gradient components $\left. \frac{1}{2} \frac{\partial J}{\partial C} \right|_{C=C_j}$, is then computed at this point on the MSE surface. Given the present coefficient vector C_j , by subtracting off a term proportional to the error gradient, the resultant tap vector should be closer to C_{opt} . This is because the gradient of the error is a vector in the direction of maximum increase of the error. Moving a short distance in the opposite (negative) direction to that of the gradient vector should therefore reduce the error. On the other hand, moving too far in this direction might actually overshoot the

* Although this is an unrealistic assumption, it is useful to examine the solution for this case as a motivation for the SG algorithm which is oriented toward the case of unknown statistics.

minimum and result in instability. Thus succeeding values of the coefficient vector are obtained according to the relation [22]

$$C_{j+1} = C_j - \Delta G_j \quad j = 0, 1, 2, \dots \quad (5.3)$$

where Δ is a small adaptation constant or step size which controls the size of the change in C_j at each update. Referring back to Equations (4.9) and (4.12), we see that the gradient vector is

$$\begin{aligned} G_j &= \left. \frac{\partial J}{2\partial C} \right|_{C=C_j} = \frac{\partial}{2\partial C} E[e_j^2] \\ &= E \left[e_j \frac{\partial}{\partial C_k} e_j \right] \\ &= -E[e_j x_j] \end{aligned} \quad (5.4)$$

where the vector C_j represents the set of coefficients at time j and $e_j = -y_j$ is the error signal at the j th iteration.

From Equation (4.14), the gradient vector can be expressed as

$$G_j = AC_j \quad (5.5)$$

Note that the division by two in the gradient vector is included to avoid a factor of two in the subsequent adaptation algorithm.

The gradient algorithm is explicitly

$$\begin{aligned} C_{j+1} &= C_j - \Delta AC_j \\ &= (I - \Delta A)C_j \end{aligned} \quad (5.6)$$

or,

$$C_{j+1} = C_j + \Delta E[e_j x_j] \quad (5.7)$$

where I is the identity matrix.

The basic difficulty in determining the optimum tap weights is the lack of knowledge of the gradient vector G_j , which depends on the covariance matrix A . In turn, these quantities depend on the channel characteristics and on the covariance of the information sequence and the additive noise, all of which may be unknown at the receiver. Thus the MMSE gradient solution is not applicable to the practical situation in which the channel characteristics are unknown.

5.2. The LMS stochastic gradient algorithm

The stochastic gradient (SG) algorithm overcomes a problem of the MMSE solution described in Section 5.1, namely, that the gradient vector is usually unavailable since taking the expectation required knowledge of the ensemble statistics. The approach taken in this section is to circumvent this problem by, in effect, substituting a time average for the ensemble average. The troublesome part of Equation (5.7) is the expectation operator. The principle behind the SG algorithm is to ignore this operator. The quantity which is left, while random, has an expected value equal to the desired gradient. Thus, it is an unbiased estimate of the gradient. This “noisy” or “stochastic” gradient is substituted for the actual gradient in the algorithm of (5.7) resulting in the SG algorithm* [22]

$$\hat{C}_{j+1} = \hat{C}_j - \Delta \hat{G}_j \quad (5.8)$$

* “^” means estimated value

Consequently, an estimate of G_j is

$$\hat{G}_j = -e_j x_j \quad (5.9)$$

Incorporating (5.9) into (5.8) gives

$$\hat{C}_{j+1} = \hat{C}_j + \Delta e_j x_j \quad (5.10)$$

Note that a constraint was imposed on the tap coefficient vector in Equation (4.23), as was described in Section 4.3 of the previous chapter. The tap coefficients were normalized, suggesting a normalization of the coefficient updating vector (Equation (5.10)) at each iteration as well. The stochastic gradient algorithm just derived does not take this extra step into account. However, the behavior of the SG algorithm as is will be used as an approximation for purposes of studying the convergence behavior of the adaptive prefilter.

5.3. Convergence behavior of the SG transversal algorithm

The convergence of an adaptation algorithm can be measured in two ways. The first way is to determine analytically or empirically how the output error decreases with time. This is usually done by calculating the output mean square error (MSE). The second method is to determine analytically or empirically how the filter coefficients approach their optimum values with time. Since the coefficients are actually fluctuating, even asymptotically for a fixed step-size algorithm, convergence is usually measured in terms of the mean values of the filter coefficients. The two measures are of course closely related to one another [23].

5.3.1. Mean coefficient vector

Before the output MSE of the filter as a function of time is investigated, it is simple and instructive to examine the mean value trajectories of the coefficients. Recall from chapter 4 that, at the n th iteration the mean square error is defined as (Equation (4.14))

$$\langle e_n^2 \rangle \triangleq E |\hat{y}_n|^2 = \langle (C_n^T x_n)^2 \rangle = C_n^T A C_n$$

To proceed with the analysis, it is convenient to reformulate the SG algorithm as follows:

- (1) Eliminate the error signal from Equation (5.10)

$$\begin{aligned} \hat{C}_{n+1} &= \hat{C}_n - \Delta x_n x_n^T \hat{C}_n \\ &= [I - \Delta x_n x_n^T] \hat{C}_n \end{aligned} \quad (5.11)$$

- (2) Define an estimated coefficient error vector

$$\hat{P}_n = \hat{C}_n - C_{\text{opt}}$$

and use it to eliminate \hat{C}_n from the right hand side of (5.11)

$$\hat{C}_{n+1} = \hat{P}_n - \Delta x_n x_n^T \hat{C}_n + C_{\text{opt}} \quad (5.13)$$

$$\hat{P}_{n+1} = [I - \Delta x_n x_n^T] \hat{P}_n - \Delta x_n e_{n_{\text{opt}}} \quad (5.14)$$

where $e_{n_{\text{opt}}}$ is the instantaneous error when the taps are at their optimum setting. This is normally quite small and would be zero if perfect equalization were possible. Taking the expectation of both sides of Equation (5.14), and assuming $e_{n_{\text{opt}}} = 0$, we get

$$E \left[\hat{P}_{n+1} \right] = E \left[(I - \Delta x_n x_n^T) \hat{P}_n \right] \quad (5.15)$$

To facilitate this analysis it is therefore assumed that \hat{P}_n is approximately independent of the input sequence $\{X_n\}$. The analytical results based upon this assumption have proved that the mean value of the filter coefficients vector does converge, provided that the algorithm step size is small enough [24]. Equation (5.15) is therefore rewritten as

$$E \left[\hat{P}_{n+1} \right] \cong E [I - \Delta A] \hat{P}_n \quad (5.16)$$

(3) Represent the correlation matrix A in terms of its eigenvalues and associated eigenvectors, as shown by

$$Q^T A Q = \Lambda \quad (5.17)$$

where the diagonal matrix Λ consists of the eigenvalues of A and the columns of the unitary matrix Q are the associated eigenvectors. Thus, premultiplying both sides of Equation (5.16) by Q^T , we get

$$\begin{aligned} Q^T E \left[\hat{P}_{n+1} \right] &= Q^T (I - \Delta A) \hat{P}_n \\ &= Q^T \hat{P}_n - \Delta Q^T A \hat{P}_n \end{aligned} \quad (5.18)$$

Define the transformed coefficient error vector

$$V_n = Q^T \hat{P}_n \quad (5.19)$$

Also, using the property of the unitary matrix that $Q Q^T = I$, we may write

$$\begin{aligned} Q^T A \hat{P}_n &= Q^T A I \hat{P}_n \\ &= Q^T \Lambda Q Q^T \hat{P}_n \\ &= \Lambda V_n \end{aligned} \quad (5.20)$$

and $E[\hat{P}_{n+1}] = QE[Q^T \hat{P}_{n+1}]$. Accordingly, we may rewrite Equation (5.18) in the form

$$V_{n+1} = (I - \Delta\Lambda)V_n \quad (5.21)$$

This equation represents a system of uncoupled scalar-valued first-order difference equations, the k th one of which may be written as

$$v_k(n+1) = (1 - \Delta\lambda_k)v_k(n) \quad k = -N, \dots, N \quad (5.22)$$

and simply iterating this equation gives

$$v_k(n) = (1 - \Delta\lambda_k)^n v_k(0) \quad k = -N, \dots, N \quad (5.23)$$

where $v_k(0)$ is the initial value of the k th element of the transformed coefficient error vector. The numbers generated by this solution form a geometric series with ratio

$$r_k = 1 - \Delta\lambda_k \quad (5.24)$$

For stability or convergence of the SG algorithm the geometric ratio must have magnitude less than one for all k . This ensures that, regardless of the initial conditions, the transformed coefficient error vector v_n approaches zero as the number of iterations n approaches infinity and, correspondingly, that the coefficient vector C_n of the filter approaches the optimum value C_{opt} . Therefore, a necessary and sufficient condition for the stability of the algorithm in terms of the mean coefficient vectors is that the step size parameter Δ satisfy the condition

$$0 < |1 - \Delta\lambda_k| < 1 \quad k = -N, \dots, N \quad (5.25)$$

Since all the eigenvalues of the correlation matrix A are real and almost always positive, the algorithm is stable if and only if

$$0 < \Delta < \frac{2}{\lambda_{\max}} \quad (5.26)$$

where λ_{\max} is the largest eigenvalue of the matrix A . This determines the largest possible value of Δ , but of more interest is the Δ corresponding to the fastest convergence of the gradient algorithm.

5.3.2. The time constant

The time constant τ_k denotes the time required for the amplitude of the k th natural mode $v_k(n)$ to decay to $1/e$ of its initial value $v_k(0)$, where e is the base of natural logarithms [23], i.e.,

$$\tau_k = \exp \left[-\frac{1}{\tau_k} \right] \quad (5.27)$$

Hence, from Equations (5.25) and (5.27), we find that the k th time constant can be expressed as follows:

$$\tau_k = -\frac{1}{\ln(1 - \Delta\lambda_k)} \quad (5.28)$$

For small Δ , we may approximate τ_k as

$$\tau_k \approx \frac{1}{\Delta\lambda_k} \quad \Delta \ll 1 \quad (5.29)$$

The mean value of the tap vector can therefore be approximated as (after Equation (5.23)):

$$E \left[\hat{C}_n \right] \cong C_{\text{opt}} + Q(I - \Delta\Lambda)V_0 \quad (5.30)$$

The second term on the right hand side of the equation is the error term which decays to zero if Δ is small enough. In particular, the convergence of the mean coefficient vector is limited by the smallest eigenvalue λ_{\min} , which produces the largest time constant τ_{\max} . It therefore takes more iterations to make adjustments in C along the direction of the eigenvector corresponding to λ_{\min} than along any other direction in the $(2N + 1)$ -dimensional coefficient vector space. As will be explained in detail in the next section, there is a convenient relation between the input signal spectrum and the rate of convergence of $E(\hat{C}_n)$. The more uneven the signal spectrum is, the larger the ratio of largest to smallest eigenvalues of A is likely to be, thereby causing the SG transversal algorithm to converge more slowly according to (5.30).

5.3.3. Normalization of step size

The SG algorithm displays an undesirable dependence of speed of convergence on input signal power. This can be seen from Equation (5.7): an increase in the size of the input signal of a factor α is equivalent to an increase in the step size Δ of a factor α^2 (to $\Delta\alpha^2$). The same effect can be deduced from the eigenvalues of A . Thus, the speed of convergence of the SG algorithm is strongly affected by the size of the input signal. A serious consequence of this is that if the input signal grows too large then the adaptation algorithm becomes unstable. One way to ameliorate this problem is to add an automatic gain control circuit before sampling the data. Another frequently used solution is to normalize the step size of the algorithm. The sizes of the updates can be kept approximately equal if the update is normalized using an estimate of the input signal variance. In particular, consider the choice of step size

$$\Delta(n) = \frac{1}{\sigma^2(n)} = \frac{1}{(1 - \alpha)\sigma^2(n - 1) + x^2(n)} \quad (5.31)$$

where α is a constant close to zero and $\sigma^2(n)$ is an estimate of the input signal power at iteration n . So, as n tends to infinity,

$$E \left[\frac{1}{\Delta(n)} \right] \rightarrow \frac{E [x^2(n)]}{\alpha} \quad (5.32)$$

In this case α controls both the speed of convergence (averaging time) and the scaling factor of the variance estimate. In order to approximately ascertain the effect of this normalization on the convergence of the mean coefficient vector, the step size Δ can be replaced by $E[\Delta(n)] \approx 1/E[\sigma^2(n)]$ in Equations (5.29) and (5.30). This is a reasonable approximation as long as the constant α is small enough to “smooth out” the statistical fluctuations in the step size $\Delta(n)$, so that it can be regarded as virtually independent of the data samples. If the step size at time $n = 0$ is initialized at the asymptotic value α/ϕ_0 , where $\phi_0 = E[x^2(n)]$, then $E[\Delta(n)] \approx \alpha/\phi_0$ for all n , and the time constant (Equation (5.29)) becomes

$$\tau_j = \frac{\phi_0}{\alpha\lambda_j} = \frac{\sum_{j=-N}^N \lambda_j}{(2N+1)\alpha\lambda_j} \quad (5.33)$$

where the fact that

$$\sum_{j=-N}^N \lambda_j = \text{trace } A = (2N+1)\phi_0$$

has been used. The normal mode time constants are therefore proportional to the ratio λ_{av}/λ_j where $\lambda_{av} = \phi_0$ is the “average” eigenvalue of A . A change in the input signal variance therefore causes a much less dramatic change in convergence speed, and furthermore, the stability constraint Equation (5.26) is satisfied as long as

$$0 < \alpha < \frac{2\lambda_{av}}{\lambda_{\max}} \quad (5.34)$$

To apply the above stability constraint to a system which uses a fractionally spaced equalizer (which is the case for the adaptive prefilter discussed in this thesis), some of the terms must be appropriately interpreted. Results were proven mathematically indicating that for $T_p = T/2$ and infinitely long FSEs half the eigenvalues are zero. It was also shown that for practical (finite length) FSEs, half the eigenvalues are quite small (close to zero). The computer results, as shown in the next chapter, show that for $T_p = T/4$ only one quarter of the eigenvalues are effective in the sense that they contain around 95 percent of the energy. So, since three-quarters of the eigenvalues will be approximately zero, we have as a tight bound that

$$\lambda_{av} \cong \sum_{j=-N/4}^{N/4} \frac{2\lambda_j}{N} \quad (5.35)$$

5.3.4. Output mean square error

Having described the behavior of the mean coefficient vector, we now investigate the filter's output MSE. Ideally, the MMSE (\mathcal{E}_{\min}) is realized when the coefficient vector C_n approaches the optimum value C_{opt} , defined by the matrix form of the quadratic equations. As shown in Section 5.1, the gradient algorithm does realize this idealized condition as the number of iterations, n , approaches infinity. This algorithm has the capability to do this because it uses exact measurements of the gradient vector at each iteration. On the other hand, the SG (LMS) algorithm relies on a noisy estimate for the gradient vector, with the result that the coefficient vector C_n only approaches the optimum value C_{opt} after a large number of iterations and then fluctuates about C_{opt} . Consequently, use of the SG algorithm, after a large number of iterations, results in a mean squared error \mathcal{E}_{∞} which is greater than the

MMSE (\mathcal{E}_{\min}). The amount by which the actual value of \mathcal{E}_{∞} is greater than \mathcal{E}_{\min} is called the excess mean squared error. Recall from Equation (4.11) that

$$\text{MSE} \triangleq \mathcal{E} = C^T A C = (C - C_{\text{opt}})^T A (C - C_{\text{opt}}) - C_{\text{opt}}^T A C_{\text{opt}} + 2C^T A C_{\text{opt}} \quad (5.36)$$

Comparing the basic formula for the eigenvalues and eigenvectors of the matrix A

$$A x_i = \lambda_i x_i \quad (5.37)$$

with Equation (4.20)

$$C_{\text{opt}}^T A C_{\text{opt}} = \lambda_{\min}$$

leads to the homogeneous equation

$$A C_{\text{opt}} = \lambda_{\min} C_{\text{opt}} \quad (5.38)$$

Then, using Equation (5.38) to eliminate $A C_{\text{opt}}$ from Equation (5.35) gives

$$\mathcal{E}_{\min} = (C - C_{\text{opt}})^T A (C - C_{\text{opt}}) - \lambda_{\min} + 2\lambda_{\min} C^T C_{\text{opt}} \quad (5.39)$$

If we define

$$q_n \triangleq (C - C_{\text{opt}})^T A (C - C_{\text{opt}}) \quad (5.40)$$

to represent the excess mean squared error, then we have

$$\mathcal{E}_n = q_n - \mathcal{E}_{\min} + 2\mathcal{E}_{\min} C^T C_{\text{opt}} \quad (5.41)$$

Recall

$$\epsilon \triangleq (C - C_{\text{opt}}) \quad (5.42)$$

Then

$$\begin{aligned}\mathcal{E}_n &= q_n - \mathcal{E}_{\min} + 2\mathcal{E}_{\min}(1 + \epsilon^T C_{\text{opt}}) \\ &= q_n + \mathcal{E}_{\min}(1 + 2\epsilon^T C_{\text{opt}})\end{aligned}\quad (5.43)$$

Assuming that ϵ is vanishing error close to zero, Equation (5.43) can be approximated as

$$\mathcal{E}_n = q_n + \mathcal{E}_{\min} \quad (5.44)$$

As a compromise, Gitlin *et. al.* [25] select the step size such that the residual excess mean squared error (5.40) is an acceptable fraction of the minimum attainable steady-state error, i.e.,

$$q_{\infty} = \gamma \mathcal{E}_{\min} \quad (5.45)$$

where $0 \leq \gamma \leq 1$.

The resulting steady-state step size is

$$\begin{aligned}\Delta &= \frac{2\bar{\lambda}}{\lambda_{\max}} \cdot \frac{1}{(2N+1)\phi_0} \cdot \frac{q_{\infty}}{q_{\infty} + \mathcal{E}_{\text{opt}}} \\ &= \frac{\gamma}{1+\gamma} \Delta_{\max}\end{aligned}\quad (5.46)$$

where

$$\Delta_{\max} = \frac{2}{\lambda_{\max}} \frac{\bar{\lambda}}{(2N+1)\phi_0}$$

As mentioned before, convergence behaviour for the LMS algorithm is only an approximation to the prefilter convergence case because of the constraint which was imposed in Equation (4.19). In the next chapter, we will study the convergence results of the adaptive prefilter after using this approximation for a given ratio of the excess mean square error to the MMSE.

Chapter 6

Computer Simulation Study

A FORTRAN program was written to simulate the operation of the iterative least mean square error prefilter developed in this thesis in a synchronous data communication system. The prefilter tap gain coefficients are driven toward their optimal values which, tends to minimize the mean square distortion at the output. The program was implemented on a DEC VAX-11/780 computer.

In the next section, the results are presented but the computer implementation is described in Appendix D.

6.1. Computer simulation results

In order to study the performance of the prefilter adaptive algorithm, it is simulated using two types of channels:

- 1) A data-quality telephone channel, and
- 2) A channel with severe intersymbol interference.

6.1.1. Simulation results for a data-quality telephone channel

The equivalent discrete-time characteristic and the spectral characteristic for the channel under consideration are shown in Figures 6.1 and 6.2 respectively [17]. The

resulting eigenvalues of the channel autocorrelation matrix with signal-to-noise ratio (S/N) of 30dB are given in Table 1. A piece of the input and output waveforms of the adaptive prefilter after 5000 iterations is shown in Figures 6.3 and 6.4 respectively. A step size parameter Δ of 0.02 was used to adjust the coefficients. This choice of Δ was obtained from several experiments which showed that this step size provides the quickest compromise convergence for this channel. The equivalent convergence rate of the prefilter equalizer, with 41 tap coefficients, for 5000 iterations is shown in Figure 6.5.

One trial for choosing the step size is applying the results developed for the LMS (Chapter 5), which conclude that

$$\Delta_{\max} = \frac{2}{\lambda_{\max}} \cdot \frac{\bar{\lambda}}{(2N + 1)\phi_0}$$

and

$$\Delta_{\text{steady state}} = \frac{\gamma}{1 + \gamma} \Delta_{\max}$$

The resulting last 100 iterations of the output waveform for 5000 iterations and the rate of convergence for $\gamma = 0.01$ or, equivalently, $\Delta = 0.004$, with 41 tap coefficients, are shown in Figures 6.6 and 6.7 respectively.

Eigenvalue number	Eigenvalue
1	9.1347407E-04
2	9.1352989E-04
⋮	⋮
21	9.6581952E-04
22	1.1128789E-03
⋮	⋮
30	7.2502294E-03
31	0.1020430
32	0.2697555
33	0.4903785
34	0.6115322
35	0.7943755
36	0.9479132
37	0.9824494
38	1.070048
39	1.118529
40	1.494431
41	1.497907

Table 1. The eigenvalues of the autocorrelation matrix corresponding to a data-quality telephone channel (Figure 6.1)

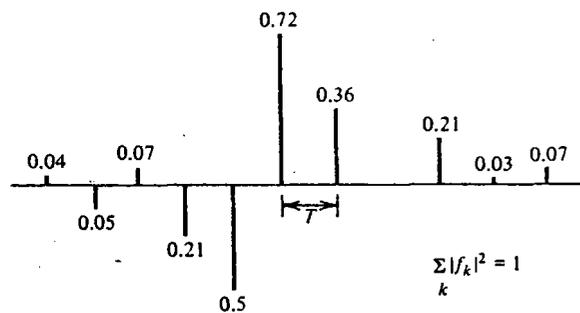


Figure 6.1. The equivalent discrete-time characteristic for a data-quality telephone channel

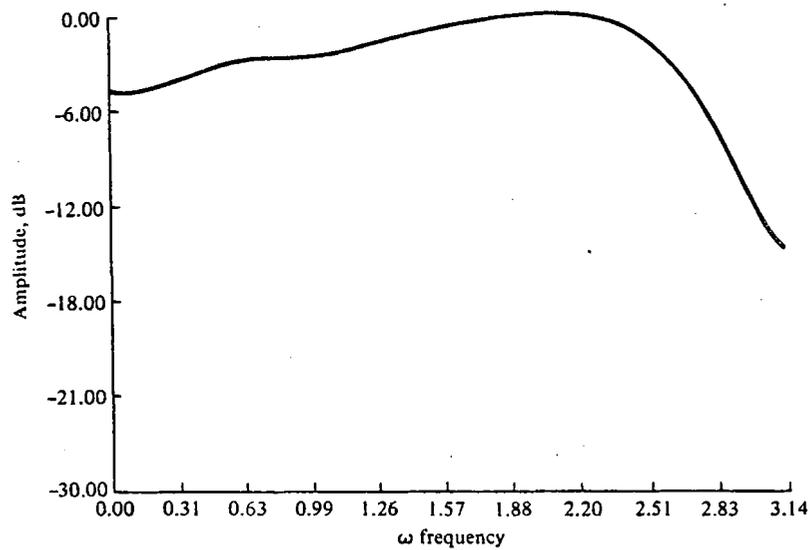


Figure 6.2. Amplitude spectrum for the channel shown in Figure 6.1

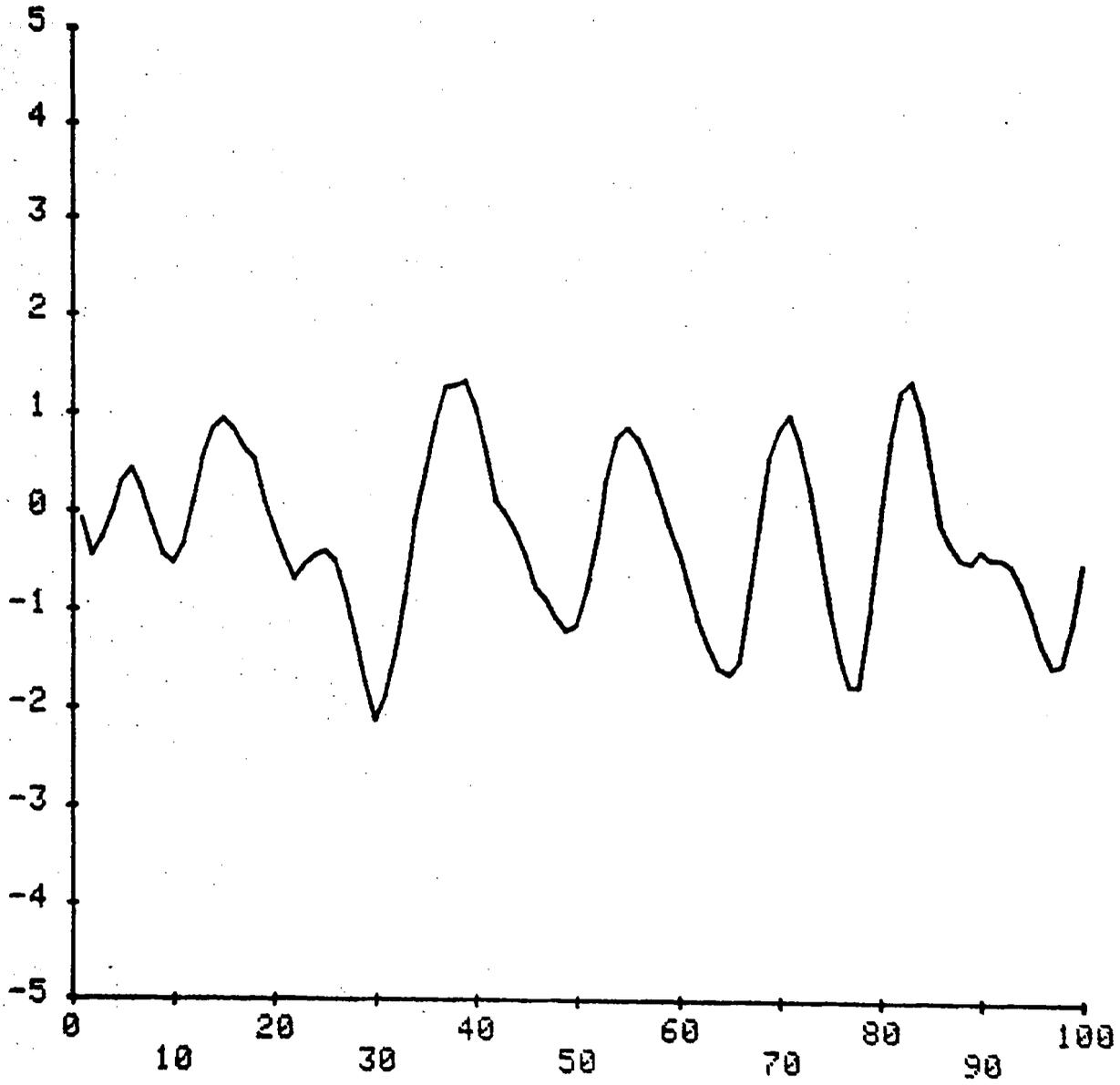


Figure 6.3. The last 100 samples of the input prefilter waveform from 5000 samples, using the channel of Figure 6.1

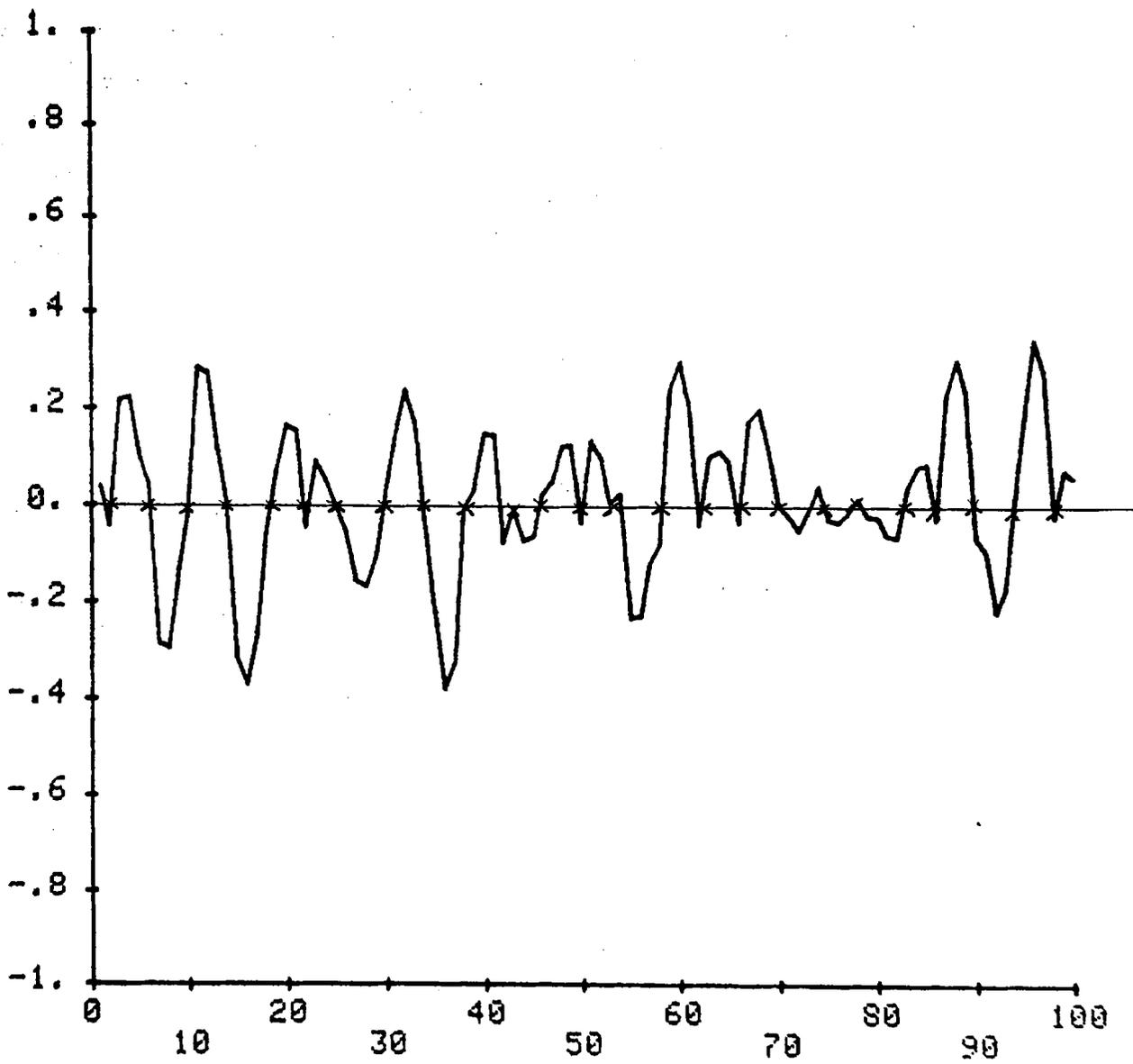


Figure 6.4. The result of the last 100 iterations of the output signal of the prefilter for 5000 iterations for $\Delta_n = 0.02$ using the channel of Figure 6.1

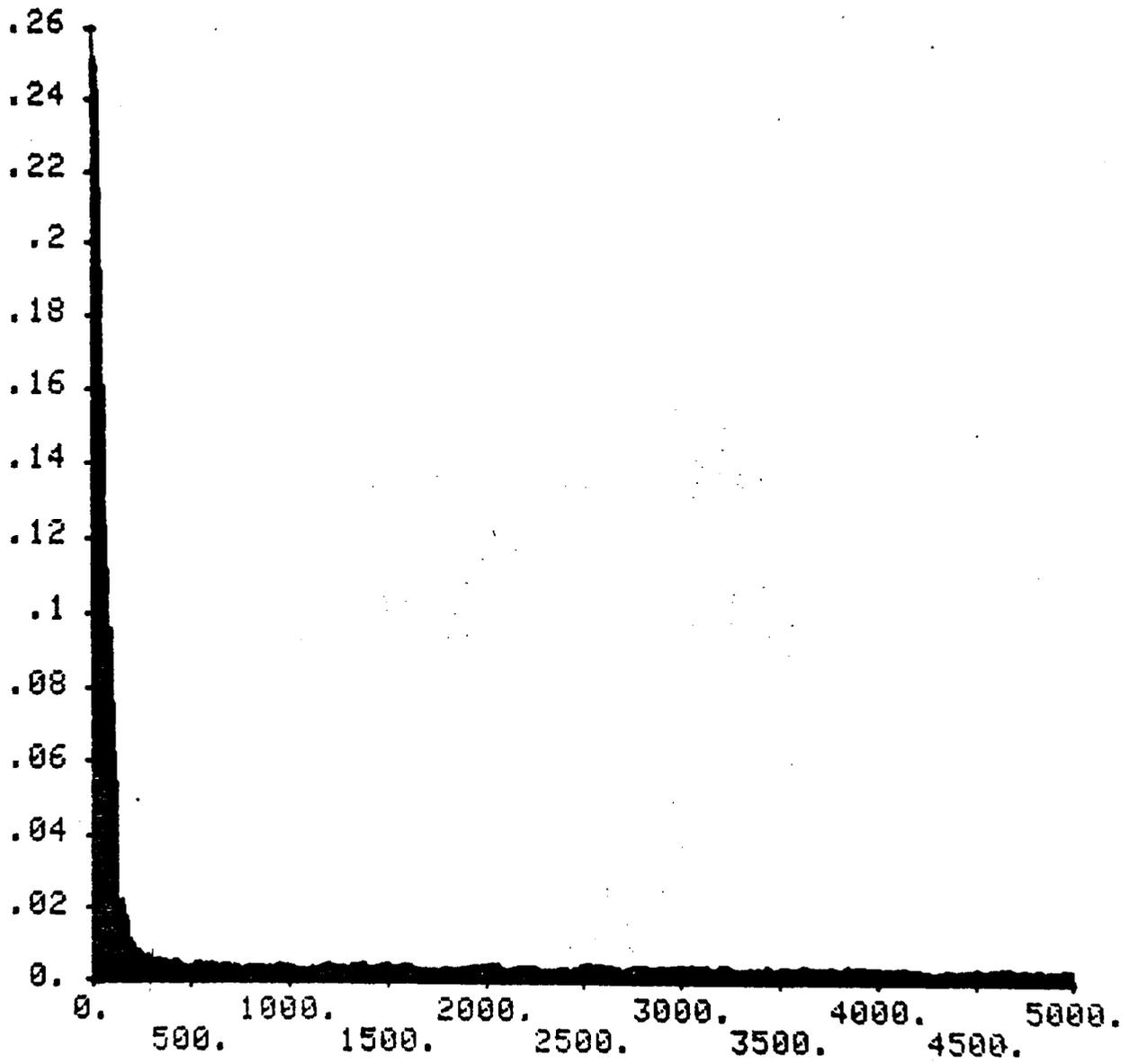


Figure 6.5. The tap MSE convergence rate for $\Delta = 0.02$ and 41 taps using the channel of Figure 6.1

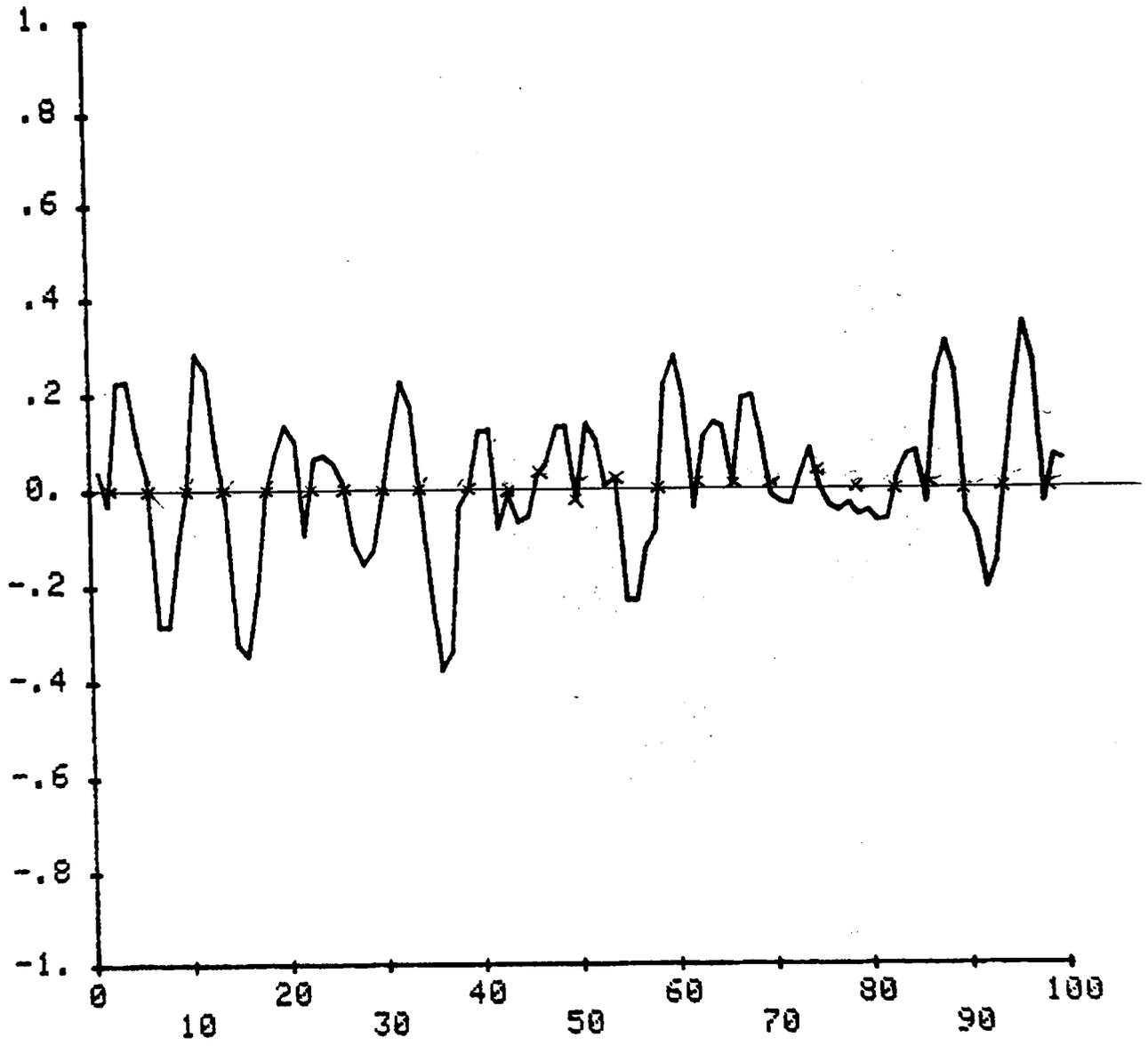


Figure 6.6. The result of the last 100 iterations of the output waveform for 5000 iterations for $\Delta = 0.004$ and 41 taps for the channel of Figure 6.1

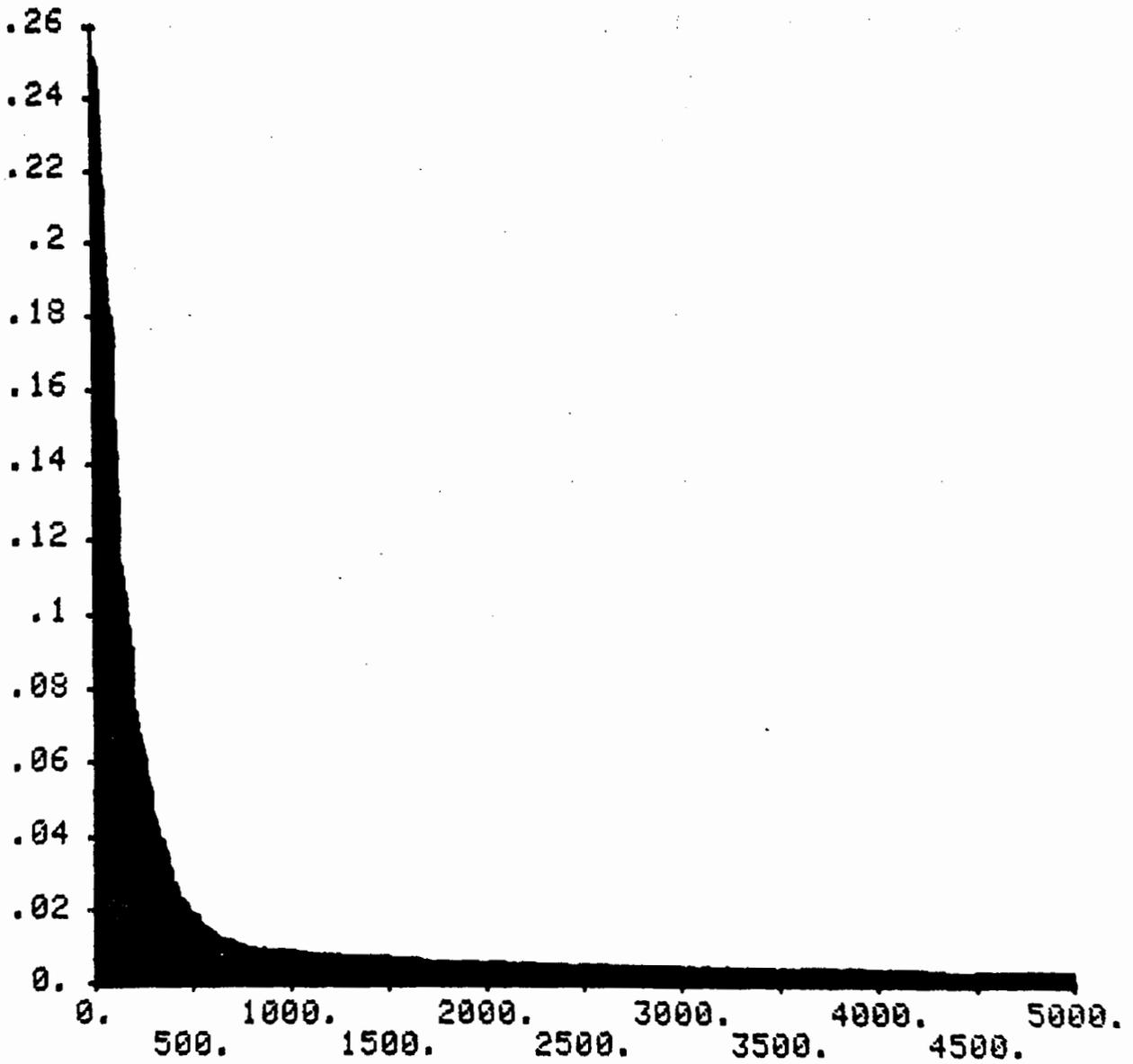


Figure 6.7. The tap MSE convergence rate for $\Delta = 0.004$ and 41 taps, using the channel of Figure 6.1

By examining these results, we can see that

1. From Table 1, half the eigenvalues are almost equal to the noise power in the channel, which is 9.1364×10^{-4} . The values of these eigenvalues are zero when there is no noise in the system. Half of the remaining eigenvalues are very small, which represents the transition band of the channel characteristic. The last quarter of the eigenvalues are the non-zero ones.
2. Comparison between Figures 6.3 and 6.4 shows that the adaptive prefilter improves the uniformity of the zero crossings in the output signal. From the numerical results, the magnitude of the output signal varies between ± 0.003 and ± 0.05 at the desired zero-crossing point.
3. For a step size of 0.004, the magnitude of the output signal varies between ± 0.008 and ± 0.03 at the desired zero-crossing points (Figure 6.6).
4. Comparing the results for the rate of convergence demonstrates that for the case of $\Delta = 0.02$ (Figure 6.5), it takes approximately 300 iterations for the tap MSE to reach its steady-state value. For the case of $\Delta = 0.004$ (Figure 6.7), it takes approximately 900 iterations. These results indicate that the algorithm converges faster for larger values of Δ , but the MSE strays farther from its steady-state value due to the large step size. This behavior is to be contrasted with that shown in Figure 6.7, which is for the case $\Delta = 0.004$.

As a compromise, we choose a step size between 0.02 and 0.004. When $\Delta = 0.01$, the tap MSE reached its steady-state value after 500 iterations while remaining relatively very smooth, as illustrated in Figure 6.8. The magnitude of the output

signal varies between ± 0.005 and ± 0.04 at the desired zero-crossing points (Figure 6.9).

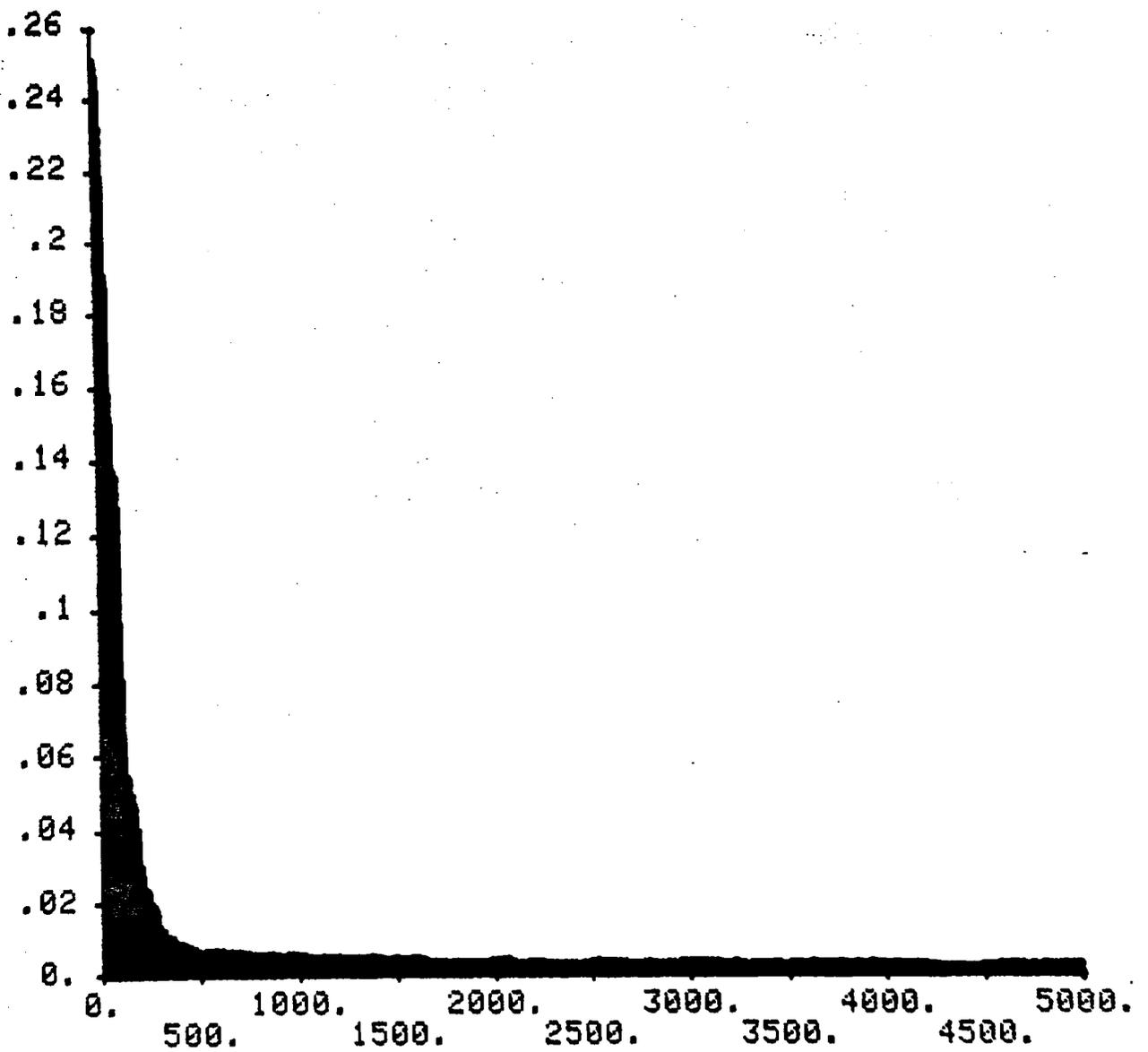


Figure 6.8. The tap MSE convergence rate for $\Delta = 0.01$ using the channel given in Figure 6.1

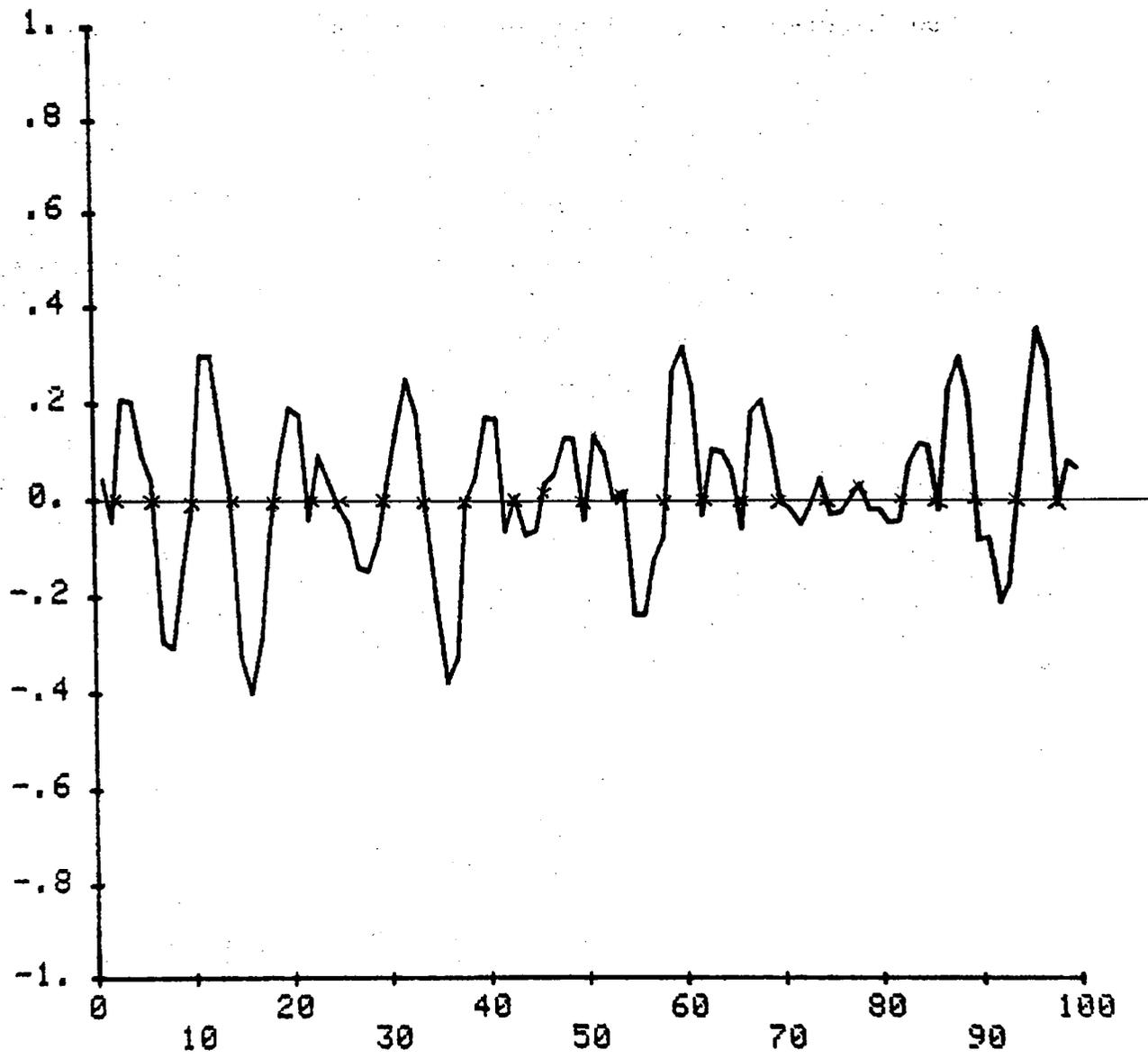


Figure 6.9. The resulting last 100 iterations of the output waveform for 5000 iterations for $\Delta = 0.01$ using the channel from Figure 6.1

6.1.2. Simulation results for a channel with severe intersymbol interference

Examples of the equivalent discrete-time characteristic and the spectral characteristic for this type of channel are shown in Figures 6.10 and 6.11 respectively. The computed eigenvalues of this channel autocorrelation matrix with a S/N ratio of 30dB are given in Table 2. The same investigation steps are taken in this example as in the previous example. So, the input and output waveforms, after 5000 iterations, with an experimental step size of 0.115, are shown in Figures 6.12 and 6.13 respectively. Figure 6.14 shows the convergence characteristics of the MSE of the adaptive prefilter with 9 taps. The results of the last 100 iterations of the output waveform for 5000 iterations and the convergence characteristics of the MSE for $\gamma = 0.01$, or, equivalently, $\Delta = 0.012$, which due to applying the results of the LMS algorithm, are shown in Figures 6.15 and 6.16 respectively.

Eigenvalue number	Eigenvalue
1	9.8612322E-04
2	9.8613382E-04
3	9.8615745E-04
4	9.8618923E-04
5	8.4129293E-03
6	2.8830357E-02
7	6.6886090E-02
8	0.5492329
9	1.572700

Table 2. The eigenvalues of the autocorrelation matrix corresponding to a channel with severe intersymbol interference (Figure 6.7)

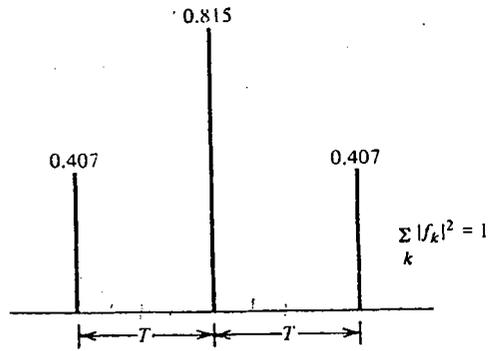


Figure 6.10. The equivalent discrete-time characteristic for a channel with severe intersymbol interference

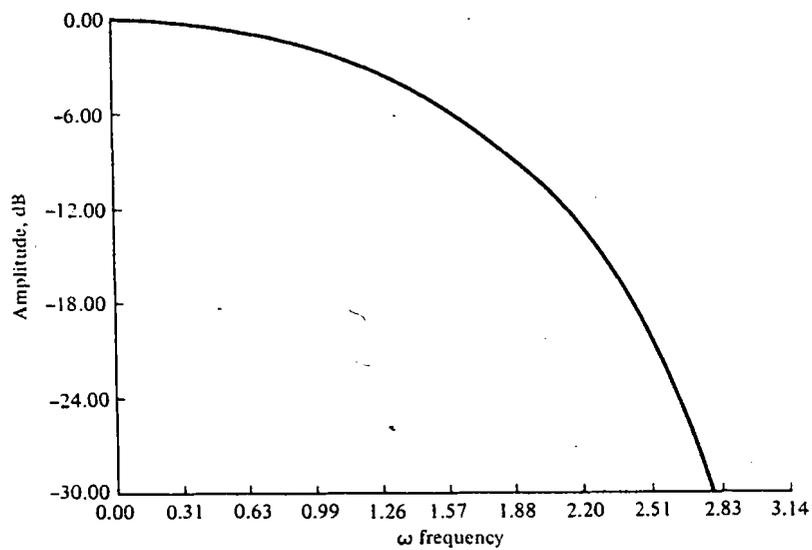


Figure 6.11. Amplitude spectrum for the channel of Figure 6.10

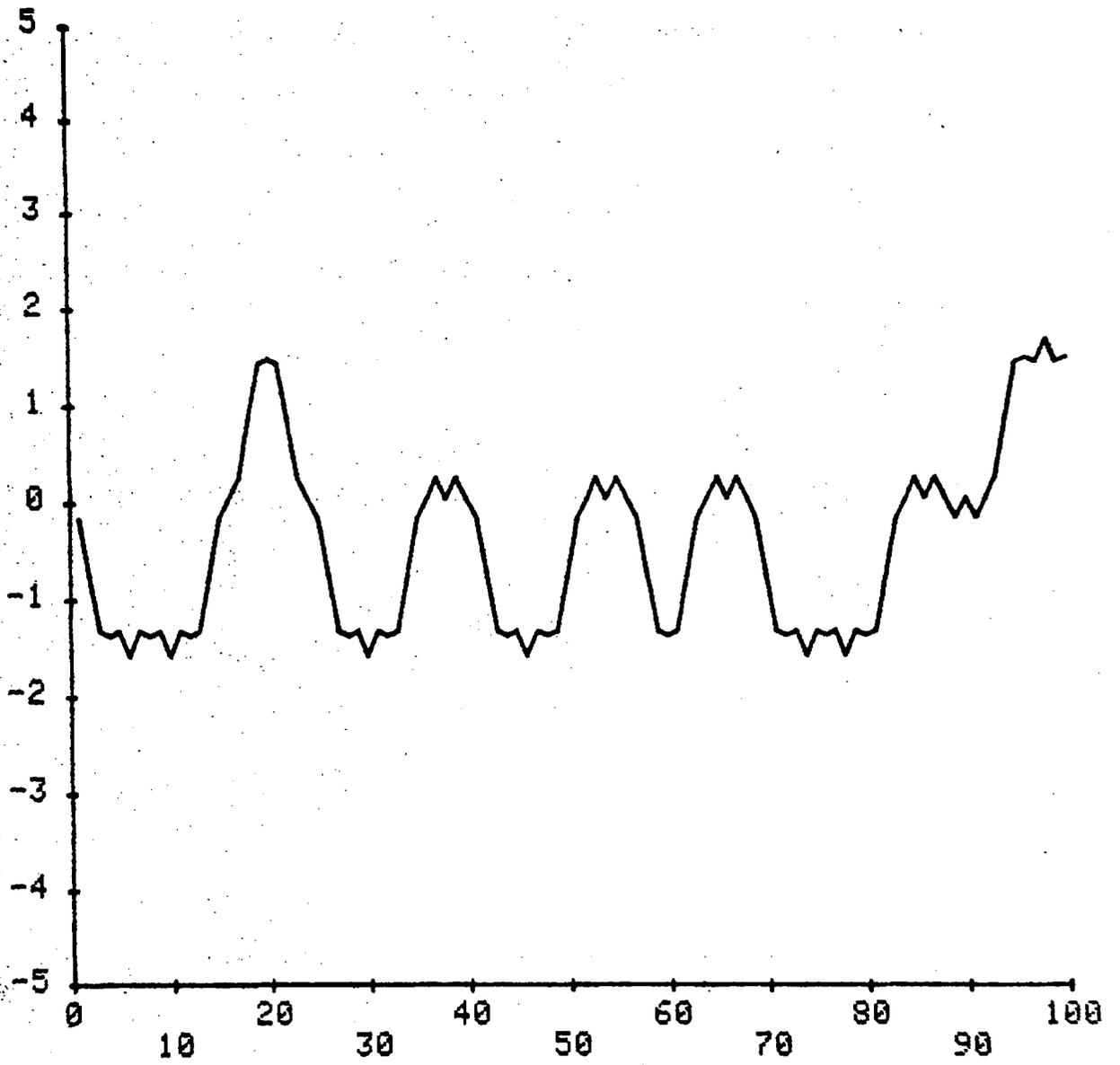


Figure 6.12. The last 100 samples of the input prefilter waveform, from 5000 samples using the channel given in Figure 6.10

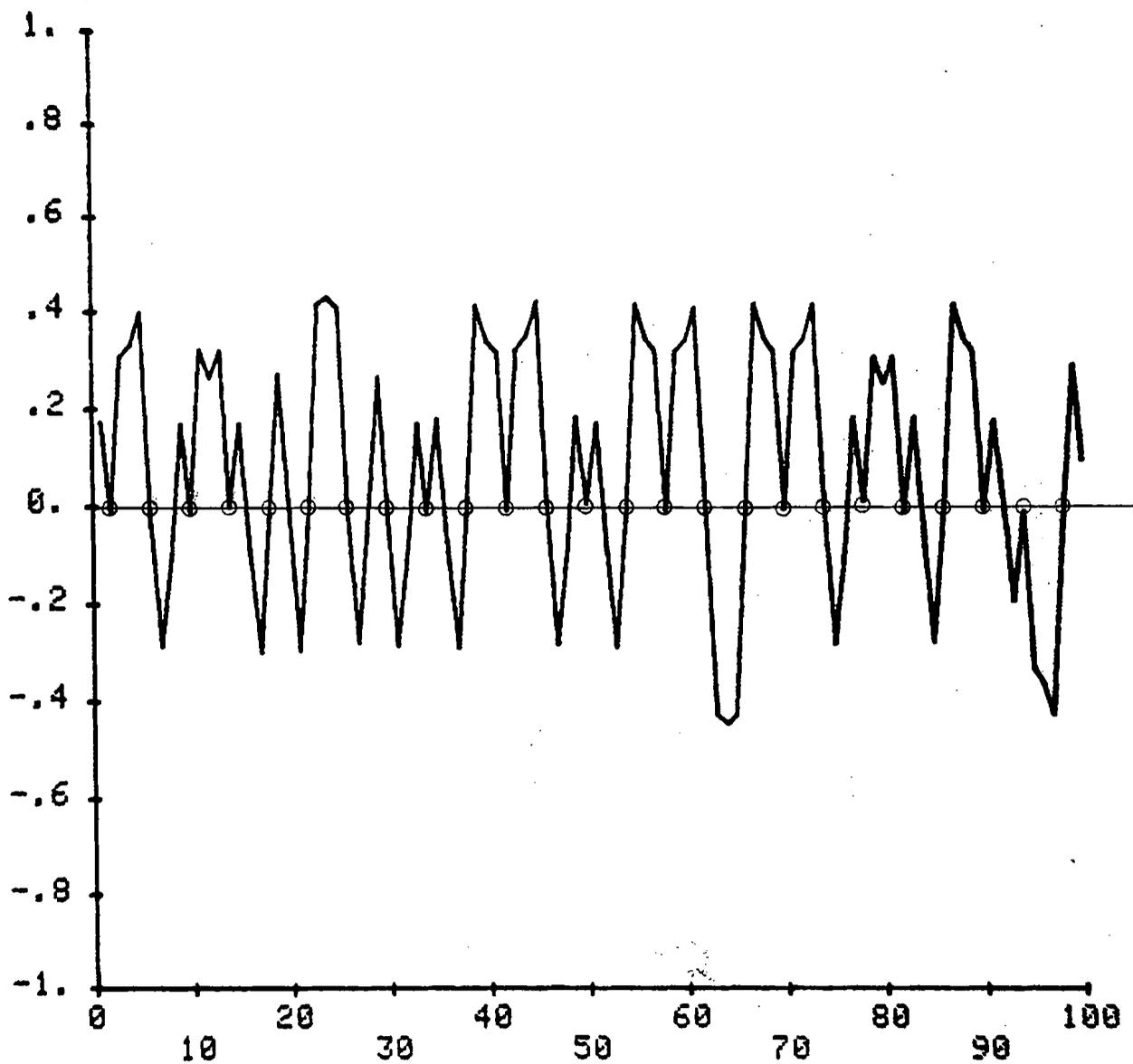


Figure 6.13. The result of the last 100 iterations of the output prefilter waveform, from 5000 iterations, for $\Delta = 0.115$, using the channel of Figure 6.10

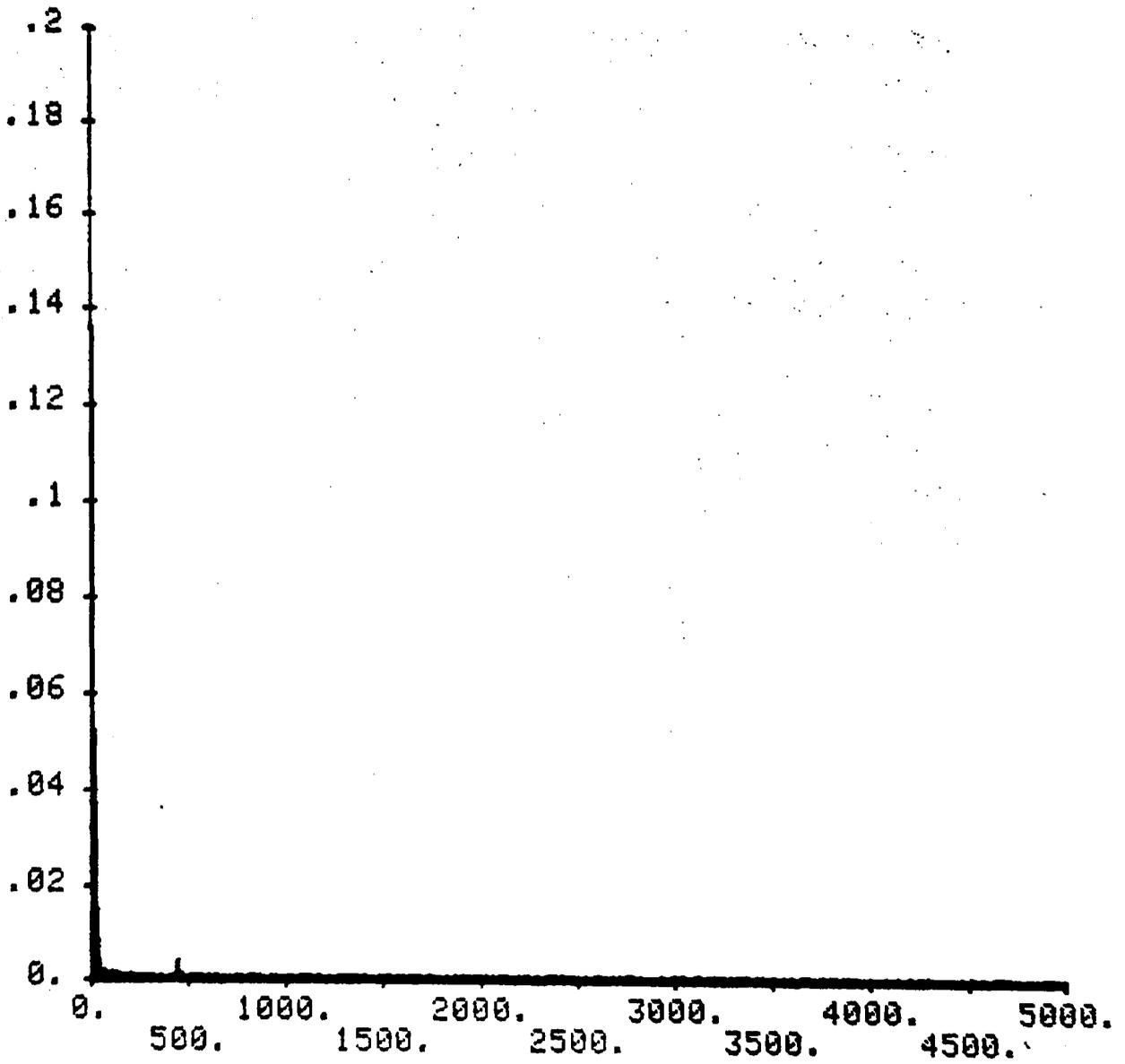


Figure 6.14. The convergence characteristic of MSE of the prefilter with 9 taps for $\Delta = 0.115$, using the channel of Figure 6.10

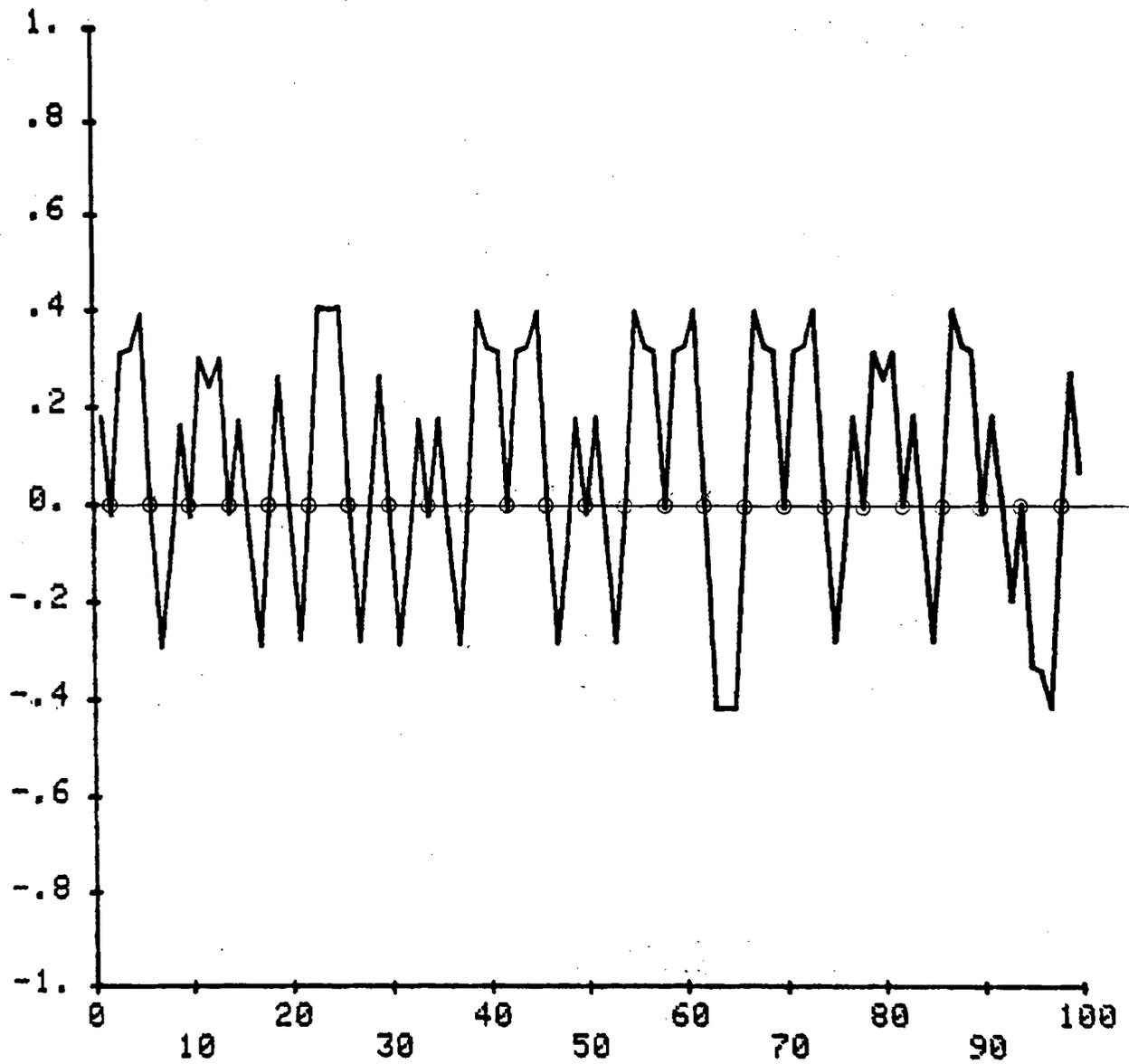


Figure 6.15. The result of the last 100 iterations of the output waveform, from 5000 iterations, for $\Delta = 0.012$ and the channel of Figure 6.10

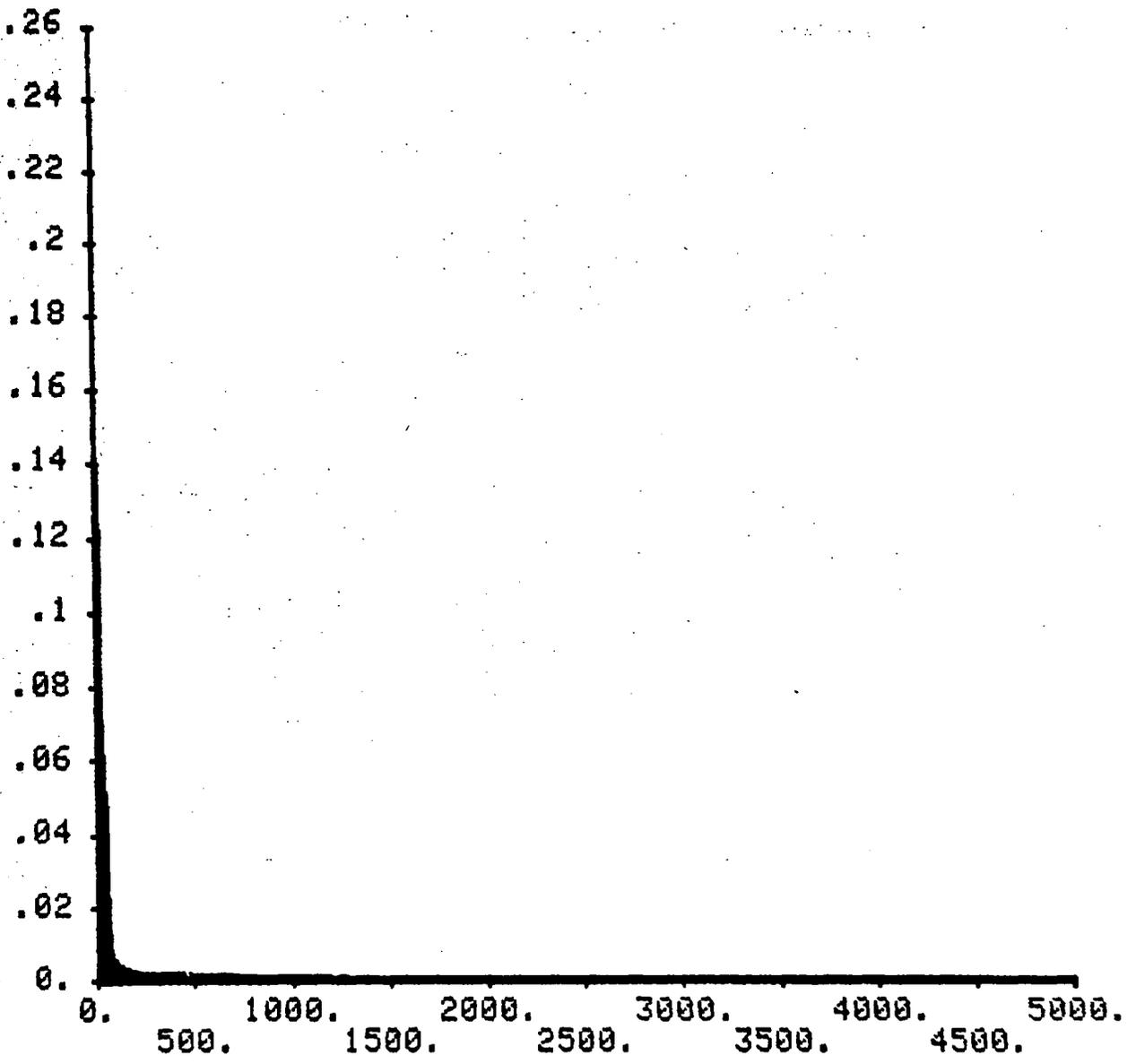


Figure 6.16. The convergence characteristics of the MSE for $\Delta = 0.012$ and the channel of Figure 6.10

A careful study of these results illustrates that

1. A similar observation can be made for the eigenvalues [Table 2] as in the first example, i.e., only one quarter of them are not zero.
2. The uniformity and the accuracy of the zero crossings in the output signal, with both step sizes, is very clear. From the numerical results, the magnitude of the output waveform, for $\Delta = 0.115$ (Figure 6.13), varies between zero and ± 0.001 at the desired zero crossing points. In the case of $\Delta = 0.012$, the variation is between ± 0.001 and ± 0.002 (Figure 6.15).
3. In Figure 6.14, which is the case $\Delta = 0.115$, approximately 80 iterations are required to reach steady-state operation and there is almost no deviation from steady-state operation once it has been reached. For $\Delta = 0.012$ the MSE steady-state value is obtained after 150 iterations with no deviations. Comparing the results for both channels leads to a question: Why are more accurate results obtained for the channel with intersymbol interference? To answer this question we need to compare the channel characteristics, which are shown in Figures 6.1 and 6.10. Note that the discrete-time characteristic of the channel with intersymbol interference is symmetrical, which is not the case for the other channel. The output waveform from the symmetrical channel has certain symmetry characteristics, even with the intersymbol interference, as indicated in Figure 6.12, which the other channel is missing completely.

This observation indicates that an important factor for uniform zero crossings of the timing wave is the symmetry of the pulse shape entering the timing path and that the intersymbol interference has no real effect on the uniformity of the zero crossings.

Chapter 7

Summary and Conclusions

This study has proposed a fast-converging technique to improve the timing recovery method for synchronous PAM data signals. This technique is based on digitizing the analog timing recovery scheme by using an adaptive digital prefilter.

The timing circuit under consideration consists of a square-law device followed by a narrow-band filter tuned to the pulse repetition frequency preceded by a prefilter for reshaping the pulse entering the timing path. The output of the timing circuit is a nearly sinusoidal timing wave whose zero crossings indicate the appropriate sampling instants for demodulation of the PAM signal. For a random data sequence, the timing wave exhibits phase fluctuations which are strongly dependent on the shape of the pulses entering the timing path and the pass-band shape of the narrow-band filter. Expressions for the rms phase fluctuations in the timing wave are represented as a function of the prefiltering characteristics of the filter preceding the square-law device and which have a form especially suitable when the signal is band-limited to frequencies less than the pulse repetition frequency. From the condition on prefiltering to optimize the pulse shaping and to give jitter-free timing recovery the performance of the self-adjusting prefilter is presented. In the trend towards digitization of the timing circuit, the signal is sampled at the input to the timing path. Sampling is performed

at a high enough rate taking into account the effect of the square-law device, which doubles the frequency range occupied by the signal. The resulting sampling rate, which is four times the baud rate, requires the prefilter to be a fractional-tap prefilter equalizer. The advantage of this kind of equalizer over T -spaced equalizers is explained. Several interesting aspects of the optimum tap-coefficient problem are revealed by examining the minimum mean-squared error (MMSE), which depends on the channel autocorrelation matrix. The least-mean-square (LMS) algorithm is studied as an approximation to a iterative algorithm for adjusting the tap-gain coefficients. The convergence behavior of the adaptive prefilter passed on the LMS algorithm was measured in two ways:

- 1) The behavior of the mean coefficient vector, and
- 2) The excess mean square error.

The convergence properties of several examples were also studied by computer simulation. The behavior exhibited by such tap gains for different values of Δ and two types of channel indicate that the fractionally-tapped prefilter equalizer can adapt the zero-crossings of the timing wave in fewer than 600 iterations (600 transmitted pulses). In the current trend towards fully digital receivers using medium and large scale integration (MSI and LSI) technology, the present technique can be easily used to achieve both fast convergence and almost complete freedom from jitter.

Appendix A

Derivation of the variance of the timing wave

Our objective is to obtain a convenient expression for either $E[z^2(t)]$ or the variance $\text{Var } z(t)$ of the timing wave, which will be manageable even when the degree of pulse overlapping is very large. We assume that the data mean value is zero and that the elements of the data sequence are statistically independent. From Equation (3.1),

$$u^2(t) = \sum_k \sum_m a_k a_{m+k} P_m(t - kT) \quad (\text{A.1})$$

so then

$$z(t) = \sum_k \sum_m a_k a_{m+k} q_m(t - kT) \quad (\text{A.2})$$

and

$$E[z^2(t)] = \sum_k \sum_m \sum_j \sum_\ell E[a_k a_{k+m} a_{k+j} a_{k+j+\ell}] q_m(t - kT) q_\ell(t - kT - jT)$$

where

$$q_m(t) = [h \otimes p_m](t) \quad (\text{A.3})$$

But we have

$$E[a_k a_{k+m} a_{k+j} a_{k+j+\ell}] = \begin{cases} \overline{a^4} & \text{if } m = \ell = j = 0 \\ \overline{a^2}^2 = R_{aa}(0) & \text{if } m = \ell = 0, j \neq 0 \\ \cdot & \text{if } m = j \neq 0, \ell = -j \\ \cdot & \text{if } m = \ell \neq 0, j = 0 \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.4})$$

Now using (A.4) to collect the nonzero terms in (A.3) in the proper order, the result may be written as

$$\begin{aligned}
E[z^2(t)] = & \left\{ R_{aa}(0) \sum_k q_0(t - kT) \right\}^2 + \left(\overline{a^4} - 3R_{aa}^2(0) \right) \sum_k q_0^2(t - kT) \\
& + R_{aa}^2(0) \sum_k \sum_m q_m(t - kT) q_{-m}(t - kT - mT) \\
& + R_{aa}^2(0) \sum_k \sum_m q_m^2(t - kT) \quad (A.5)
\end{aligned}$$

From (A.2), the first term in (A.5) is the square of the mean. Also, it can be shown that the last two terms in (A.5) are equal, so the variance can be written as

$$\text{Var } z(t) = 2R_{aa}^2(0) \sum_k \sum_m q_m^2(t - kT) + \left(\overline{a^4} - 3R_{aa}^2(0) \right) \sum_k q_0^2(t - kT) \quad (A.6)$$

Now, considering the Fourier series of (A.6), obtaining the coefficients by application of the Poisson sum formula, and defining

$$a(t) = \sum_m q_m^2(t)$$

we get that

$$\sum_k \sum_m q_m^2(t - kT) = \sum_k a(t - kT) \quad (A.7)$$

So,

$$\sum_k a(t - kT) = \frac{1}{T} \sum_{\ell} A \left(\frac{\ell}{T} \right) \exp \left[\frac{j2\pi\ell t}{T} \right] \quad (A.8)$$

and

$$\sum_k \sum_m q_m^2(t - kT) = \frac{1}{T} \sum_r A\left(\frac{r}{T}\right) \exp\left[\frac{j2\pi r t}{T}\right] \quad (\text{A.9})$$

where

$$A(f) = \sum_m [Q_m \otimes Q_m](f) = \sum_m \left[\int_{-\infty}^{\infty} Q_m(v) Q_m(f - v) dv \right] \quad (\text{A.9a})$$

$$Q_m(f) = H(f) P_m(f)$$

$$P_m(t) = g(t)g(t - mT)$$

Then

$$\begin{aligned} P_m(f) &= G(f) * [G(f) \exp[-j2\pi f_m T]] \\ &= \int_{-\infty}^{\infty} G(f - \eta) G(\eta) \exp[-j2\pi f_m T] d\eta \end{aligned}$$

so

$$Q_m(f) = H(f) \int_{-\infty}^{\infty} G(f - \eta) G(\eta) \exp[-j2\pi \eta m T] d\eta \quad (\text{A.10})$$

Now, we can rewrite (A.9a) in the form

$$\begin{aligned} A(f) &= \sum_m \int_{-\infty}^{\infty} \left[H(v) \int_{-\infty}^{\infty} G(v - \eta) G(\eta) \exp[-j2\pi \eta m T] d\eta \right. \\ &\quad \left. \cdot \int_{-\infty}^{\infty} H(f - v) G(f - \lambda - v) G(\lambda) \exp[-j2\pi \eta m T] d\lambda \right] dv \\ &= \sum_m \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(v) H(f - v) G(v - \eta) G(\eta) G(f - \lambda - v) G(\lambda) \\ &\quad \cdot \exp[-j2\pi \eta m T] d\lambda dv d\eta \end{aligned}$$

Now apply the Poisson sum formula to the previous equation to obtain

$$A(f) = \frac{1}{T} \sum_{\ell} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(v)H(f-v)G\left(v-\eta-\frac{\ell}{T}\right) G(\eta)G\left(\frac{\ell}{T}-\eta\right)G(f-\eta-v)dv d\eta \quad (A.11)$$

For the other term in Equation (A.6)

$$\sum_k q_0^2(t-kT) = \sum_k B(t-kT) = \frac{1}{T} \sum_r B\left(\frac{r}{T}\right) \exp\left[\frac{j2\pi r t}{T}\right] \quad (A.12)$$

Define

$$b(t) = q_0^2(t)$$

$$B(f) = [Q_0 \otimes Q_0](f)$$

Then from (A.10),

$$Q_0(f) = H(f) \int_{-\infty}^{\infty} G(f-v)G(v)dv$$

Then

$$B(f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(v)H(f-v)G(v-\eta)G(\eta)G(\lambda)G(f-\lambda-v)d\lambda d\eta dv \quad (A.13)$$

After comparing Equations (A.9) and (A.12) with (A.6), we can write $\text{Var } z$ in the form

$$\text{Var } z(t) = \sum_r V_r \exp\left[\frac{j2\pi r t}{T}\right] \quad (A.14)$$

where

$$V_r = 2 \frac{R_{aa}(0)}{T} A\left(\frac{r}{T}\right) + \frac{\bar{a}^4 - 3R_{aa}(0)}{T} B\left(\frac{r}{T}\right)$$

Appendix B

Proof of Equation (4.23)

Theorem:

The following three problems are equivalent, where A is a Hermetian matrix:

(i) Minimize $1/(x, x)$ subject to $(x, Ax) = 1$.

(ii) Minimize $(x, Ax)/(x, x)$ for $x \neq 0$.

(iii) Minimize (x, Ax) subject to $(x, x) = 1$.

Proof:

Let ξ , η , and ζ be the three minimum values from (i), (ii), and (iii) and suppose that these occur at $x = u_0$, v_0 , and w_0 respectively. The definition of u_0 gives, for any other vector u such that $(u, Au) = 1$,

$$F = \frac{1}{(u_0, u_0)} \leq \frac{1}{(u, u)} \quad (\text{with } (u_0, Au_0) = (u, Au) = 1) \quad (1)$$

Set $z = pu_0$ and $u = qv$ for any nonzero p and q . Then

$$\xi = \frac{\bar{p}p}{(z, z)} \leq \frac{qq}{(v, v)} \quad (\text{with } (z, Az) = \bar{p}p \text{ and } (v, Av) = \bar{q}q)$$

Hence,

$$\xi \leq \frac{(v, Av)}{(v, v)} \quad (2)$$

This is true for any $v \neq 0$, and, in particular, for $v = v_0$. Hence

$$\xi \leq \eta \quad (3)$$

Now we reverse the argument. The definition of v_0 gives, for any other vector $v \neq 0$,

$$\eta = \frac{(v_0, Av_0)}{(v_0, v_0)} \leq \frac{(v, Av)}{(v, v)} \quad (4)$$

Set $y = v_0 / \|v_0\|$ and $u = v / \|v\|$. Then (4) gives

$$\eta = (y, Ay) \leq (u, Au) \quad \text{with } (y, y) = (u, u) = 1$$

This is true for any u such that $(u, u) = 1$, and, in particular, for u_0 , which implies $\eta \leq \xi$. Combining this with (3) we obtain $\eta = \xi$, which demonstrates the equivalence of (i) and (ii). The equivalence of (ii) and (iii) is proved similarly.

Appendix C

Proof of Equation (4.24)

Theorem:

If A is a Hermetian matrix with eigenvalues $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, then $\lambda_1 \leq \rho \leq \lambda_n$ and $\lambda_1 = \min_{x \neq 0} \frac{(x, Ax)}{(x, x)}$, $\lambda_n = \max_{x \neq 0} \frac{(x, Ax)}{(x, x)}$ where ρ is defined to be the Rayleigh quotient corresponding to the Hermetian matrix A , which may be expressed as

$$\rho = \frac{(x, Ax)}{(x, x)}$$

for any $x \neq 0$.

Proof:

If A is Hermetian, an orthonormal set of eigenvectors exists, say x_1, x_2, \dots, x_n , where x_i corresponds to λ_i . Suppose that the expansion of an arbitrary vector in terms of the x_i is

$$x = \sum_{i=1}^n \alpha_i x_i$$

Then

$$Ax = \sum_{i=1}^n \alpha_i \lambda_i x_i \tag{1}$$

and

$$\rho = \frac{(x, Ax)}{(x, x)} = \frac{\lambda_1 \bar{\alpha}_1 \alpha_1 + \lambda_2 \bar{\alpha}_2 \alpha_2 + \cdots + \lambda_n \bar{\alpha}_n \alpha_n}{\bar{\alpha}_1 \alpha_1 + \bar{\alpha}_2 \alpha_2 + \cdots + \bar{\alpha}_n \alpha_n} \quad (2)$$

So we have

$$\rho - \lambda_1 = \frac{(\lambda_2 - \lambda_1) \bar{\alpha}_2 \alpha_2 + \cdots + (\lambda_n - \lambda_1) \bar{\alpha}_n \alpha_n}{\bar{\alpha}_1 \alpha_1 + \bar{\alpha}_2 \alpha_2 + \cdots + \bar{\alpha}_n \alpha_n}$$

Since $\lambda_i - \lambda_1 \geq 0$ for all i , we have $\rho \geq \lambda_1$. Also, if we choose $x = x_1$, this gives $\rho = \lambda_1$, which proves the first statement

$$\lambda_1 = \min_{x \neq 0} \frac{(x, Ax)}{(x, x)}$$

The remainder of the theorem is proved similarly by considering $\rho - \lambda_n$.

Appendix D

Computer Implementation

The contents of the computer program can be grouped into four categories.

D.1. The optimal tap-gain coefficients calculations

The object of this simulation is to determine C_{opt} by computing the minimum eigenvalue for the autocorrelation matrix A and its corresponding eigenvector. The program reads in the parameters of the equalizer, i.e., the number of taps and the signal-to-noise ratio. It also reads in the channel samples, which are given at each band rate sampling $1/T$. Before computing the channel autocorrelation matrix, the channel samples were interpolated by the cubic spline interpolation subroutine IC-SCCU to calculate their values at $T_p = T/4$. Then the CORR subroutine calculates the autocorrelation matrix A . The system subroutine EIGRS was used to compute the smallest eigenvalue for the matrix, which is the minimum mean square error (MMSE). This subroutine gives also the corresponding eigenvector, which is the optimum tap coefficient vector C_{opt} .

D.2. Generate the prefilter's input signal at each time interval $T/4$

The input data for this package are the channel samples after interpolation, the desired variance, and the desired number of samples.

Subroutine GGUBS generates pseudo-random independent uniformly distributed (-1 or 1) symbols.

Subroutine GGNML generates pseudo-random independent samples from a zero mean and desired variance Gaussian distribution.

This noise vector is interpolated with the ICSCCU subroutine to determine the values of the noise samples at each $T_p = T/4$. Then, the convolution of the channel samples with the input digits is computed in the CONVOL subroutine with adding the noise samples at each time interval $T/4$.

The first three subroutines used in this package are drawn from the IMSL library and the last (CONVOL) was written by the author.

D.3. Generate the prefilter's output signal

The program reads in the tap coefficients and the input samples which were generated by the previous package. Note that the tap coefficients are updated each four time intervals T_p (i.e., each T). So, for the first calculation of the output signal the program reads in the initial tap coefficients which it is given as input data. Then after each updating iteration, it reads in the new coefficient values from the output data and uses them to compute the prefilter's output signal.

The convolution sum of the tap coefficients with the prefilter's input signal is calculated by the CONVOL subroutine to generate the output signal.

D.4. Updating the prefilter's coefficients

This program and the previous one are strongly related in the sense that we are using the output of the previous program at each nT^* to calculate the updating vector according to Equation (5.10). For each iteration n , we are calling subroutine UPDAT, in which the step size Δ (given data) is multiplied by the output signal at nT . The resulting value is then multiplied by each element of the input vector of the same time interval nT resulting in the gradient vector. Each element of this vector is added to its corresponding tap coefficient, resulting in the updated coefficient vector.

In subroutine NORMTAP, we are summing the square of each element of the updating vector, then dividing each element by the sum to determine the normalized updated coefficient vector.

In subroutine MSE, we are doing matrix multiplication to calculate the mean square error (MSE) according to the formula $C^T A C$, namely, the product of a $1 \times (2N + 1)$ row vector C^T (the transpose of the coefficient vector) with the $(2N + 1) \times (2N + 1)$ autocorrelation matrix is then multiplied by a $(2N + 1) \times 1$ column vector C resulting in the MSE.

Note that the matrix is symmetric, so by doing the previous multiplication operations on the diagonal elements and the lower (higher) triangular matrix of matrix A only, we decrease the calculations by almost half. All the subroutines used in this package were written by the author.

* $n = 0, 1, 2, \dots$ is the iteration index

References

- [1] W. R. Bennett and J. R. Davey, *Data Transmission*, New York, McGraw-Hill, 1965.
- [2] J. J. Stiffler, *Theory of Synchronous Communication*, Englewood Cliffs, NJ, Prentice-Hall, 1971.
- [3] B. R. Saltzberg, "Timing recovery for synchronous binary data transmission", *Bell Syst. Tech. J.*, Vol. 46, No. 3, March 1967, pp. 593–622.
- [4] R. D. Gitlin and J. Salz, "Timing recovery in PAM systems", *Bell Syst. Tech. J.*, Vol. 50, No. 5, May-June 1971, pp. 1645–1669.
- [5] R. D. Gitlin, J. F. Hayes, "Timing recovery and scramblers in data transmission", *Bell Syst. Tech. J.*, Vol. 54, No. 3, March 1975, pp. 569–593.
- [6] W. R. Bennett, "Statistics of regenerative digital transmission", *Bell Syst. Tech. J.*, Vol. 37, No. 6, November 1958, pp. 1501–1542.
- [7] Y. Takasaki, "Systematic jitter due to imperfect equalization", *IEEE Trans. Comm.*, Vol. COM-19, No. 6, December 1971, pp. 1275–1276.
- [8] Y. Takasaki, "Timing extraction in baseband pulse transmission", *IEEE Trans. Comm.*, Vol. COM-20, No. 5, October 1972, pp. 877–884.

- [9] L. E. Franks and J. P. Bubrouski, "Statistical properties of timing jitter in a PAM timing recovery scheme", *IEEE Trans. Comm.*, Vol. COM-22, No. 7, July 1974, pp. 913–920.
- [10] K. H. Muller and M. Müller, "Timing recovery in digital synchronous data receivers", *IEEE Trans. Comm.*, Vol. COM-24, No. 5, May 1976, pp. 516–531.
- [11] L. E. Franks, *Signal Theory*, Englewood Cliffs, NJ, Prentice-Hall, 1969.
- [12] G. L. Cariolaro and F. Todero, "A general spectral analysis of time jitter produced in a regenerative repeater", *IEEE Trans. Comm.*, Vol. COM-25, No. 4, April 1977, pp. 417–426.
- [13] C. J. Byrne, B. J. Karafin, and D. B. Robinson, "Systematic jitter in a chain of digital regenerators", *Bell Syst. Tech. J.*, Vol. 42, No. 6, November 1963, pp. 2679–2741.
- [14] J. M. Manley, "The generation and accumulation of timing noise in PCM systems—An experimental and theoretical study", *Bell Syst. Tech. J.*, Vol. 48, No. 3, March 1969, pp. 541–613.
- [15] L. E. Franks, "Carrier and bit synchronization in data communications—A tutorial review", *IEEE Trans. Comm.*, Vol. COM-28, No. 8, August 1980, pp. 1107–1120.
- [16] K. Feher, *Digital Communication Satellite/earth station engineering*, Prentice-Hall, Englewood Cliffs, NJ, 1983.
- [17] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, 1983.

- [18] G. Ungerboeck, "Fractional tap-spacing equalizer and consequences for clock recovery in data modems", *IEEE Trans. Comm.*, August 1976, Vol. COM-24, No. 8, pp. 856-864.
- [19] R. D. Gitlin and S. B. Weinstein, "Fractionally-spaced equalization: An improved digital transversal equalizer", *Bell Syst. Tech. J.*, Vol. 60, No. 2, February 1981, pp. 275-296.
- [20] B. Noble, *Applied Linear Algebra*, Prentice-Hall, Englewood Cliffs, NJ, 1969.
- [21] R. D. Gitlin, H. C. Meaders, and S. B. Weinstein, "The tap-leakage algorithm: An algorithm for the stable operation of a digitally implemented, fractionally spaced adaptive equalizer", *Bell Syst. Tech. J.*, Vol. 61, No. 8, October 1982, pp. 1817-1838.
- [22] J. G. Proakis and J. H. Miller, "Adaptive receiver for digital signaling through channels with intersymbol interference", *IEEE Trans. Comm.*, Vol. 15, No. 4, July 1969, pp. 484-496.
- [23] M. L. Honig and Messerschmitt, *Adaptive Filters: Structures, Algorithms, and Applications*, Academic, New York, 1984.
- [24] S. Haykin, *Introduction to Adaptive Filters*, Macmillan, New York, 1984.
- [25] R. D. Gitlin and S. B. Weinstein, "On the required tap-weight precision for digitally implemented, adaptive, mean-squared equalizers", *Bell Syst. Tech. J.*, Vol. 58, No. 2, February 1979, pp. 301-321.