A BANDWIDTH EFFICIENT DIGITAL RADIO SYSTEM

By

Ka-Sing Sammy Yeung, B. Eng.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Engineering

McMaster University

October 1979

A BANDWIDTH EFFICIENT DIGITAL RADIO SYSTEM

MASTER OF ENGINEERING (1979) McMASTER UNIVERSITY

(Electrical Engineering) Hamilton, Ontario

TITLE: A Bandwidth Efficient Digital Radio System

AUTHOR: Ka-Sing Sammy Yeung, B. Eng. (McMaster University)

SUPERVISOR: Dr. D. P. Taylor

NUMBER OF PAGES: ix, 111

ABSTRACT

A feasibility study has been performed on a 4 bits/s/Hz digital radio system. This system utilizes a 4-level quadrature amplitude modulation scheme with Duobinary filtering to achieve such high bandwidth efficiency. The error rate performance of such systems is evaluated. A repeater structure including the transmitter and the receiver of such systems is developed and analyzed. A decision directed carrier recovery loop is presented and analyzed using stochastic techniques. Expressions for the effective phase detector characteristic and the steady state phase error probability density function are derived. A symbol timing recovery structure is also presented and exhaustive simulations have been performed to analyze the performance of the timing recovery structure.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the advice, guidance and encouragement given by his supervisor, Dr. D.P. Taylor, throughout the course of this work.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (continued)

LIST OF ILLUSTRATIONS (continued)

CHAPTER 1

INTRODUCTION

Since the discovery of the use of a new natural resource, the radio spectrum, the demand for communications has grown tremendously. The spectrum is now overcrowded by millions of radio transmitters. According to one prediction, the entire available radio spectrum up to 20 GHz will be fully occupied by the end of the century [1]. In order to combat this congestion problem, we have to conserve this precious resource while looking for new discoveries. Conservation would mean better management of existing allocations, and the use of spectrally efficient modulation. The latter means the transmission of as much information as possible in a fixed time and a fixed bandwidth.

Microwave techniques have been employed extensively for radio communication since World War II. Because of the non-linearities of microwave amplifiers, microwave transmission is unsuitable for use with amplitude modulation (AM) systems. However, it is possible to change the frequency of the carrier in a substantially linear fashion. Furthermore, it is well known that Frequency Modulation (FM) systems have better noise performance than AM systems. This is why most

existing microwave systems employ FM techniques.

With the recent advances in digital circuit and microelectronics technology, most analog equipment will be converted to its digital counterpart eventually. This applies to communications equipment including radio repeaters. One of the most important advantages in switching from analog to digital radio is that digital signals can be regenerated at each repeater. Hence there is no cumulative performance degradation due to noise, amplitude and phase distortion of the channel, whereas, analog signals cannot be regenerated because they are indistinguishable in form from noise. With digital transmissions, there is no distinction between data, PCM voice or digitized TV signals. This provides an easy means of mixing different types of signals. Furthermore, digital equipment usually has built-in monitoring and diagnostic capabilities, which will identify equipment failure in much less time than in an analog system. In addition, the advances in digital switching technology have stimulated the switch to digital equipment, because a complete digital communications and switching hierachy can provide for easy interfacing between different equipment. The early digital radio links usually operated using on-off keying of the radio carrier [27]. The number for voice channels handled, was quite restricted because of the relatively large

bandwidth required to accommodate the spectrum. Somewhat later, phase-shift keying (PSK) types of modulation schemes were used. With these modulation schemes, the information is contained in either the phase or the frequency of the carrier. There is absolutely no information carried by the amplitude of the carrier, and the carrier has a constant envelope. These modulations were chosen to avoid the problem of the nonlinear characteristics of the traveling wave tube amplifier (TWTA) used at the transmitter. However, these modulation schemes can only achieve spectral efficiencies of typically 1 bit/s/Hz at the noise margins required on radio links.

A great deal of research and development has been done in trying to find better modulation schemes. As people became more aware of the spectrum congestion, better systems were designed and built. One such system is the MDR-11 digital radio sytem designed by the Collins Radio Group [2]. This system employs 8-PSK modulation with further band-limiting or filtering to satisfy the FCC emission requirement. This technique was then modified to achieve a bandwidth efficiency of 3 bits/s/Hz [3]. Another approach to improving the spectral efficiency was employed by the GTE systems [4]. This system uses a 7-level correlated FM, to achieve an efficiency of approximately 2 bits/s/Hz. A third approach was taken by Bell Northern Research in designing

the DRS-8 digital radio system [5,6,7]. The Canadian DRS-8 digital radio system was designed and built to replace the analog radio system, and uses Partial Response techniques combined with PSK. Economic viability plays a great role in upgrading already existing equipment. This is why the system was designed to use the existing sites and horn reflector antennae of the analog radio system. The DRS-8 system operates at a rate of 91 Mb/s in a 41 MHz RF channel, or with a bandwidth efficiency of approximately 2.25 bits/s/Hz.

In this thesis, an extension of the DRS-8 system is considered, with the objective of achieving an even higher bandwidth efficiency. Theoretically, this new system will have a spectral efficiency of 4 bits/s/Hz. In practice, further filtering or bandlimiting is possible to achieve an efficiency slightly greater than that of 4 bits/s/Hz. The proposed system will be able to operate at a data rate of 180 Mb/s in the same 41 MHz RF bandwidth as the DRS-8 system, or at rate of 90 Mb/s in half the bandwidth - 20 MHz. As will be seen later, this change in data rate and improved spectral efficiency can be achieved without changing the basic repeater structure.

1.1 SCOPE OF THESIS

The thesis reports on a feasibility study and conceptual design of a 4 bits/s/Hz digital radio system as mentioned earlier. Chapter 2 presents the analysis of the basic system idea. Next, the different methods for generating the signal are considered, after which, a brief description of the receiver structure is given. Finally, the theoretical signal to noise requirements and error rate performance are given. Chapter 3 analyzes a decision directed carrier recovery loop. Simulation results for the carrier recovery loop are also presented. Chapter 4 presents the basic concepts of a timing recovery loop. Detailed simulation results are presented. Simulations have been done to evaluate both the steady state performance and the acquisition behaviour of the timing recovery loop. Chapter 5 summarizes the new digital radio system concepts and gives suggestions for future work.

CHAPTER 2

SYSTEM DESCRIPTION

The system considered here uses a 4-level quadrature amplitude modulation (QAM) Duobinary (class I Partial Response) filtered signal to achieve a spectral efficiency of 4 bits/s/Hz. This technique is better known as Quadrature Partial Response Signalling (QPRS). It combines two methods to improve the spectral efficiency. The first method is to reshape the spectrum of the transmitted signals, so as to concentrate their energy in a smaller frequency band. Secondly, a multi-level signal set is used to increase the information contained in one symbol. This is also a suppressed carrier type modulation, hence no power is wasted in transmitting the carrier which carries no information about the signal.

First, a brief discussion will be presented explaining why this particular type of modulation is chosen in favour of others. Next, a brief description of the transmitter and receiver structure will be presented. Finally, the error performance of the Quadrature Partial Response Signalling scheme will be evaluated.

2.1 SIGNALLING SCHEME DESCRIPTION

There has been considerable debate over the merits of various modulation schemes for a digital microwave radio system for achieving 2 bits/s/Hz [2,5]. The primary candidates are heavily filtered 8-PSK and binary QPRS. These two schemes provide comparable performance at the 2 bits/s/Hz level.

If the spectral width is taken to be between the first nulls of the $\sin^2 x/x^2$ shape spectrum, typical of phase shift keying, the PSK approach requires a 256-PSK modulation, to achieve 4 bits/s/Hz. When the MDR-11 approach [2] is considered, a 64-PSK signal is required to achieve 4 bits/s/Hz. From both a performance point of view, and the realization point of view, this is clearly unacceptable. If the QPRS approach is followed, a 4-level QAM signal with Duobinary shaping is required to achieve 4 bits/s/Hz.

2.1.1 Partial Response and Duobinary Signalling

It is well known that ideal Pulse Amplitude Modulation (PAM) with minimum Nyquist bandwidth filtering cannot be implemented in practice because of its intolerance to timing error. In order to overcome this problem, non-minimum bandwidth filters have been used. One class of such filters which has been used in practice is the so-called raised-cosine filters [9]. However, Nyquist rate signalling

can no longer be achieved. Partial Response Signalling makes it possible to transmit at the Nyquist rate again, by introducing some intersymbol interference (ISI). By reshaping the signal spectrum, some controlled intersymbol interference is introduced. Knowing the exact amount of ISI introduced, the ISI can be removed at the receiver and the signal can be decoded without difficulty. The first PRS scheme was introduced by Lender [10]. This is known as Duobinary [10] or class I Partial Response [11]. This scheme trades off symbol timing intolerance for higher required signal to noise ratio (SNR) to achieve equal performance.

A Duobinary signalling system can be characterized by the encoding equation:

$$y_n = a_n + a_{n-1} \tag{2.1}$$

where $\{y_n\}$ are the encoded symbols

and $\{a_n\}$ are the input symbols.

If the input symbols are binary, the encoded symbols have three levels which must be distinguished at the receiver instead of two. This is why a higher SNR is required to achieve the same performance as in ideal pulse amplitude modulation.

2.1.2 Generation of Partial Response Signals

From equation (2.1), it is clear that a Duobinary encoder consists of a single tapped delay line, or transversal filter structure with a filter in cascade to limit the spectrum because of the periodic nature of the frequency response of the transversal filter (Figure 2.1). In general, any Partial Response signal can be generated in this way, only that more taps are required, and also non-unity feedback might be required. In Figure 2.1, T is the sampling interval and $G(\omega)$ is an ideal rectangular filter with frequency response:

$$G(\omega) = \begin{cases} 1 & |\omega| < \frac{\pi}{T} \\ 0 & \text{otherwise} \end{cases} \qquad (2.2)$$

It has been shown that [12] the structure shown in Figure 2.1 can be replaced by a filter (Figure 2.2) with frequency response $H(\omega)$, where:

$$H(\omega) = \begin{cases} 2\,T\cos(\omega T/2) & |\omega| < (\pi/T) \\ 0 & \text{otherwise} \end{cases} \qquad (2.3)$$

Both methods can be used to generate the Partial Response signal. However, the first approach is not practical for microwave radio applications because of the nonlinear characteristics of the traveling wave tube

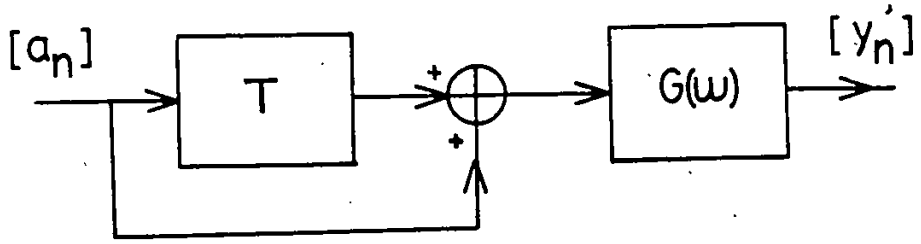


FIGURE 2.1 ENCODER STRUCTURE OF A DUOBINARY SYSTEM

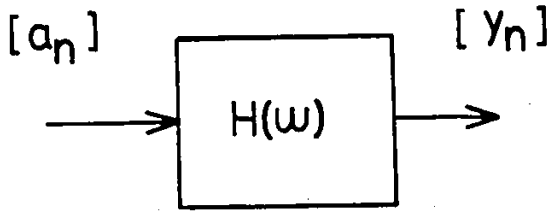


FIGURE 2.2 EQUIVALENT ENCODER STRUCTURE FOR A DUOBINARY SYSTEM

amplifier (TWTA) in each repeater. Since the QPR signal has multiple amplitudes, the TWTA has to operate with a considerable amount of backoff to achieve the linear amplification the QPR signal requires, or else the signal will be distorted to an unacceptable level. From the power efficiency point of view, the TWTA should be operated in the saturation mode. This problem can be overcome by using the second method (Figure 2.3).

As long as the RF bandpass filter has the same equivalent baseband or low-pass characteristics, the resultant waveform is the same. In the next section, the generation of the 16-QAM signal will be described.

In order to further the analysis, the following assumptions about the system model will be made. In practice the pulse shaping will be split in some optimal fashion between transmitters and receivers, so as to meet the bandwidth allocation mask specified by various government agencies. Physical realizability also requires the pulse shaping to be split between the transmitter and receiver. But for ease of analysis in this thesis, all the pulse shaping will be assumed to be done at the transmitter. At the receiver, the only filter involved is an ideal zonal filter to limit out of band noise. The channel is assumed to be ideal except for additive white Gaussian noise with two-sided spectral density of $N_o/2$ W/Hz. The data sequence

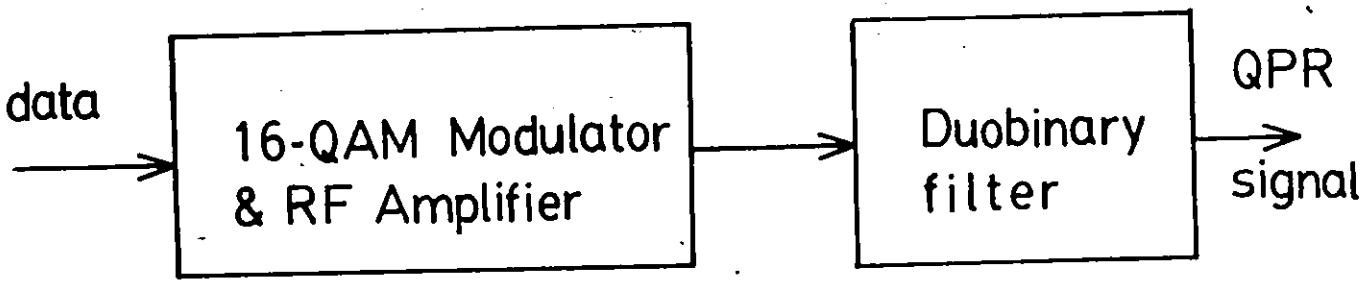


FIGURE 2.3 GENERATION OF A QPR SIGNAL USING RF SHAPING

$\{a_n\}$ are 4-level narrow pulses with amplitudes ± 1, ± 3. Figure 2.4 only shows the baseband system model. The modulator and demodulator have been left out, because PAM is a linear modulation method, and its performance can be deduced from its equivalent baseband system.

2.2 TRANSMITTER STRUCTURE

QAM signals usually have amplitude variation except in the case when there are only two data levels ± 1. Therefore straightforward power amplification cannot be done using a nonlinear TWTA, without backoff. However there is a simple way to overcome this as shown by Miyauchi et al [13]. This method is based on the superposition of two constant envelope QPSK signals. The generation of the 16-QAM signals is best illustrated in Figure 2.5. It consists of summing two coherent QPSK signals, one of which has an amplitude which is twice as large as the other. Many other multi-amplitude signal sets can be generated using a similar approach. This superposition technique has been employed for generating the multi-amplitude minimum shift keying signal [14] as well as other M-level amplitude phase shift keying signals [15].

With the 16-QAM signal in RF, the Duobinary shaping can be done without difficulty by a bandpass filter with a corresponding baseband filter response as described in

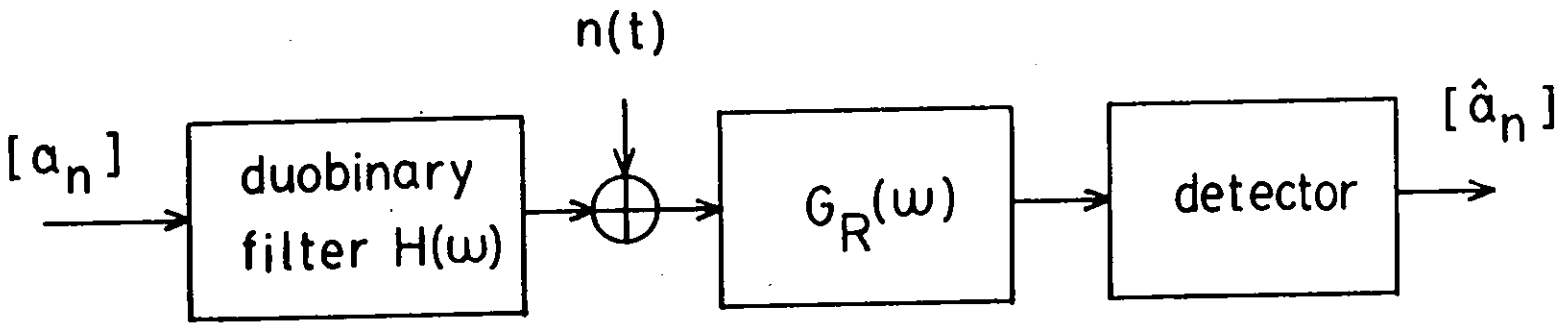


FIGURE 2.4 BASEBAND SYSTEM MODEL

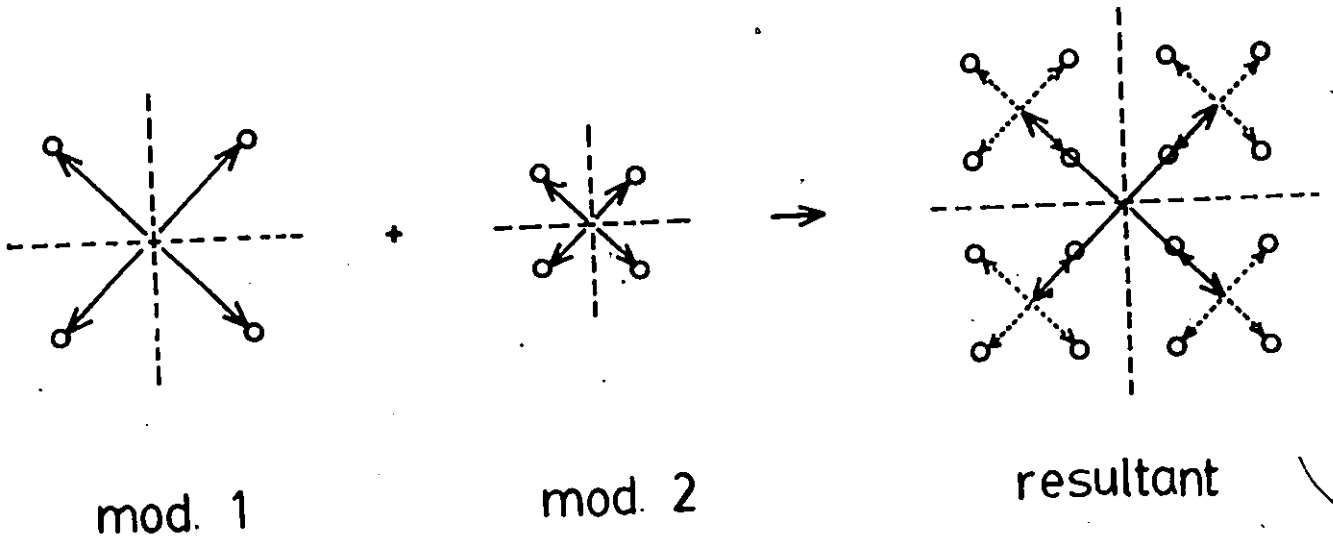


FIGURE 2.5 16-QAM GENERATION SCHEME

Equation (2.3). A block diagram of the resulting QPRS transmitter is shown in Figure 2.6, showing the modulation method and the pulse shaping filter for the QPR system [8]. With this configuration, one bandpass shaping filter is required instead of the two if shaping is done at baseband. However, this filter has to operate at microwave frequencies. The two traveling wave tube amplifiers have to be matched and have the same saturated gain characteristics. The output of the modulators 1 and 2 can be expressed in any signalling interval as:

$$s_1(t) = \sqrt{\frac{E_s}{T}} \left[a_1 \cos \omega_c t - b_1 \sin \omega_c t \right] \qquad (2.4)$$

$$s_2(t) = \sqrt{\frac{E_s}{T}} \left[a_2 \cos \omega_c t - b_1 \sin \omega_c t \right] \qquad (2.5)$$

where a_1, a_2, b_1, b_2 are input binary symbols having values ± 1, and ω_c is the angular carrier frequency. After the two signals are amplified, they are added with an amplitude ratio of 1:2 or a power ratio of 1:4. The resultant signal is:

$$s'(t) = 2 s_1(t) + s_2(t)$$

$$= \sqrt{\frac{E_s}{T}} \left[(2a_1 + a_2) \cos \omega_c t - (2b_1 + b_2) \sin \omega_c t \right] \qquad (2.6)$$

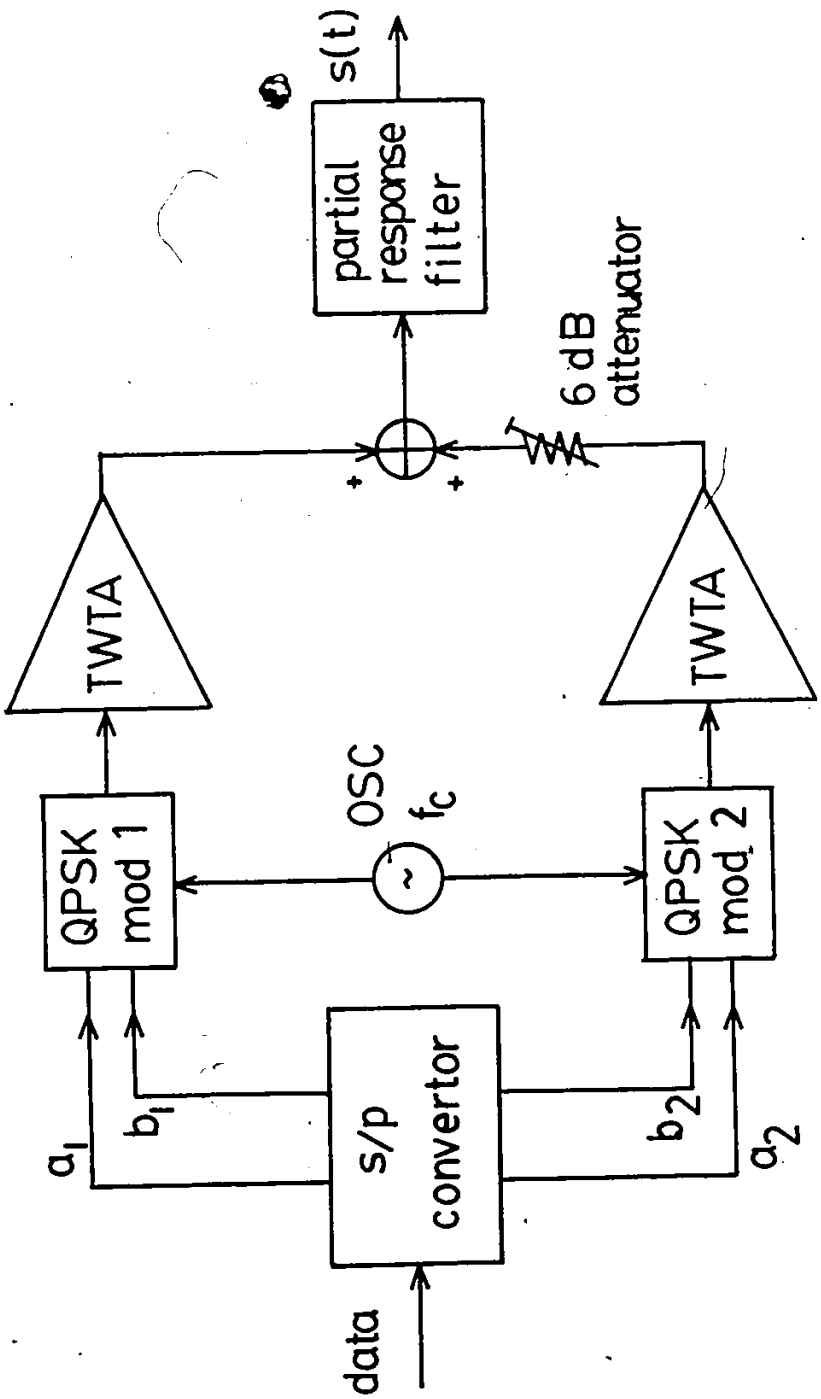


FIGURE 2.6 TRANSMITTER STRUCTURE

Since s_1 and s_2 both have constant envelope, they can be amplified separately using TWTA's operating at saturation to obtain the maximum possible power efficiency.

2.3 RECEIVER STRUCTURE DESCRIPTION

Coherent demodulators are the most efficient in performance and are almost always used in high speed data transmission systems. QPR and QAM are suppressed carrier modulations, and for coherent demodulation, a phase coherent carrier must be extracted from the received signal. A decision directed loop structure will be used in the system under consideration. This loop structure uses estimates of the data symbols to remodulate the incoming signal, so as to remove the signal component, and to reconstruct a spectral line at the carrier frequency. With the coherent carrier, the received signal can then be coherently demodulated to baseband where further decoding and processing can be carried out. Each of the in-phase and quadrature-phase baseband signals is a Duobinary encoded signal having seven encoded levels. If these are sampled at the correct times, the ISI is controlled and can be subtracted out by using previous correct decisions. However, if precoding [10] is used at the transmitter, modulo-4 arithmetic has to be used for decoding. If no precoding is done at the transmitter, the decoder ideally consists of only one delay element and a

4-level slicer as shown in Figure 2.7.

In practice, better spectral efficiency can be obtained by further bandlimiting the signal spectrum. However this will introduce more ISI. Also when conventional Chebyshev filters are used to approximate the Duobinary filter response, a considerable amount of additional ISI will be introduced. When the ideal Duobinary response cannot be achieved due to one reason or another, more decision feedback taps can be added to equalize or suppress the trailing echoes [7]. This second approach will be taken because in practice the ideal response cannot be achieved. However a new problem arises, namely error propagation, in the decision feedback decoder. This will be analyzed in the next section under the assumption of ideal Duobinary filtering. The decision feedback decoder will be used within the carrier recovery loop to provide estimates of the data used for extracting the phase information of the carrier. This loop structure will be analyzed in the next chapter.

The decoding of the Duobinary signal requires a knowledge of the correct sampling times. This information has to be extracted from the received signal as well. Due to the special characteristic of the Duobinary signal, ordinary threshold crossings or signal derivative methods [31,32], will not provide any information about the sampling

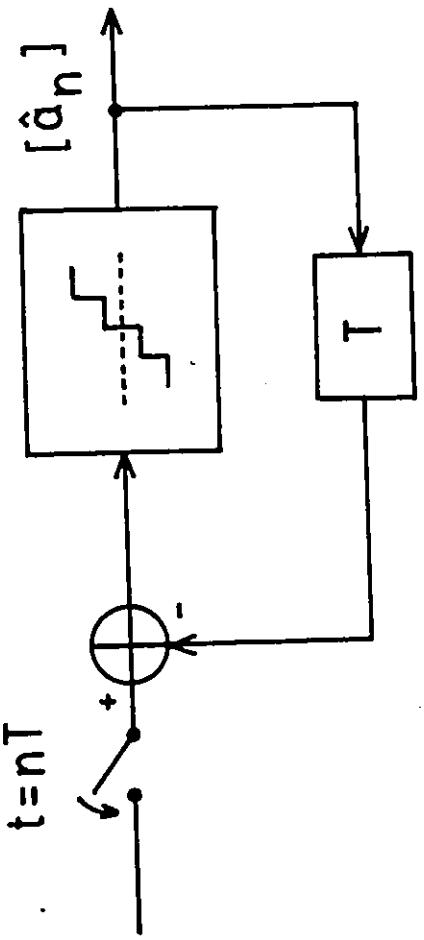


FIGURE 2.7 DECODER FOR A DUOBINARY SYSTEM ASSUMING IDEAL FILTERING (DECISION FEEDBACK EQUALIZER)

times. Some nonlinear process such as the 'slicer' method used in modified Duobinary systems [16] has to be used to extract the timing information. This problem will be discussed in Chapter 4.

2.4 ERROR PERFORMANCE ANALYSIS

If a decision feedback equalizer is used for decoding, there will be a degradation in performance due to the effect of error propagation. Error propagation arises because previous estimates of data symbols are used to calculate the ISI to be subtracted from the present sample. However a lower bound on the error performance can be obtained by ignoring error propagation [17] as:

$$P_{eL} = 2 \left(1 - \frac{1}{K} \right) Q[d/\sigma] \qquad (2.7)$$

where K is the number of input levels

$2d$ is the distance between input levels

σ^2 is the noise variance

and $Q[x] = \frac{1}{(2\pi)^{1/2}} \int_x^\infty \exp(-u^2/2)\, du$

In order to include the effects of error propagation, the system can be modelled as a Markov Chain [17] and the exact probability of error can then be evaluated. The decision feedback decoder is modelled as a Markov Chain with states corresponding to the different levels of error,

including the no error state. The transition probability from one error state to another can then be written out explicitly. This is known as the state transition matrix. The probability of each state at time k, $\underline{S}^k$ is related to the previous state probabilities by the state transition matrix (STM) by:

$$\underline{S}^k = [STM]\ \underline{S}^{k-1} \tag{2.8}$$

The stationary probabilities of all the error states can be found numerically without difficulty. Each of these stationary probabilities, describes the probability of occurrence of each state on the average. Therefore, the probability of error, including the effect of error propagation, can be evaluated by summing up all the stationary probabilities corresponding to error states not equal to zero. The detailed evaluation is given in the appendix, and the error performance curves for the decoder of Figure 2.7 are plotted in Figure 2.8. In this figure, the performance curves for 2-level and 4-level ideal PAM are also included for comparison. Also included is the Partial Response curve for binary inputs. In making this comparison, the transmit power P_s is kept constant for all cases. The average SNR for the PAM case is defined as:

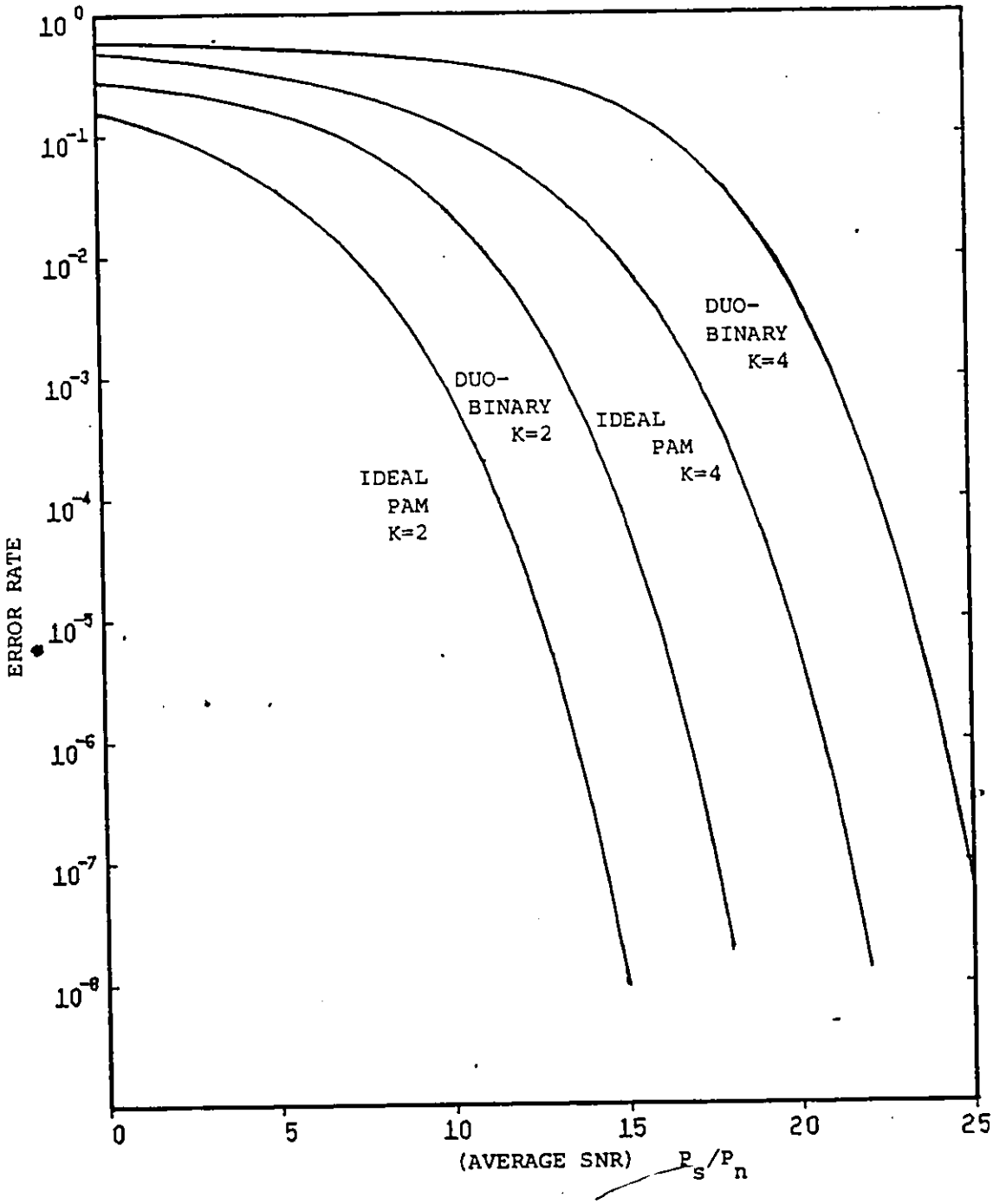


FIGURE 2.8 ERROR PROBABILITY FOR DUOBINARY SYSTEMS INCLUDING ERROR PROPAGATION EFFECT

$$\frac{P_s}{P_n} = \frac{(K^2 - 1)\ \Delta^2}{3} \qquad (2.9)$$

whereas the average SNR for the Duobinary case is defined as:

$$\frac{P_s}{P_n} = \frac{2(K^2 - 1)\cdot\Delta^2}{3}$$

where $\Delta = d/\sigma$.

From Figure 2.8, we can see that in the 4-level case, an error rate of 10^{-5} can be obtained at a reasonable SNR, approximately 23.5 dB. When compared with 4-level PAM, there is only a 3 dB degradation, however the bandwidth requirement is much less in the Duobinary case. An increase of 7.5 dB in SNR can bring the efficiency from 2 bits/s/Hz to 4 bits/s/Hz, when 2-level Duobinary is converted to 4-level. Furthermore, the effect of error propagation is small at reasonably high SNR. At an error rate of 10^{-5}, the degradation due to error propagation is less than 0.9 dB. Therefore, a decision feedback decoder can be used without too much degradation. In the case when real filtering is used to achieve the Partial Response spectrum as in the DRS-8 system, decision feedback decoding must be used. The degradation due to error propagation in this latter case will be very little different from the results for the ideal decoder [22].

CHAPTER 3

CARRIER RECOVERY LOOP ANALYSIS

Coherent detection is known to perform much better than incoherent detection by 3 dB [28]. However coherent detection requires a locally generated carrier which is in phase with the received modulated carrier. This phase information has in most cases to be extracted from the received signal. The QPR system under consideration is a suppressed carrier system. There is no pilot carrier available at the receiver. But the phase information can be recovered easily using a decision-directed phase-locked loop structure as shown in Figure 3.1. A decision directed loop is chosen because in general it gives significant improvement in tracking performance over conventional phase-locked loops [25, p 64]. The structure proposed here is similar to that proposed by Simon and Smith for quadrature amplitude shift keying systems [18], except that it is modified to include the decision feedback decoder within the loop structure so as to provide estimates of the data symbols.

In this structure, an error voltage is produced by cross-correlating the in-phase (I) channel baseband samples

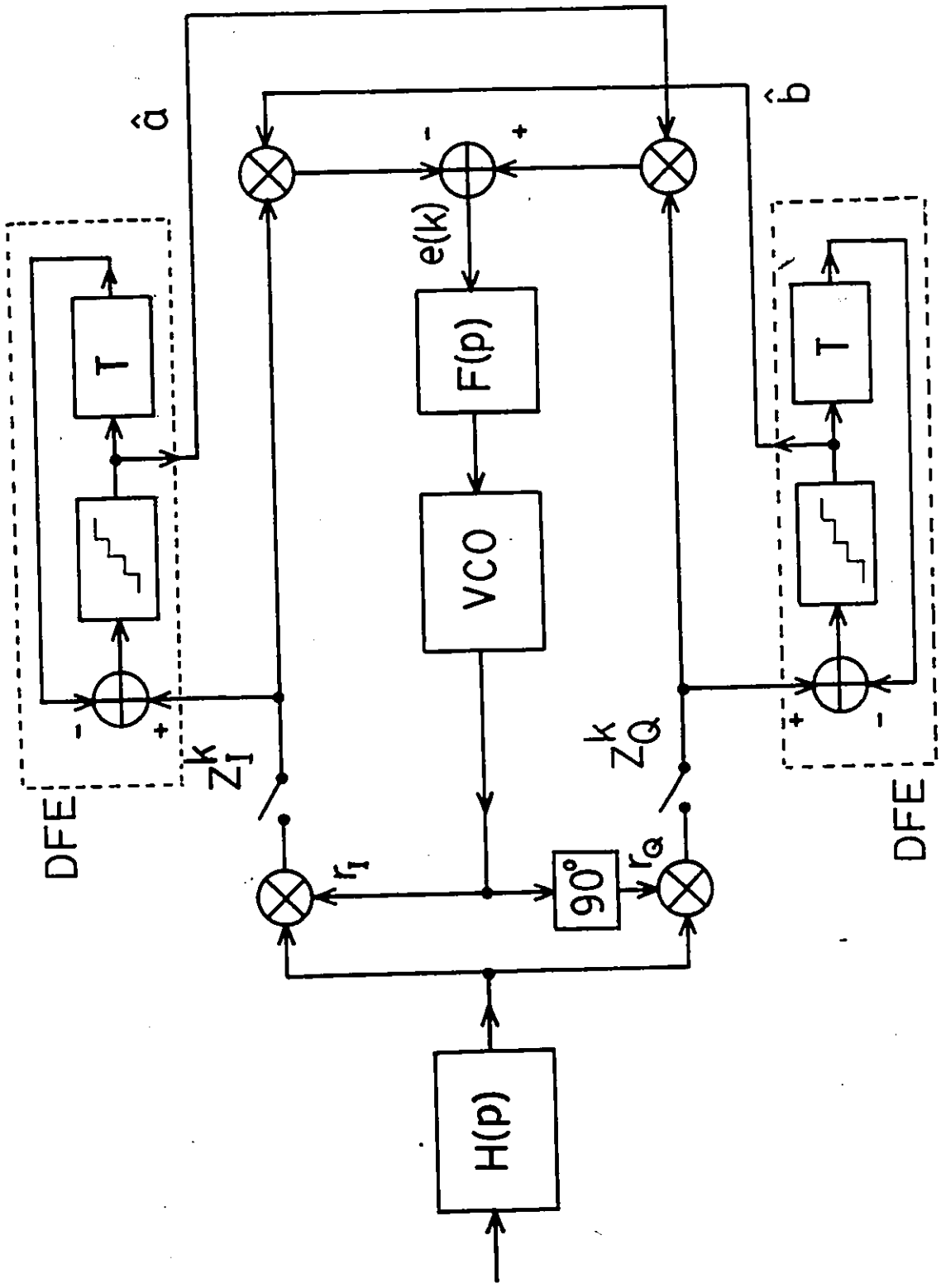


FIGURE 3.1 CARRIER RECOVERY LOOP STRUCTURE FOR A DUOBINARY MODULATED SYSTEM (4 level data)

with the estimates of the quadrature (Q) channel signal and vice versa. These products are used to generate an error voltage which after further averaging is used to drive a voltage controlled oscillator (VCO) to correct the phase error between the locally generated carrier and the received modulated carrier. The following analysis will parallel that of Simon and Smith [18, 19] and also Taylor and Cheung [20], who investigated a similar loop structure for offset QPRS signals.

3.1 DEVELOPMENT OF STOCHASTIC INTEGRO-DIFFERENTIAL EQUATION FOR LOOP STRUCTURE

The received signal is first bandpass filtered to remove out of band noise. The bandwidth W_i of the filter $H(p)$ should be sufficiently wide that the signal component will pass through without distortion. At the output of the bandpass filter, the noise process can be considered as a narrow band Gaussian process. That is, at the output of the filter the signal can be expressed as:

$$x(t) = s\,[t, \theta(t)] + n(t) \tag{3.1}$$

$$\text{where } n(t) = \sqrt{2}\,\{\, n_c(t) \cos [\omega_o t + \theta(t)] - n_s(t) \sin [\omega_o t + \theta(t)]\,\}$$

This signal is then multiplied by the two locally generated signals $r_I(t)$ and $r_Q(t)$ where r_I and r_Q are in

quadrature, and may be represented by:

$$r_I(t) = \sqrt{2}\, K_1 \cos [\omega_o t + \hat{\theta}(t)]$$
$$r_Q(t) = -\sqrt{2}\, K_1 \sin [\omega_o t + \hat{\theta}(t)] \qquad (3.2)$$

where K_1 is the voltage controlled oscillator RMS amplitude and,

$\hat{\theta}(t)$ is the local estimate of $\theta(t)$.

After multiplying by the two reference signals, we obtain Z_I and Z_Q, the I channel and Q channel baseband signals. Ignoring the double frequency terms, Z_I and Z_Q can be written as:

$$Z_I(t) = K_m\, x(t)\, r_I(t)$$
$$= K_1 K_m \{ \sum_n a_n h(t-nT) \cos \phi(t) -$$
$$\sum_n b_n h(t-nT) \sin \phi(t) + N_I [t, \phi(t)]\} \qquad (3.3a)$$

$$Z_Q(t) = K_m\, x(t)\, r_Q(t)$$
$$= K_1 K_m \{ \sum_n a_n h(t-nT) \sin \phi(t) +$$
$$\sum_n b_n h(t-nT) \cos \phi(t) + N_Q[t, \phi(t)]\} \qquad (3.3b)$$

where $\phi = \theta(t) - \hat{\theta}(t)$ is the phase error

K_m is the multiplier gain

$\{a_n\}$ $\{b_n\}$ are statistically independent data symbols

$h(t)$ is the Duobinary pulse shape.

Also N_I and N_Q are defined as follows:

$$N_I\ [t, \phi(t)] = n_c(t) \cos \phi(t) - n_s(t) \sin \phi(t) \tag{3.4a}$$

$$N_Q\ [t, \phi(t)] = n_c(t) \sin \phi(t) + n_s(t) \cos \phi(t) \tag{3.4b}$$

The low pass signals Z_I and Z_Q are then sampled at the times $t = nT$ (n=integer) to drive the decision feedback equalizer (DFE) and to cross-correlate with the quadrature channel data estimates so as to obtain the error signal $e(t)$.

Assuming sampling at the correct times, the sampled values of Z_I and Z_Q namely Z_I^k and Z_Q^k can be written as:

$$Z_I^k = K_1 K_m \{(a_k+a_{k-1}) \cos \phi(t) - (b_k+b_{k-1}) \sin \phi(t) + N_I[t, \phi(t)]\} \tag{3.5a}$$

$$Z_Q^k = K_1 K_m \{(a_k+a_{k-1}) \sin \phi(t) + (b_k+b_{k-1}) \cos \phi(t) + N_Q[t, \phi(t)]\} \tag{3.5b}$$

The output from the DFE's are $\hat{a}_k$ and $\hat{b}_k$. They are the estimates of the actual data symbols a_k and b_k. The estimates $\hat{a}_k$ and $\hat{b}_k$ can be expressed in the following form:

$$\hat{a}_k = Q_K [\ Z_I^k - \hat{a}_{k-1}\ ] \tag{3.6a}$$

$$\hat{b}_k = Q_K \left[z_Q^k - \hat{b}_{k-1} \right] \tag{3.6b}$$

where Q_K is the quantization characteristics of the slicer. From the above, the sampled error signal (cf. Figure 2.1) can be written in the following form:

$$\begin{aligned} e(kT) &= \hat{a}_k z_Q - \hat{b}_k z_I \\ &= K_1 K_m \{ [\hat{a}_k (a_k + a_{k-1}) + \hat{b}_k (b_k + b_{k-1})] \sin \phi \\ &\quad + [\hat{a}_k (b_k + b_{k-1}) - \hat{b}_k (a_k + a_{k-1})] \cos \phi \\ &\quad + \hat{a}_k N_Q[t, \phi] - \hat{b}_k N_I[t, \phi] \} \end{aligned} \tag{3.7}$$

The relationship between the instantaneous frequency $\dot{\hat{\theta}}(t)$ of the VCO output and the error signal e(t) is given by:

$$\dot{\hat{\theta}}(t) = K_v F(p) e(t) \tag{3.8}$$

where K_v is the VCO gain in radians/V/s

and $\dot{\hat{\theta}}(t) = \frac{d}{dt} \hat{\theta}(t)$

where $\hat{\theta}(t)$ is the instantaneous phase of the VCO output,

or $$\begin{aligned} \dot{\hat{\theta}}(t) &= \frac{d}{dt} [\theta(t) - \phi(t)] \\ &= \dot{\theta}(t) - \dot{\phi}(t) \end{aligned}$$

Combining Equations (3.7) and (3.8), we have the stochastic loop differential equation as:

$$\dot{\phi}(t) = \dot{\theta}(t) - K_v F(p)\, e(t)$$

$$= \dot{\theta}(t) - K_o F(p) \{[\hat{a}_k(a_k+a_{k-1}) + \hat{b}_k(b_k+b_{k-1})]$$

$$\sin\phi + [\hat{a}_k(b_k+b_{k-1}) - \hat{b}_k(a_k+a_{k-1})] \cos\phi$$

$$+ \hat{a}_k N_Q[t, \phi(t)] - \hat{b}_k N_I[t, \phi(t)]\} \qquad (3.9)$$

where $K_o = K_1 K_v K_m$ is known as the open loop gain of the loop.

In order to further analyze the loop, we make the following practical assumption:

$$W_L << \frac{1}{T} << \frac{\omega_o}{2\pi}$$

where W_L is the loop bandwidth. Based on these assumptions, we can to a very good approximation [18] treat the noise process at the output of the filter $H(p)$ as narrow band white noise. Also, the phase process will vary much more slowly than the signal or the noise process. Therefore, the statistical average of the stochastic Equation (3.9) can be taken over the data to obtain:

$$\dot{\phi}(t) = \dot{\theta}(t) - K_o F(p) [\, G(\phi) + H^{1/2}(\phi)\, N_1(t)] \qquad (3.10)$$

where $N_1(t)$ is approximately white Gaussian noise of single-sided spectral density N_o W/Hz, and

$$G(\phi) = E\{[\hat{a}_k(a_k+a_{k-1}) + \hat{b}_k(b_k+b_{k-1})]|\phi\} \sin\phi + E\{[\hat{a}_k(b_k+b_{k-1}) - \hat{b}_k(a_k+a_{k-1})]|\phi\} \cos\phi \quad (3.11)$$

$$H(\phi) = E[\hat{a}_k^2|\phi(t)] + E[\hat{b}_k^2|\phi(t)] \quad (3.12)$$

$G(\phi)$ is known as the loop phase detector characteristic or S-curve, and $H(\phi)$ is known as the noise function. Next, for ease of comparison, $G(\phi)$ is normalized to have unity slope at $\phi = 0$, that is:

$$g(\phi) = G(\phi)/G'(\phi) \quad (3.13)$$

$$h(\phi) = H(\phi)/H(0) \quad (3.14)$$

These normalized characteristics will be evaluated in the next two sections.

Substituting Equations (3.13) and (3.14) into (3.10) and rewriting the integro-differential Equation (3.10) in terms of the normalized characteristics, we obtain:

$$\dot{\phi}(t) = \dot{\theta}(t) - K_o F(p) [A g(\phi) + h^{1/2}(\phi) N_1(t)] \quad (3.15)$$

where $K_1 = K_o H^{1/2}(0)$

$A = G'(0)/H^{1/2}(0)$

3.2 EVALUATION OF THE NORMALIZED LOOP PHASE DETECTOR CHARACTERISTICS (S-CURVE)

The normalized loop phase detector characteristic is defined as in Equation (3.13).

Rewriting Equation (3.13) using (3.11), we have:

$$g(\phi) = g_1 \sin \phi + g_2 \cos \phi \tag{3.16}$$

where $g_1 = E\{[\hat{a}_k(a_k+a_{k-1}) + \hat{b}_k(b_k+b_{k-1})]|\phi\}/G'(0)$

$$g_2 = E\{[\hat{a}_k(b_k+b_{k-1}) - \hat{b}_k(a_k+a_{k-1})]|\phi\}/G'(0) \tag{3.17}$$

The evaluation of the expectations will be done in the following manner:

$$\begin{aligned}
&E\{\hat{a}_k(a_k+a_{k-1})|\phi\} \\
&= \sum_{ijkmn} \hat{a}_k(a_i+a_j)\, P[\hat{a}_k a_i a_j b_m b_n|\phi] \\
&= \sum_{ijk\ell mn} \hat{a}_k(a_i+a_j)\, P[(Z_I-\hat{a}_\ell)|a_i a_j b_m b_n \hat{a}_\ell \phi] \\
&\qquad P[a_i a_j b_m b_n \hat{a}_\ell] \\
&= \sum_{ijk\ell mn} \hat{a}_k(a_i+a_j)\, P[(Z_I-\hat{a}_\ell)|a_i a_j b_m b_n \hat{a}_\ell \phi] \\
&\qquad P[\hat{a}_\ell|a_i a_j b_m b_n]\, P[a_i a_j b_m b_n] \\
&= \frac{1}{K^4} \sum_{ijk\ell mn} \hat{a}_k(a_i+a_j)\, P[(Z_I-\hat{a}_\ell)|a_i a_j b_m b_n \hat{a}_\ell \phi)
\end{aligned} \tag{3.18}$$

The evaluation of $P[\hat{a}_\ell|a_i a_j b_m b_n]$ involves a Markov process, and is extremely difficult. The following assumption will simplify the analysis:

$P[\hat{a}_\ell|a_i a_j b_m b_n] = 1/K$ for all ℓ possible cases.

This is a pessimistic assumption at high SNR, because with this assumption, the probability that $\hat{a}_\ell$ is a correct estimate of the data symbol is only $1/K$. Making use of this assumption, we have:

$$
\begin{aligned}
& E\,[\hat{a}_k(a_k+a_{k-1})\,|\,\phi] \\
&= (1/K^5) \sum_{ijk\ell mn} \hat{a}_k(a_i+a_j)\, P[Z_I - \hat{a}_\ell | a_i a_j b_m b_n \hat{a}_\ell \phi] \\
&= (1/K^5) \sum_{ij\ell mn} \{(i+j)\ \hat{a}_k\, P[Z_I - \hat{a}_\ell | a_i a_j b_m b_n \hat{a}_\ell \phi]\} \\
&= (2/K^5) \sum_{ij\ell mn} \{(i+j) \sum_{k'} Q\{[k'+\ell+(m+n)\sin\phi \\
&\qquad - (i+j)\cos\phi]/\sigma\}\}
\end{aligned}
\tag{3.20}
$$

where $i, j, k, \ell, m, n = \pm 1, \pm 3, \ldots \pm(K-1)$

$k' = 0, \pm 2, \ldots \pm(K-2)$

For 4-level data symbols, we have $K=4$, and therefore:

$$
\begin{aligned}
& E\,[\hat{a}_k(a_k+a_{k-1})\,|\,\phi] \\
&= (2/4^5) \sum_{ij\ell mn} (i+j)\ \{Q[(X-2)/\sigma] + Q[X/\sigma] \\
&\qquad + Q[(X+2)/\sigma]\}
\end{aligned}
\tag{3.21}
$$

where $X = (m+n)\sin\phi - (i+j)\cos\phi + \ell$

Similarly, the other expectations can be evaluated:

$$
E\,[\hat{b}_k(b_k+b_{k-1})\,|\,\phi] = E\,[\hat{a}_k(a_k+a_{k-1})\,|\,\phi] \tag{3.22}
$$

$$
E\,[\hat{a}_k(b_k+b_{k-1})\,|\,\phi] = (2/K^5) \sum_{ij\ell mn} \{(m+n) \{ \sum_{k'} Q\{[k'+X)/\sigma]\}\} \tag{3.23}
$$

and $$E\,[\hat{b}_k(a_k+a_{k-1})\,|\,\phi] = -\,E\,[\hat{a}_k(b_k+b_{k-1})\,|\,\phi] \qquad (3.24)$$

Substituting Equations (3.20), (3.22) to (3.24) into (3.11) and evaluating for 4 level data symbols, we have:

$$\begin{aligned} G(\phi) = {} & (1/4^4) \sum_{ij\ell mn} (i+j)\,\{Q[(X-2)/\sigma] + Q[X/\sigma] \\ & + Q[(X+2)/\sigma]\}\sin\phi \\ & + (1/4^4) \sum_{ij\ell mn} (m+n)\,\{Q[(X-2)/\sigma] + Q[X/\sigma] \\ & + Q[(X+2)/\sigma]\}\cos\phi \end{aligned} \qquad (3.25)$$

Next, the slope of $G(\phi)$ is evaluated as:

$$\begin{aligned} G'(\phi) = {} & \frac{d}{d\phi} G(\phi) \\ = {} & \sin\phi\,\frac{d}{d\phi}\{[(4/K^5) \sum_{ij\ell mn} (i+j) \sum_{k'} Q[(k'+X)/\sigma]\} \\ & + (4/K^5) \sum_{ij\ell mn} (i+j) \sum_{k'} Q[(k'+X)/\sigma]\cos\phi \\ & + \cos\phi\,\frac{d}{d\phi}\{(4/K^5) \sum_{ij\ell mn} (m+n) \sum_{k'} Q[(k'+X)/\sigma]\} \\ & - \sin\phi\,(4/K^5) \sum_{ij\ell mn} (m+n) \sum_{k'} Q[(k'+X)/\sigma] \end{aligned} \qquad (3.26)$$

from which, after some manipulation, we obtain:

$$\begin{aligned} G'(0) = {} & (4/K^5)\{ \sum_{ij\ell mn} (i+j) \sum_{k'} Q[(k'+X)/\sigma]\}\big|_{\phi=0} \\ & + (4/K^5)\,\frac{d}{d\phi}\{ \sum_{ij\ell mn} (m+n) \sum_{k'} Q[(k'+X)/\sigma]\}\big|_{\phi=0} \end{aligned} \qquad (3.27)$$

Since $$Q[u] = -\frac{1}{\sqrt{(2\pi)}} \exp(-u^2/2) \frac{du}{d\phi}$$

we have,

$$G'(0) = (4/K^5) \sum_{ij\ell mn} (i+j) \sum_{k'} Q[(k'+Y)/\sigma]$$

$$+ (4/K^5) \sum_{ij\ell mn} (m+n) \left[\frac{-1}{\sqrt{(2\pi)}}\right]$$

$$\left\{ \sum_{k'} \exp[-(k'+X)^2/2\sigma^2] \frac{d}{d\phi} \frac{(k'+X)}{\sigma} \right\}\Big|_{\sigma=0} \quad (3.28)$$

where $Y = X|_{\phi=0} = \ell - i - j$

Before G'(0) can be evaluated, the following has to be done:

$$\frac{d}{d\phi} \frac{(k'+X)}{\sigma} = \frac{d}{d\phi} [k'+(m+n) \sin\phi - (i+j) \cos\phi + \ell]/\sigma$$

$$= [(m+n) \cos\phi + (i+j) \sin\phi]/\sigma \quad (3.29)$$

Hence, we obtain:

$$G'(0) = (4/K^5) \sum_{ij\ell mn} (i+j) \sum_{k'} Q[(k'+Y)/\sigma]$$

$$+ (4/K^5) \left[-\frac{1}{\sqrt{(2\pi)}}\right] \sum_{ij\ell mn} (m+n)$$

$$\sum_{k'} \exp[-(k'+Y)^2/2\sigma^2] [(m+n)/\sigma] \quad (3.30)$$

Rewriting Equation (3.30) in a more compact form,

$$G'(0) = (4/K^5) \sum_{ij\ell mn} (i+j) \sum_{k'} Q[(k'+Y)/\sigma]$$

$$-\left[\frac{4}{\sqrt{(2\pi)}\sigma K^5}\right] \sum_{ij\ell mn} (m+n)^2 \sum_{k'} \exp[(-k'+Y)^2/2\sigma^2]$$

(3.31)

On further manipulation, we obtain:

$$G'(0) = (4/K^5) \sum_{ij\ell mn} (i+j)\{\sum_{k'} Q[(k'+Y)/\sigma]\}$$

$$-\{8(K^2-1)/[3K^3\sigma\sqrt{2\pi}\,]\} \sum_{ij\ell k'} \exp[-(k'+Y)^2/2\sigma^2] \tag{3.32}$$

The normalized phase detector characteristic is then evaluated using a digital computer and the results are plotted in Figure 3.2 for 4-level data symbols (7-level encoded data) for different values of Δ. The S-curve has a period of $\pi/2$, and has odd symmetry about $\phi = 0$. This means that a four-fold ambiguity exists in the phase detector characteristics. This type of four-fold ambiguity can be expected by just looking at the quadrature amplitude shift keying signal set. A rotation of $\pi/2$ radians produces no change in the constellation of the signal set. Although there is no difference in the constellation, for correct decoding of the data symbols, this quadrant ambiguity has to be resolved. This quadrant ambiguity can be resolved by using proper coding techniques or by using a special training sequence to ensure that the loop locks at the correct phase during startup.

As described before, the S-curve has odd symmetry about $\phi = 0$. This will be the stable lock point of the carrier recovery loop. However, there are a number of undesirable or false lock points lying between approximately

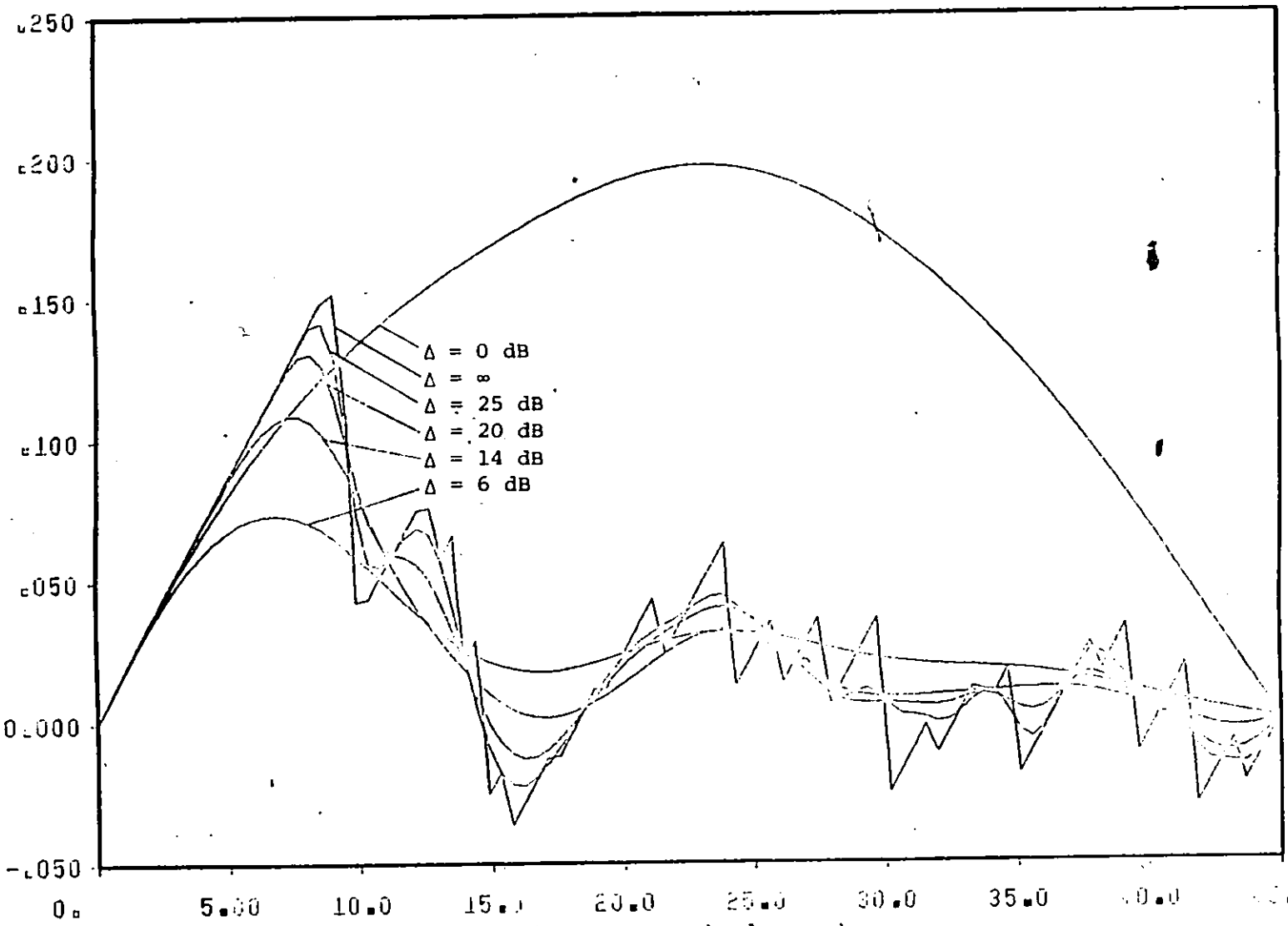


FIGURE 3.2 Normalized Phase Detector Characteristics for Carrier Recovery Loop (4 level data)

15 degrees and 45 degrees. These false lock points will create problems during initial phase acquistion or start-up of the loop. However, a special training sequence will enable the loop to lock at the correct lock point $\phi = 0$. With a sufficiently narrow band loop, the loop will remain locked with very high probability.

It is also of interest to note that the loop structure is identically the same when the data symbols are binary. In this case, the S-curve reduces to:

$$G(\phi) = (1/2^3) \sum_{ij\ell mn} \{[(i+j)\sin\phi + (m+n)\cos\phi] Q[X/\sigma]\} \tag{3.34}$$

and

$$G'(0) = (1/8) \{ \sum_{ij\ell mn} (i+j)\, Q[Y/\sigma] - \sum_{ij\ell} [\frac{8}{\sigma\sqrt{(2\pi)}}] \exp(-Y^2/2\sigma^2)\} \tag{3.34}$$

The phase detector characteristic for binary data symbols is then shown in Figure 3.3. Again the detector characteristic has a four-fold ambiguity. However, there is only one false lock point at approximately 30 degrees. This provides a larger possible lock-in range for initial acquisition, and also a higher probability of lock-in, and suggests that initial acquisition be accomplished using binary transmission.

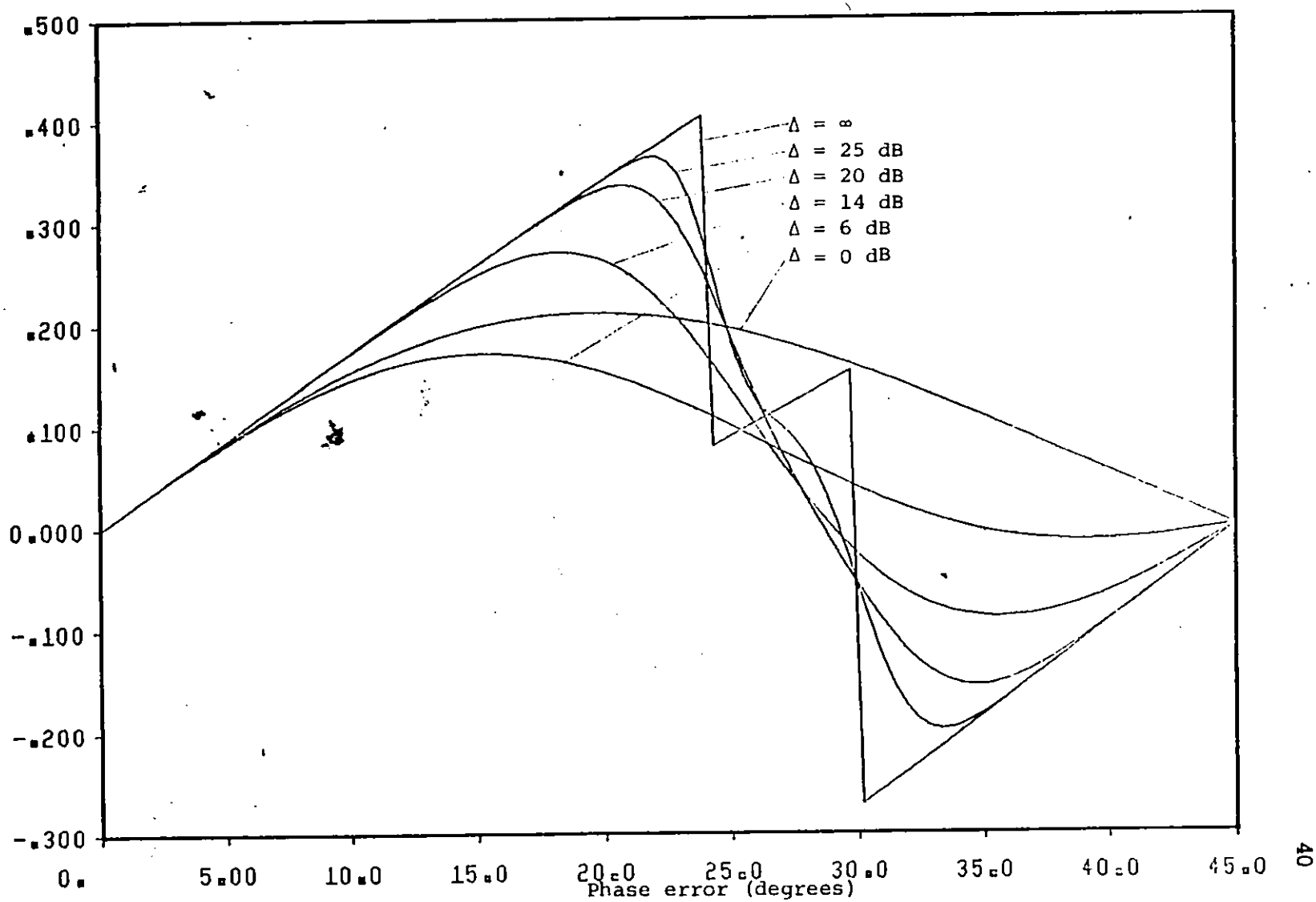


FIGURE 3.3 NORMALIZED PHASE DETECTOT CHARACTERISTIC FOR CARRIER RECOVERY LOOP (2 level data)

3.3 EVALUATION OF THE NORMALIZED NOISE FUNCTION

The noise function as defined in Equation (3.12) is

$$H(\phi) = E\,[\hat{a}_k^2|\phi] + E\,[\hat{b}_k^2|\phi] \tag{3.12}$$

This is then normalized to have unity value at $\phi = 0$. That is:

$$h(\phi) = H(\phi)/H(0) \tag{3.14}$$

The expectations $E[\hat{a}_k^2]$ and $E[\hat{b}_k^2]$ are evaluated as follows:

$$E[\hat{a}_k^2|\phi] = \sum_{ijk\ell mn} \hat{a}_k^2\, P[\hat{a}_k\hat{a}_\ell a_i a_j b_m b_n|\phi]$$

$$= \sum_{ij\ell mnk} \hat{a}_k^2\, P[\hat{a}_k|\phi\hat{a}_\ell a_i a_j b_m b_n]\; P[\hat{a}_\ell a_i a_j b_m b_n] \tag{3.35}$$

Using the same assumption (Equation 3.19) as in the evaluation of the S-curve, we have:

$$E[\hat{a}_k^2|\phi] = (1/K^5) \sum_{ij\ell mn} \{ \sum_{k'} 4k'\, Q[(k'+X)/\sigma] + (K-1)^2 \} \tag{3.36}$$

Similarly,

$$E[\hat{a}_k^2|\phi] = E[\hat{b}_k^2|\phi] \tag{3.37}$$

Hence, we obtain,

$$H(\phi) = (2/K^5) \sum_{ij\ell mn} \{ \sum_{k'} 4k'\, Q[(k'+X)/\sigma] + (K-1)^2 \} \tag{3.38}$$

and the normalized noise function is:

$$\frac{H(\phi)}{H(0)} = \frac{\sum_{ij\ell mn} \left\{ \sum_{k'} 4\,k'\,Q[(k'+X)/\sigma]+(K-1)^2 \right\}}{\sum_{ij\ell mn} \left\{ \sum_{k'} 4\,k'\,Q[(k'+Y)/\sigma]+(K-1)^2 \right\}} \tag{3.39}$$

The normalized noise function is plotted in Figure 3.4 for K=4, as a function of loop phase error for different values of Δ. When the data symbols are binary, that is K=2, the noise function reduces to a constant. This implies the loop nonlinearity will not change the noise statistics as a function of the phase error ϕ.

3.4 EVALUATION OF THE PROBABILITY DENSITY FUNCTION OF THE PHASE ERROR PROCESS

Another useful characteristic of a carrier tracking loop is the steady state probability density function (PDF) of the phase error process. Using the Fokker-Planck technique, the steady state PDF $P(\phi)$ of the modulo-2π reduced phase error will be evaluated in this section. To further simplify the analysis, we shall assume a first order loop with zero-detuning (that is $F(p) = 1$ and $\dot{\theta} = 0$). Therefore, $P(\phi)$ can be shown to satisfy the steady state Fokker-Planck equation [21].

$$\frac{d}{d\phi}\,[\,A_o\,P(\phi)] = \frac{1}{2}\,\frac{d^2}{d\phi^2}\,[\,B_o\,P(\phi)] \tag{3.40}$$

where $A_o(\phi) = -K_o\,G(\phi)$

$$B_o(\phi) = N_o K_o^2\,H(\phi)/2 \tag{3.41}$$

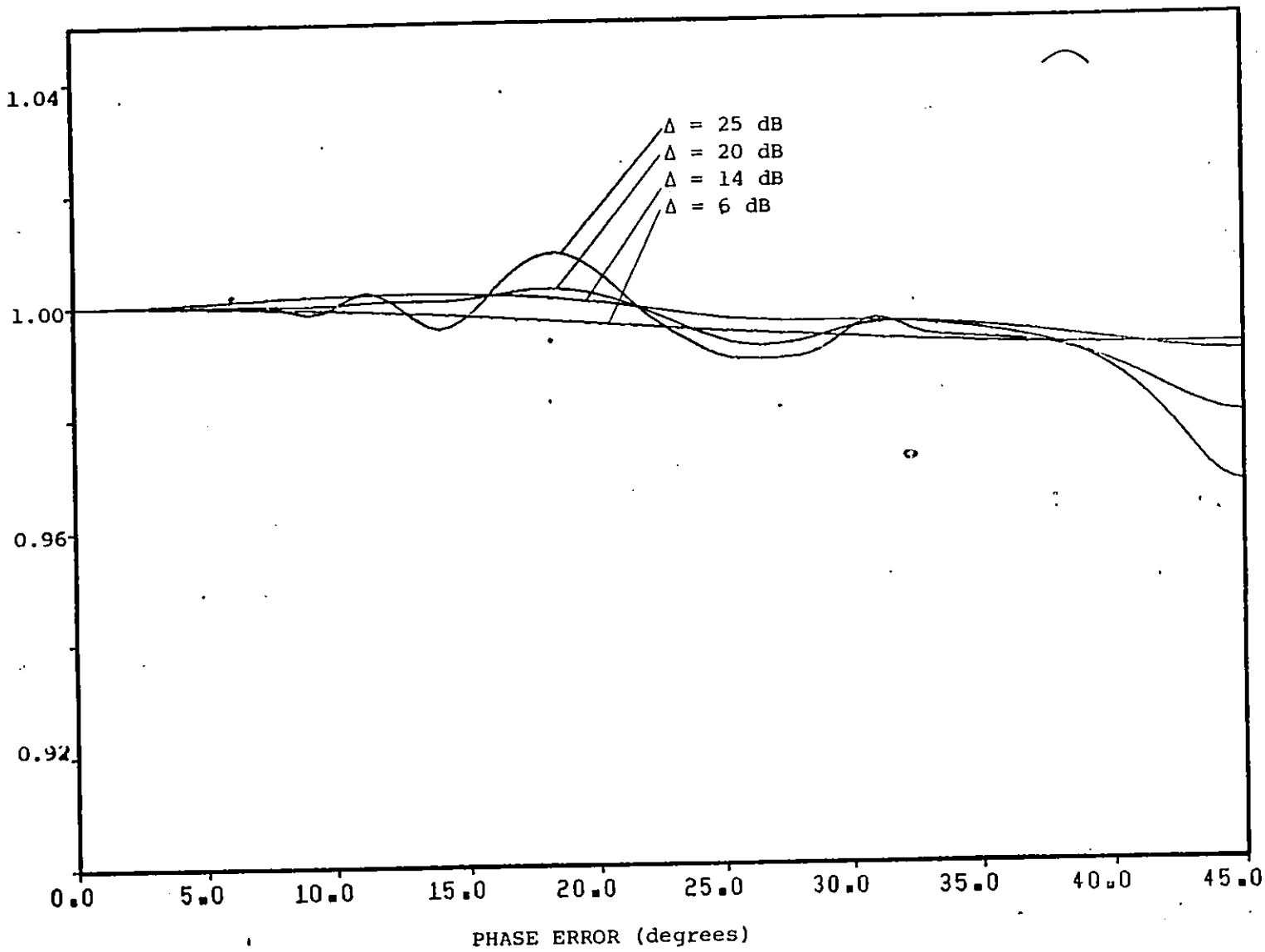


FIGURE 3.4 NORMALIZED NOISE FUNCTION $h(\phi)$ (4-LEVEL DATA)

The solution of Equation (3.40) is well known [21] and is given by:

$$p(\phi) = C_1 \exp\left[\int_0^{\phi} \frac{2 A_o(x) - B_o'(x)}{B_o(x)} \, dx \right] \qquad (3.42)$$

where C_1 is a constant.

After substituting the parameters of the loop Equation (3.42) reduces to:

$$p(\phi) = C_1 \exp\left[- \int_0^{\phi} \frac{\alpha g(x) + h'(x)}{h(x)} \, dx \right] \qquad (3.43)$$

where $$\alpha = 4A/(N_o K_1) \qquad (3.44)$$

is the loop SNR and C_1 is a normalization constant.

Using A and K_1 as defined in Equation (3.15), we have:

$$\alpha = [G'(0)]^2 \Delta^2 \delta_s / H(0) \qquad (3.45)$$

where $$\delta_s = -2/(W_L T) \qquad (3.46)$$

and $$W_L = A K_1/2 \qquad (3.47)$$

is the two sided loop bandwidth.

All the terms in Equation (3.43) have been defined or evaluated except $H'(\phi)$, which will be evaluated next.

$$\begin{aligned} H'(\phi) &= \frac{d}{d\phi} H(\phi) \\ &= (2/K^5) \sum_{ij\ell mn} \{4 \sum_{k'} k' \frac{d}{d\phi} Q[(X+k')/2\sigma]\} \\ &= \frac{8}{K^5} \sum_{ij\ell mn} \{ \sum_{k'} k' [-\frac{1}{\sqrt{(2\pi)}}] \exp[-(X+k')^2/2\sigma^2] \\ &\quad \frac{d}{d\phi} [(X+k')/\sigma]\} \end{aligned} \qquad (3.48)$$

For K = 4,

$$H'(\phi) = \left[-\frac{16}{K^5 \sigma \sqrt{(2\pi)}}\right] \sum_{ij\ell mn} [(m+n)\cos\phi + (i+j)\sin\phi]$$

$$[\exp(-u^2/2) - \exp(-v^2/2)] \qquad (3.49)$$

where $u = (X-2)/\sigma$

$v = (X+2)/\sigma$

With these, the steady state PDF $P(\phi)$ can then be evaluated numerically. The results are plotted in Figures 3.5 and 3.6, as a function of the loop phase error and for different values of Δ and δ_s.

For $\delta_s = 5$, the phase error at an average SNR of 24 dB (Δ = 14 dB) is within ±5 degrees with very high probability. When δ_s the ratio of data rate to loop bandwidth is doubled to 10, the maximum phase error reduces to approximately ±3 degrees. If the data rate is held constant, then a loop with a larger value of δ_s is equivalent to a loop with narrower loop bandwidth. When a narrow loop bandwidth is used, extra noise will be filtered out and smaller phase jitter can be achieved.

In practice, a very high value of δ_s will be used, to ensure a small RMS phase jitter. Typically a loop bandwidth of 60 kHz is used for a 90 Mb/s system [22], a compromise between phase jitter and fast acquisition of phase lock.

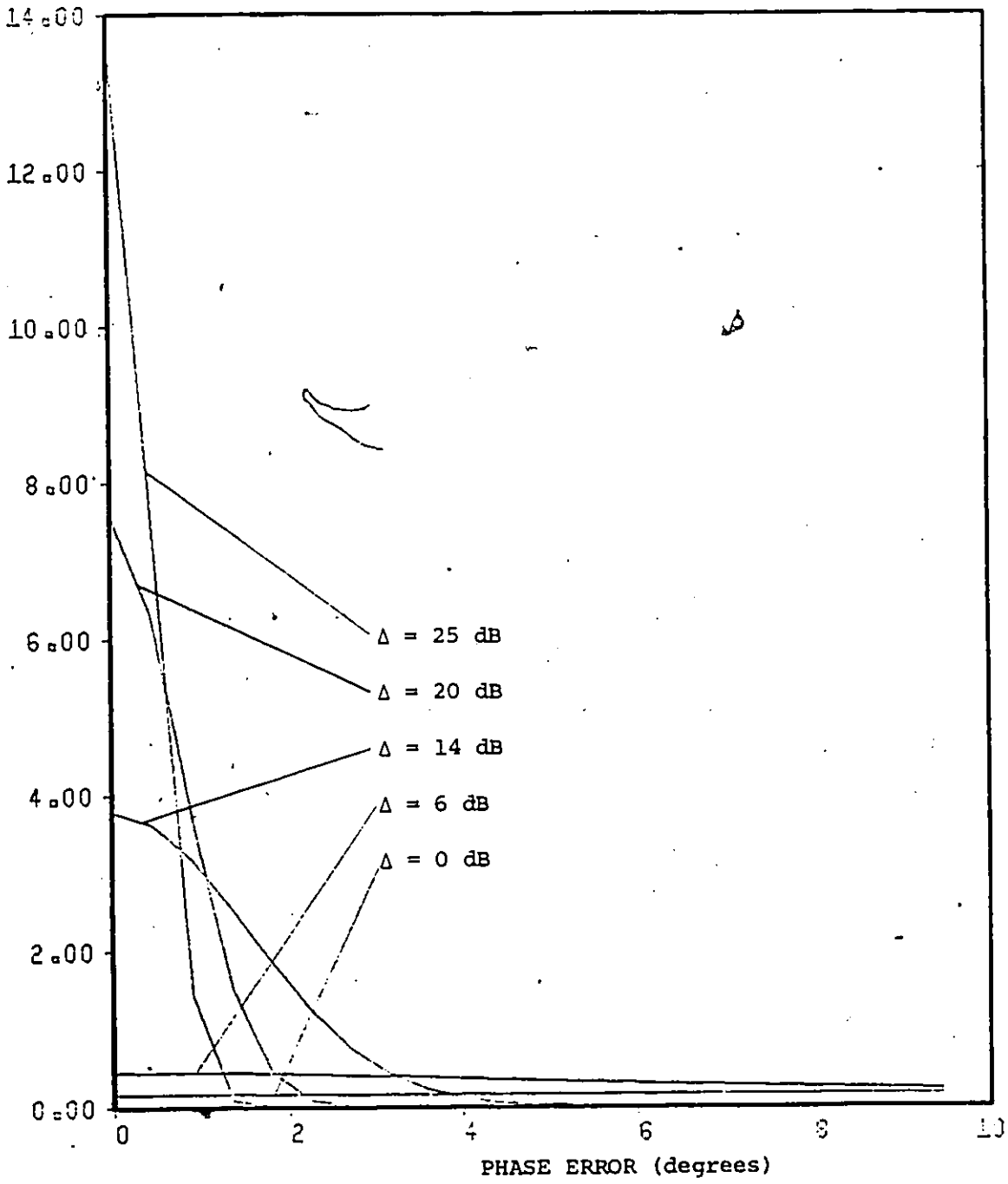


FIGURE 3.5 . STEADY STATE PDF of LOOP PHASE ERROR (δ_S=5)

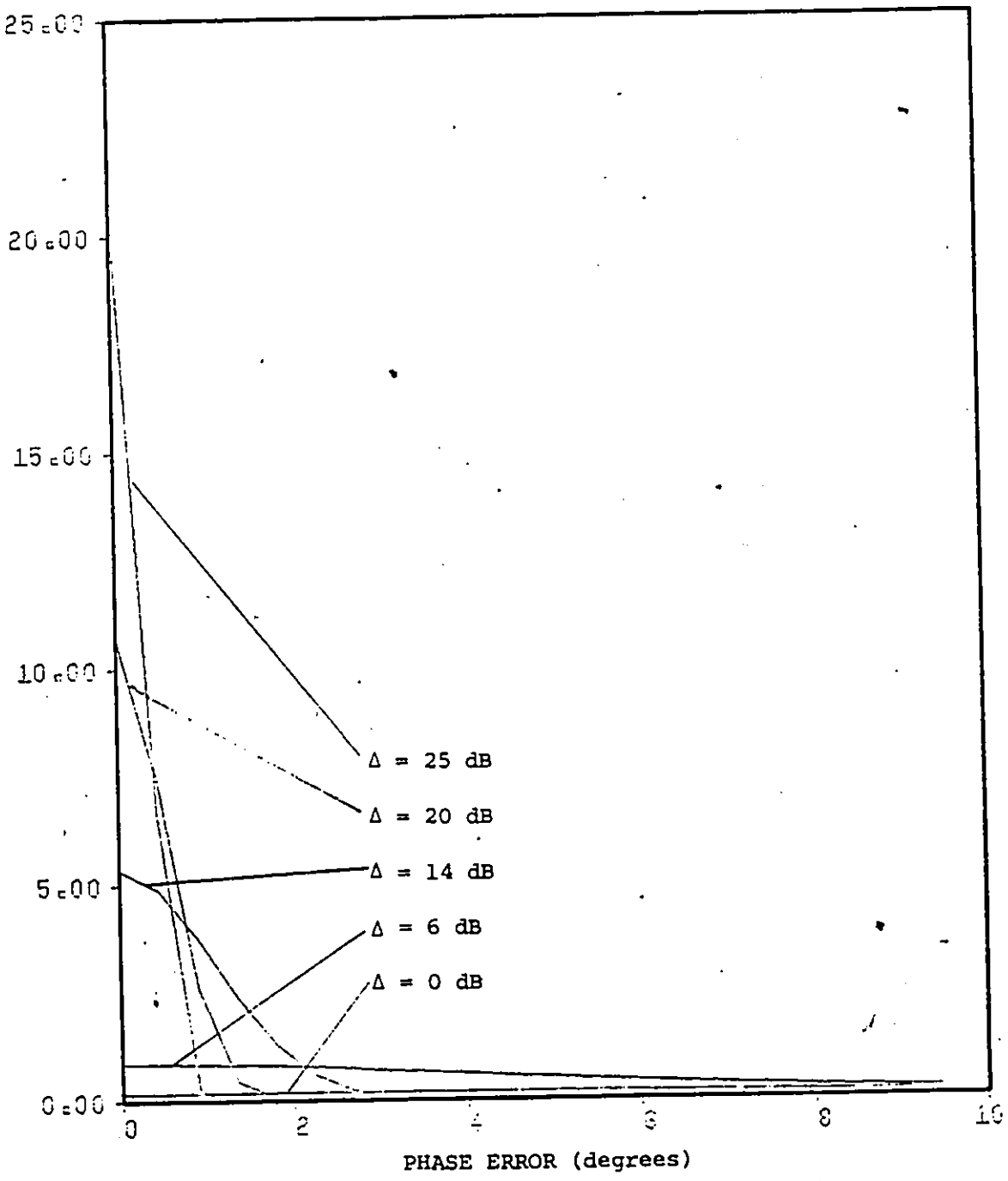


FIGURE 3.6 STEADY STATE PDF of LOOP PHASE ERROR (δ_s=10)

3.5 SIMULATION OF CARRIER RECOVERY LOOP

In order to further verify the analysis, a baseband equivalent model of the loop has been simulated using a digital computer. The baseband equivalent model is nothing more than the mechanization of the stochastic integro-differential equation of operation (Equation 3.9). A block diagram showing the simulation model is shown in Figure 3.7. By making the approximation,

$$\dot{\phi} \approx \{\phi[t+\Delta t] - \phi[t]\}/\Delta t \tag{3.50}$$

Equation (3.9) can be written as the recursive difference equation

$$\phi[t+\Delta t] = \phi[t] - \Delta t\, K_o\, e(t) \tag{3.51}$$

Since the sampling circuits in Figure 3.1 are operated at the baud rate, new values will appear at their output every T seconds. In other words, the error signal or voltage will stay constant for T seconds and will be updated every T seconds. This allows us to rewrite Equation (3.41) replacing Δt by T to get:

$$\phi\,[(n+1)T] = \phi[nT] - K_o\, T\, e[nT] \tag{3.52}$$

where e[nT] is as given in Equation (3.7).

In order to compare the simulation results with the theoretical results obtained in the previous sections, the

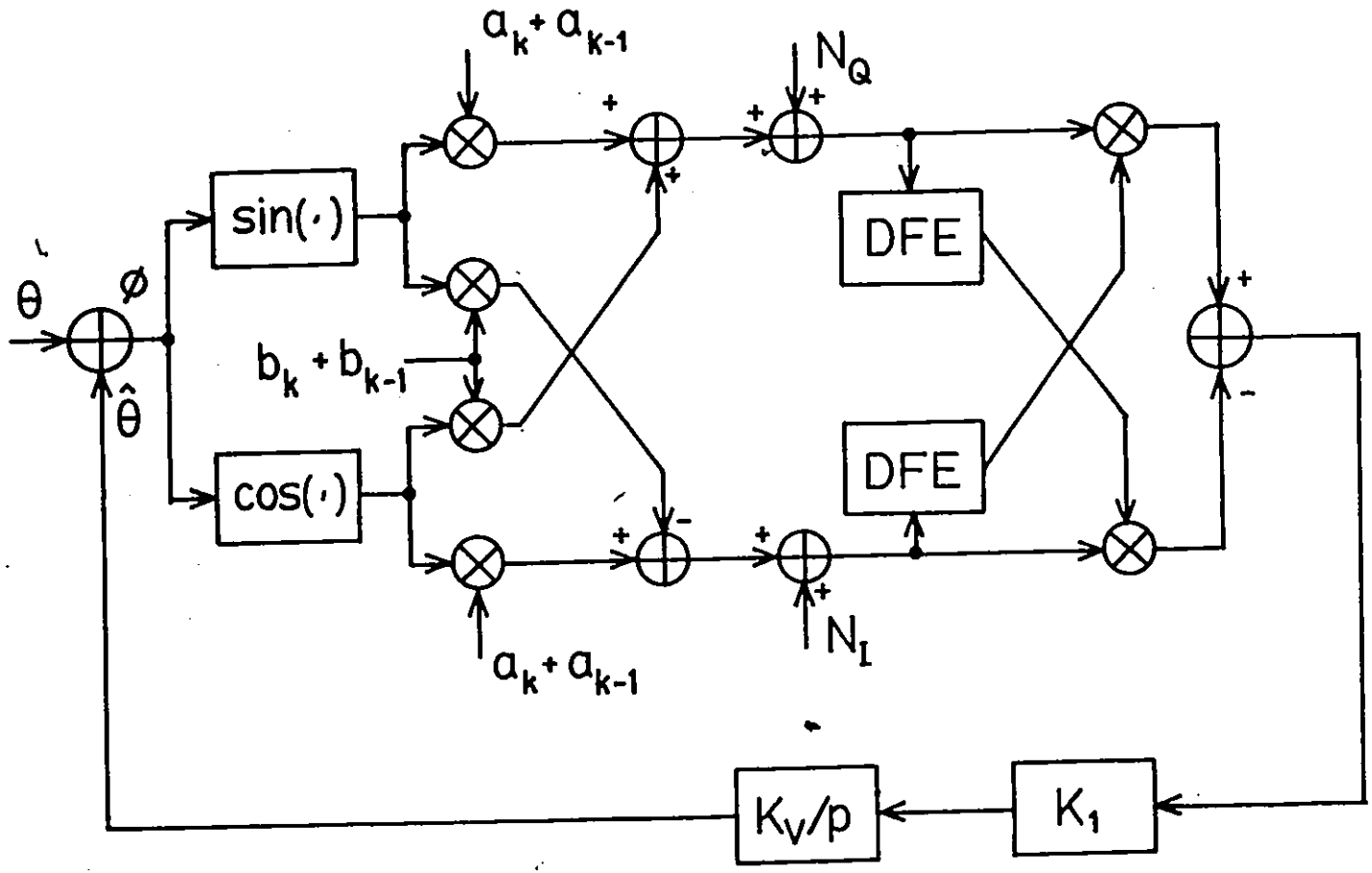


FIGURE 3.7 SIMULATION MODEL FOR CARRIER RECOVERY LOOP

relationship between K_oT in Equation (3.42) and δ_s has to be found.

Using Equations (3.45 - 3.46) and (3.15), the following is obtained:

$$\delta_s = 4/[g'(0)\ K_oT]$$
$$\text{or } K_oT = 4/[g'(0)\ \delta_s] \qquad (3.53)$$

The simulation is nothing more than the mechanization on a computer, of the difference Equation (3.52). The phase error is updated every baud according to this equation. In the simulation, the steady state performance of the loop is acquired. Histograms of the phase error process are then compiled from the simulation results. Figure 3.8 shows a histogram of the loop phase error for $\Delta = 10$ and $\delta_s = 10$. The theoretical curve is also included for comparison. The simulated results have wider distribution. This is probably due to the sample and hold nature of the loop operation, which is not considered in the theoretical analysis. Since sample and hold circuits tend to lower the loop gain [29], a wider distribution should be expected.

3.6 DISCUSSION

A decision directed carrier recovery loop of the 4 bits/s/Hz system has been analyzed. It has shown good

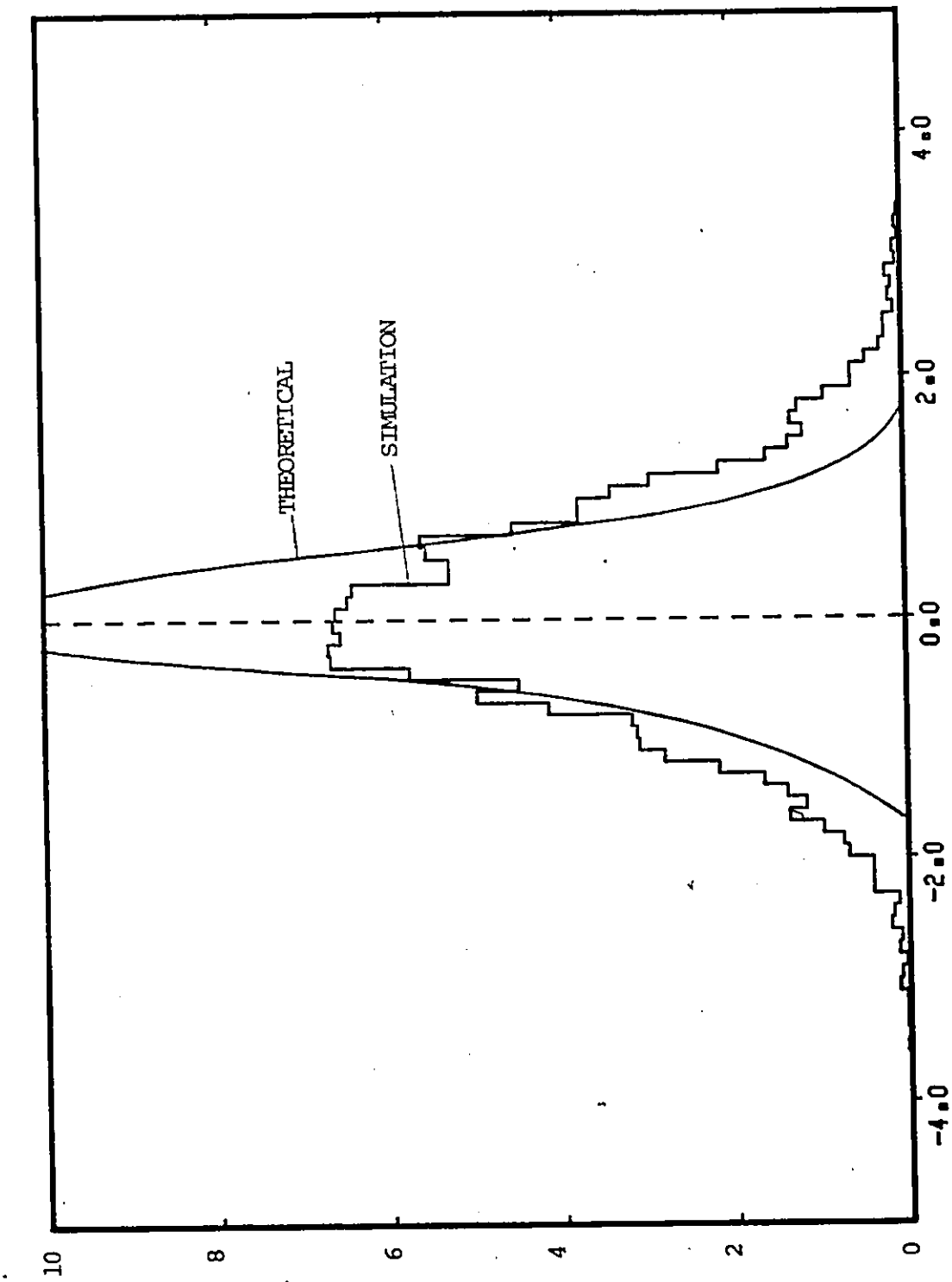


FIGURE 3.8 HISTOGRAM SHOWING LOOP PHASE ERROR PDF OF SIMULATED LOOP

performance at reasonable values of SNR and δ_s. The simulation also confirms the theoretical derivation of the phase error PDF.

It has been shown that the phase detector characteristic has a wider linear range around the lock point for binary data. Hence, it might be worthwhile to start up the system with binary data, because the average startup time will be shorter. Furthermore, this loop structure will work for both binary and 4-level data without modification.

This analysis assumes that the correct sampling times are known precisely. In practice, there will be some timing jitter, this will degrade the performance slightly. However, if the RMS timing jitter is kept small by use of a good symbol timing circuit, the degradation will be insignificant. Such a timing circuit is investigated in Chapter 4.

CHAPTER 4

TIMING RECOVERY LOOP ANALYSIS

The problem of carrier synchronization has been solved in the previous chapter, assuming that information about the correct sampling times are available. The purpose of this chapter is to develop a way to estimate the sampling instants. This problem is usually known as the symbol synchronization or clock recovery problem.

Symbol synchronization information can be obtained in many different ways. Symbol synchronization signals may be provided by a separate channel, for example, the synchronization information may be transmitted as a pilot signal at a spectral null of the signal spectrum, or at the edge of the signal spectrum. However, if the phase distortion in the synchronization and data channels is different, which is usually true, excessive timing degradation may occur so as to make this method inapplicable. Also from the power efficiency point of view, this method is highly inefficient. Zero-crossings of the signal or the derivative of the signal will also provide synchronization information, but this technique can only be applied when the signal to noise ratio is high. Another widely used method is to generate a

spectral line at a multiple of the clock frequency by a nonlinear process [23]. However with the ISI in Partial Response systems, the generated spectral line is lost in an almost continuous spectrum from which it can hardly be recovered [24], unless an extremely nonlinear process such as the 'slicer' method [16] is used. The 'slicer' method is highly dependent upon the automatic gain control circuits or level setting circuits, which is also an undesirable feature. Most symbol synchronizers extract the synchronization information from the data bearing signal directly. This type of synchronizer is usually known as a data derived type of synchronizer, and is more suitable than the other methods for high bandwidth efficient transmission systems. The principle advantage of extracting synchronization information from the data bearing signal is that no additional transmitted power or bandwidth is required [25].

The synchronizer presented here is simple and economical to implement. It also has good acquisition time. Simulations have shown that typically about fifty symbols or adjustments are required to lock in. This is usually less than 1 μs when a data rate of 180 Mb/s is used.

4.1 GENERATION OF ERROR SIGNAL.

A lot of work has been done in the area of symbol synchronization for binary baseband digital communication

systems. Most investigators have taken the Maximum a Posteriori (MAP) estimator approach. Unfortunately, this approach avoids a direct attack on the major problem of acquisition time [25].

Some of the more generally used methods have been described earlier. In a Duobinary system, there is no spectral line at the clock frequency or any multiple of this frequency for random data bit sequences. Therefore, a narrow band filter or a phase locked loop cannot be used to reproduce the timing information required to drive the digital receiver and other synchronization circuits such as parallel to serial converters or demultiplexers. In view of this, some form of nonlinear operation is required to generate a signal which is a function of timing error.

4.1.1 Timing Function with Ideal Decisions

A simple algorithm for extracting timing information from the received signal has been proposed by Mueller and Muller [26]. With this algorithm, convergence is exponential. The basic idea behind this algorithm is to take advantage of the symmetry of the baseband pulse shape. For Duobinary signalling, the pulse shape can be expressed as:

$$h(t) \triangleq \frac{4\,T^2 \cos\,[\pi(t-T/2)/T]}{\pi\,[\,T^2 - 4(t-T/2)^2\,]} \qquad (4.1a)$$

or $$h(t) = \frac{\sin(\pi t/T)}{t/T} + \frac{\sin[\pi(t-T)/T]}{(t-T)/T} \qquad (4.1b)$$

and is plotted in Figure 4.1.

The correct sampling times are 0, T, 2T,...etc., where 1/T is the baud rate.

Now let h_i denote $h(iT+\tau)$. Assuming these h_i samples of the impulse response can be obtained from the received modulated signal, the following timing function can be formed:

$$f(\tau) = (h_1 - h_{-1})/2 \qquad (4.2)$$

where τ is the timing error.

Next, substituting Equation (4.1) into (4.2), Equation (4.3) is obtained:

$$f(\tau) = \frac{2T^2}{\pi}\left\{\frac{\cos[\pi(\tau+T/2)/T]}{T^2 - 4(\tau+T/2)^2} - \frac{\cos[\pi(\tau-3T/2)/T]}{T^2 - 4(\tau-3T/2)^2}\right\} \qquad (4.3)$$

When τ is normalized to the symbol period T, the timing function becomes:

$$f(\tau) = \frac{2}{\pi}\left\{\frac{\cos[\pi(\varepsilon+1/2)]}{1 - 4(\varepsilon+1/2)^2} - \frac{\cos[\pi(\varepsilon-3/2)]}{1 - 4(\varepsilon-3/2)^2}\right\}$$

$$= \frac{2}{\pi}\left\{\frac{\sin \varepsilon\pi}{1 - 4(\varepsilon-3/2)^2} - \frac{\sin \varepsilon\pi}{1 - 4(\varepsilon+1/2)^2}\right\} \qquad (4.4)$$

where $\varepsilon = \tau/T$ is the normalized timing error.

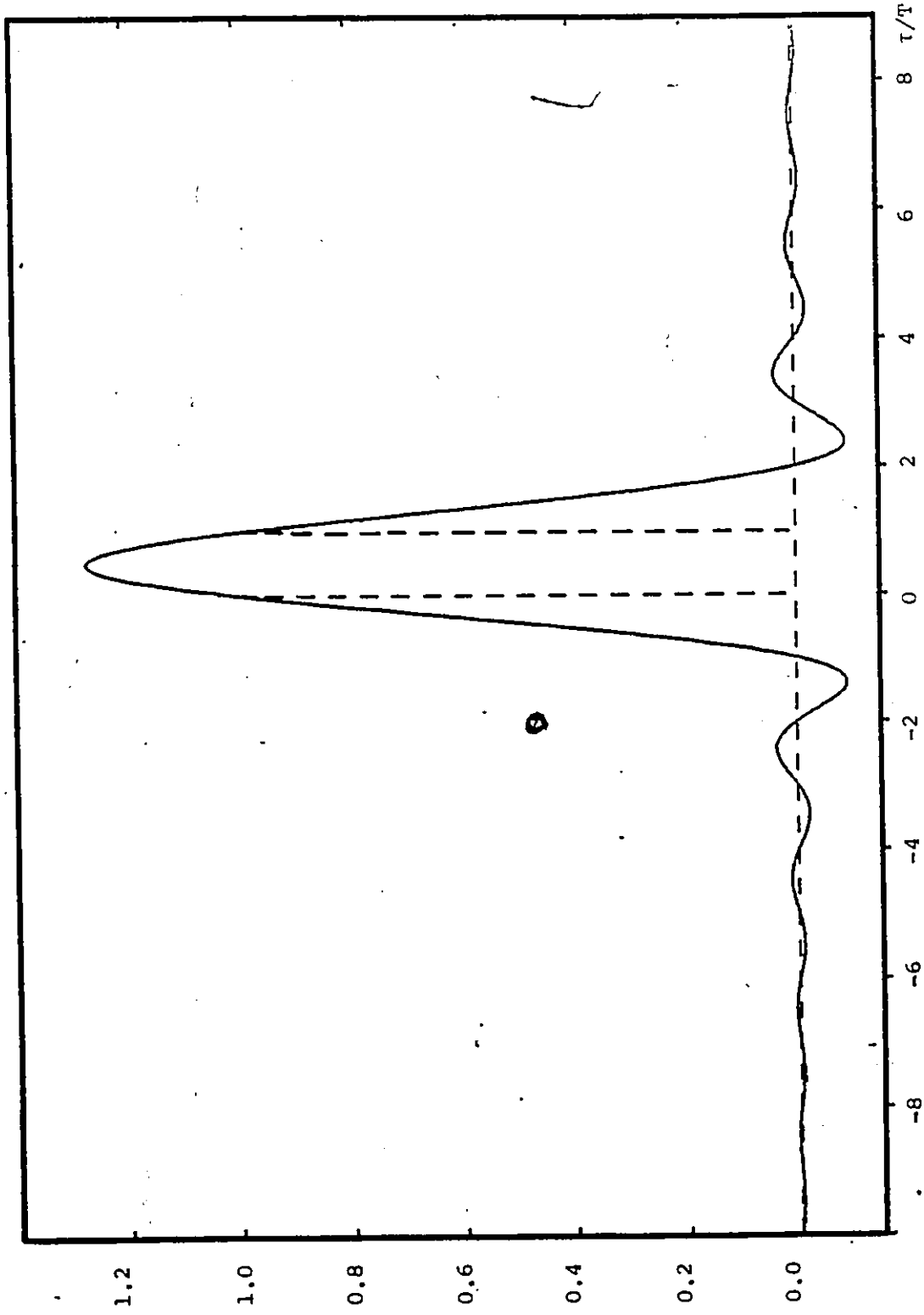


FIGURE 4.1 NORMALIZED IMPULSE RESPONSE

This is plotted in Figure 4.2. From this figure, it is clear that if this were used in a feedback loop, the loop will lock at $\epsilon = 0.5$. That is if $f(\tau)$ is used as an error signal to correct for the timing phase, the loop will lock at $\epsilon = 0.5$. This is the case because the timing function (Equation 4.4) will ensure that one of the sampling instants be at the pulse peak about which symmetry occurs. The sampling times are then T/2 seconds away from the correct instants. This will not pose too much of a problem, because a constant offset can always be added or subtracted to achieve the correct sampling times.

So far, nothing has been written about the method of obtaining h_i samples of the impulse response. Mueller et al [26] have also shown that:

$$E[x_k a_{k-1} - x_{k-1} a_k] = 2\, f(\tau)\, E[a_k^2] \tag{4.5}$$

$$\text{where } x_k(\tau) = \sum_i a_{k-i}\, h(\tau+iT) + n(\tau+kT) \tag{4.6}$$

If we then write,

$$z_k = [x_k a_{k-1} - x_{k-1} a_k] / \{ 2\, E[a_k^2] \} \tag{4.7}$$

Equation (4·5) becomes:

$$f(\tau) = E[\, z_k \,] \tag{4·8}$$

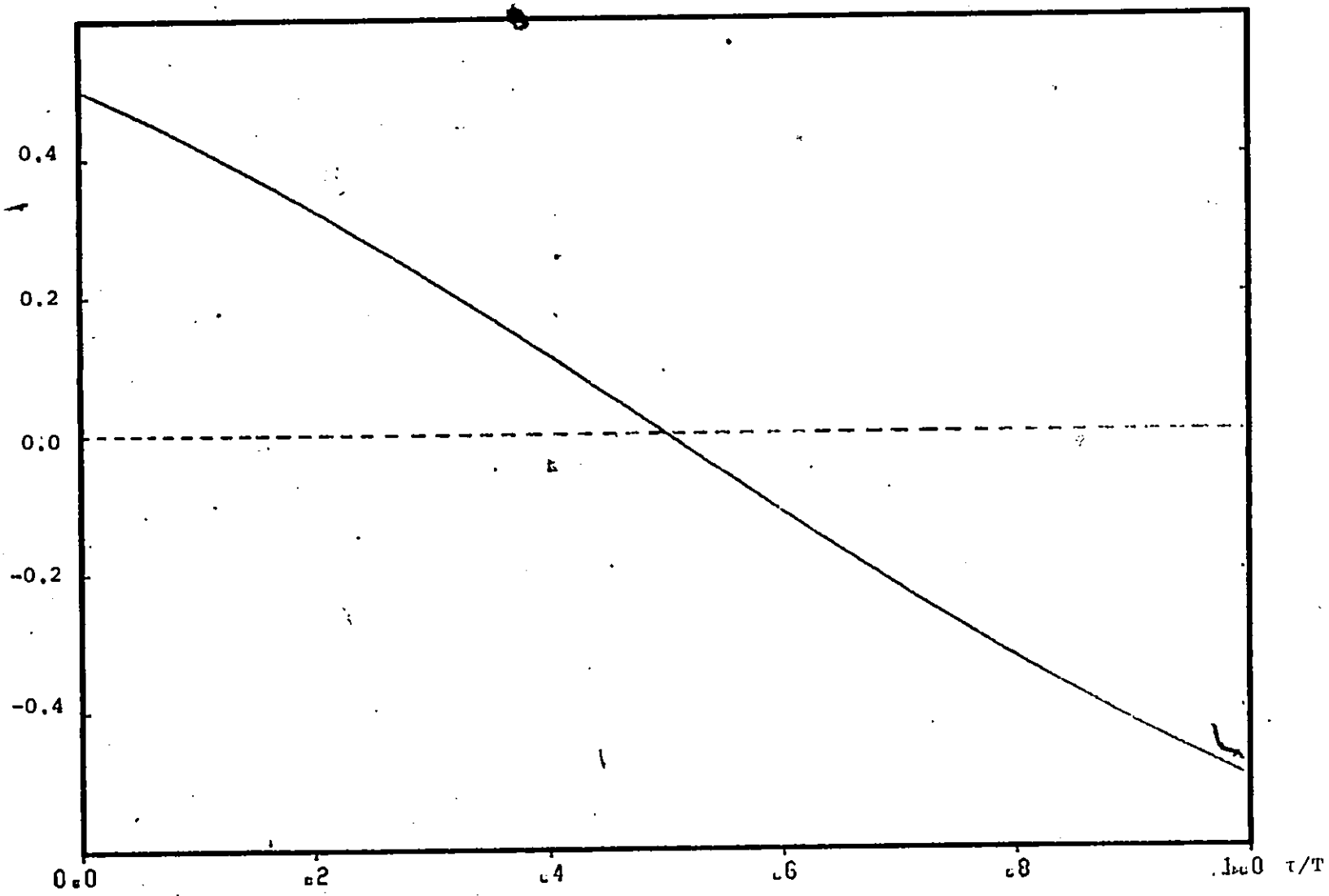


FIGURE 4.2 NORMALIZED TIMING FUNCTION f(τ) AS A FUNCTION OF TIMING PHASE ERROR

However, direct evaluation of $E[z_k]$ would be extremely difficult, if not impossible, due to the correlation between the variables. The ISI problem also makes the evaluation of the x_k's impossible except for the particular case of $\tau = 0$.

The expectation is evaluated as follows:

$$E[z_k] = E[x_k a_{k-1} - x_{k-1} a_k] / 2E[a_k^2] \tag{4.9}$$

$$= \{E[x_k a_{k-1}] - E[x_{k-1} a_k]\} / \{2E[a_k^2]\}$$

For $\tau = 0$, we obtain, for 4-level symbols a_k:

$$E[x_k a_{k-1}] = \Sigma\, x_k a_{k-1} P[a_{k-1}]\, P[a_k]$$

$$= \frac{1}{16} \Sigma\, x_k a_{k-1}$$

$$= 5 \tag{4.10}$$

$$E[x_{k-1} a_k] = E[x_{k-1}]\, E[x_k]$$

$$= 0 \tag{4.11}$$

From Equation (4.9) to (4.11), we have,

$$E[z_k(\tau=0)] = 5/\{2\, E[a_k^2]\}$$

$$= 0.5 \tag{4.12}$$

This result agrees with $f(\tau)$ obtained as in Figure 4.2, using Equation (4.4). No further attempt has been made to evaluate the expectation in Equation (4.5). However, some simulations have been made to further verify the validity of Equation (4.5). In these simulations, the statistics

including the mean and variance of z_k are compiled. A truncated pulse shape (Figure 4.3) is used instead of the ideal pulse shape for practical reasons. With this pulse shape, only four symbols from each side are allowed to interfere with the current symbol when the timing phase is in error. This is a very close approximation to the ideal, infinite time response. From Figure 4.1, it is clear that the contribution due to the rest should be insignificant, because the tails of the pulse shape dies down quite quickly. In fact, h(t) decays as $1/t^2$. A total of 1000 records have been gathered for each value of ε, and then the average is taken as $E[z_k]$. The mean and variance hence compiled is shown in Figure 4.4 and 4.5 respectively. From Figure 4.4, we can see that the simulation results fall quite close to the theoretical or the ideal timing function.

4.1.2 Timing Function with Non-Ideal Decisions

In the last subsection, the derivation of the timing function assumed that all decisions are correct even at maximum timing error. This is definitely not true in practice. For practical circuits, decision errors will inevitably be made, once the timing error is large enough that the eye* is closed. Again, the analysis taking

---

* for an explanation on the 'eye' in communication see [30,p. 92]

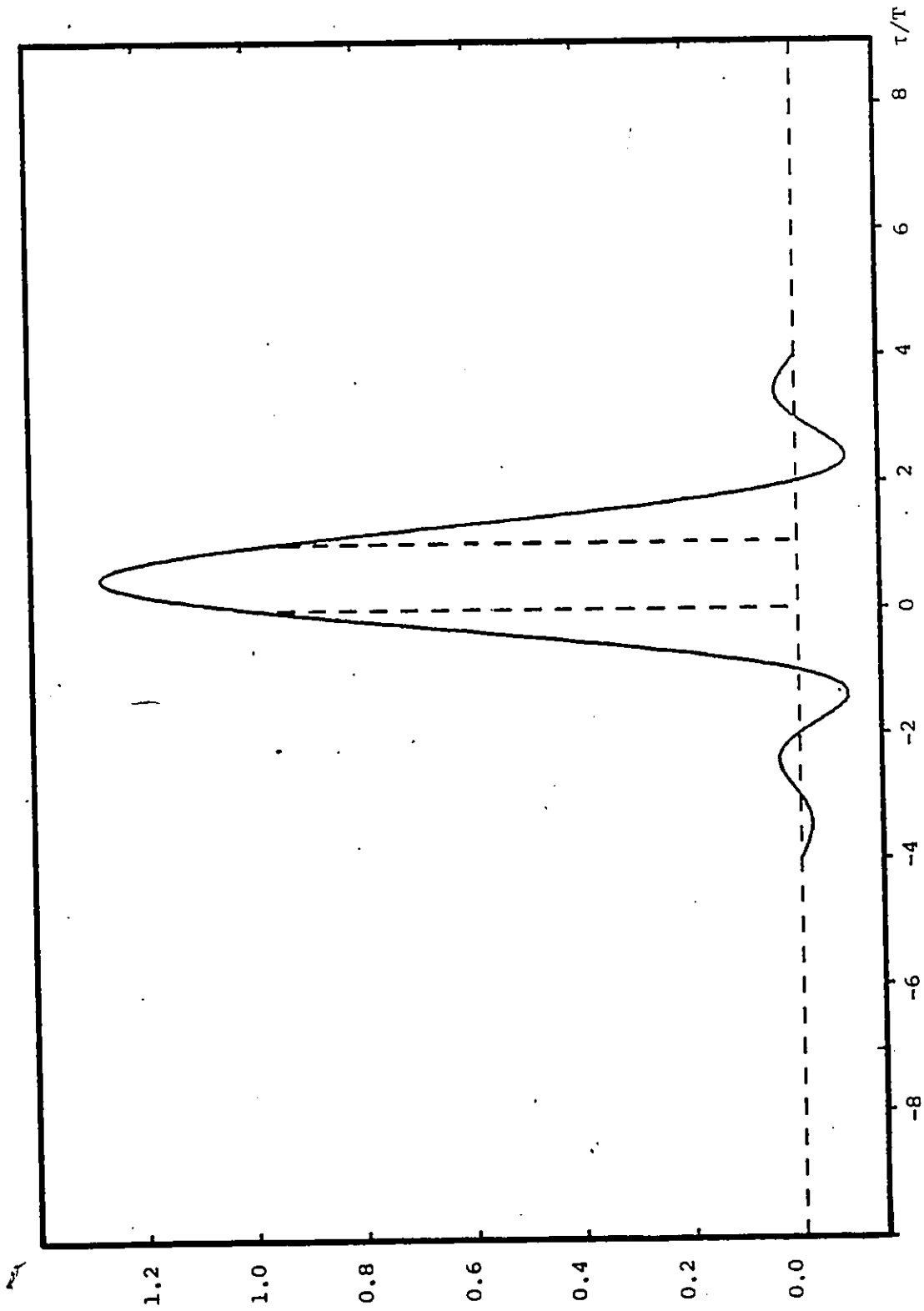


FIGURE 4.3 TRUNCATED IMPULSE RESPONSE

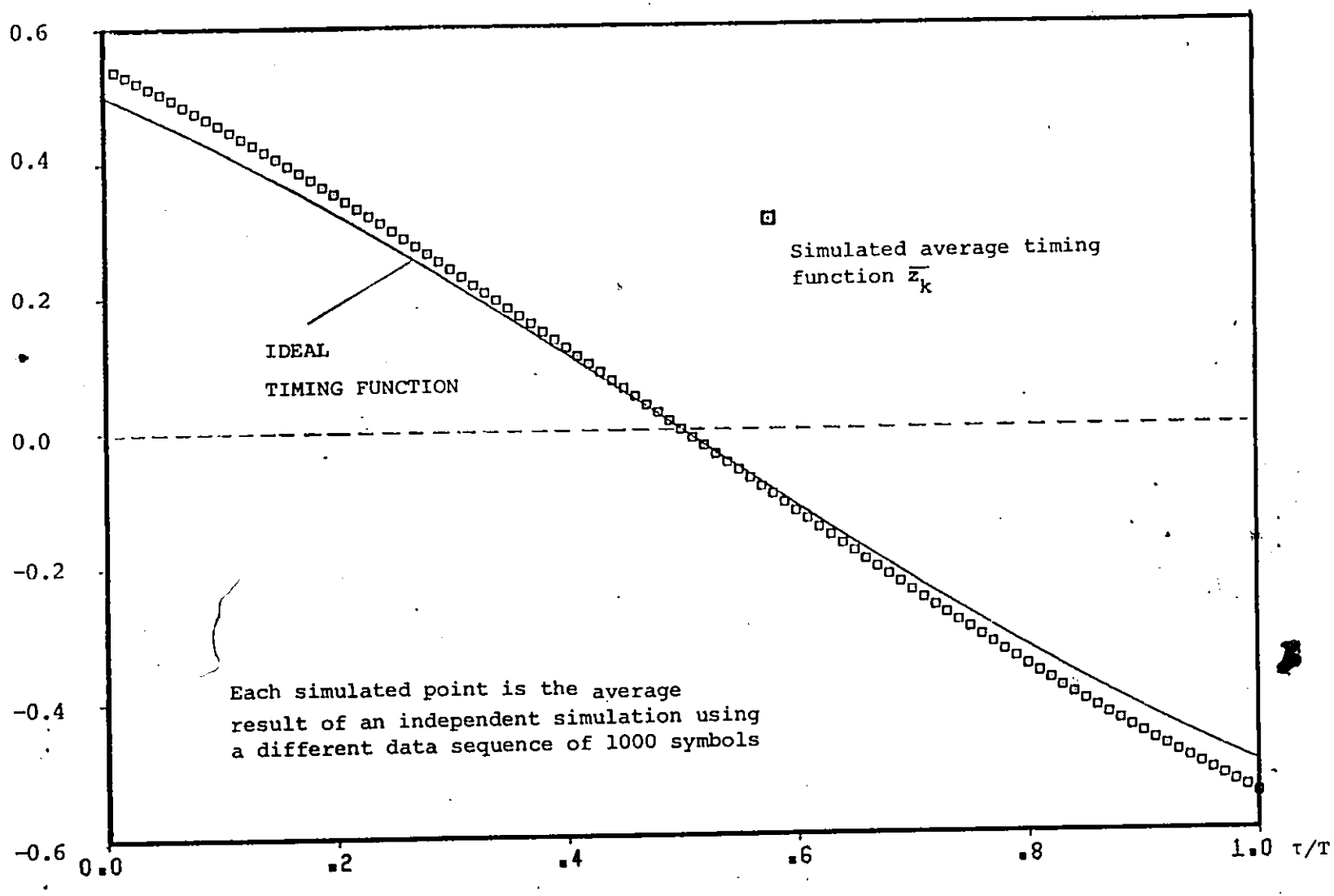


FIGURE 4.4 SIMULATED TIMING FUNCTION WITH IDEAL DECISIONS (4 level data)

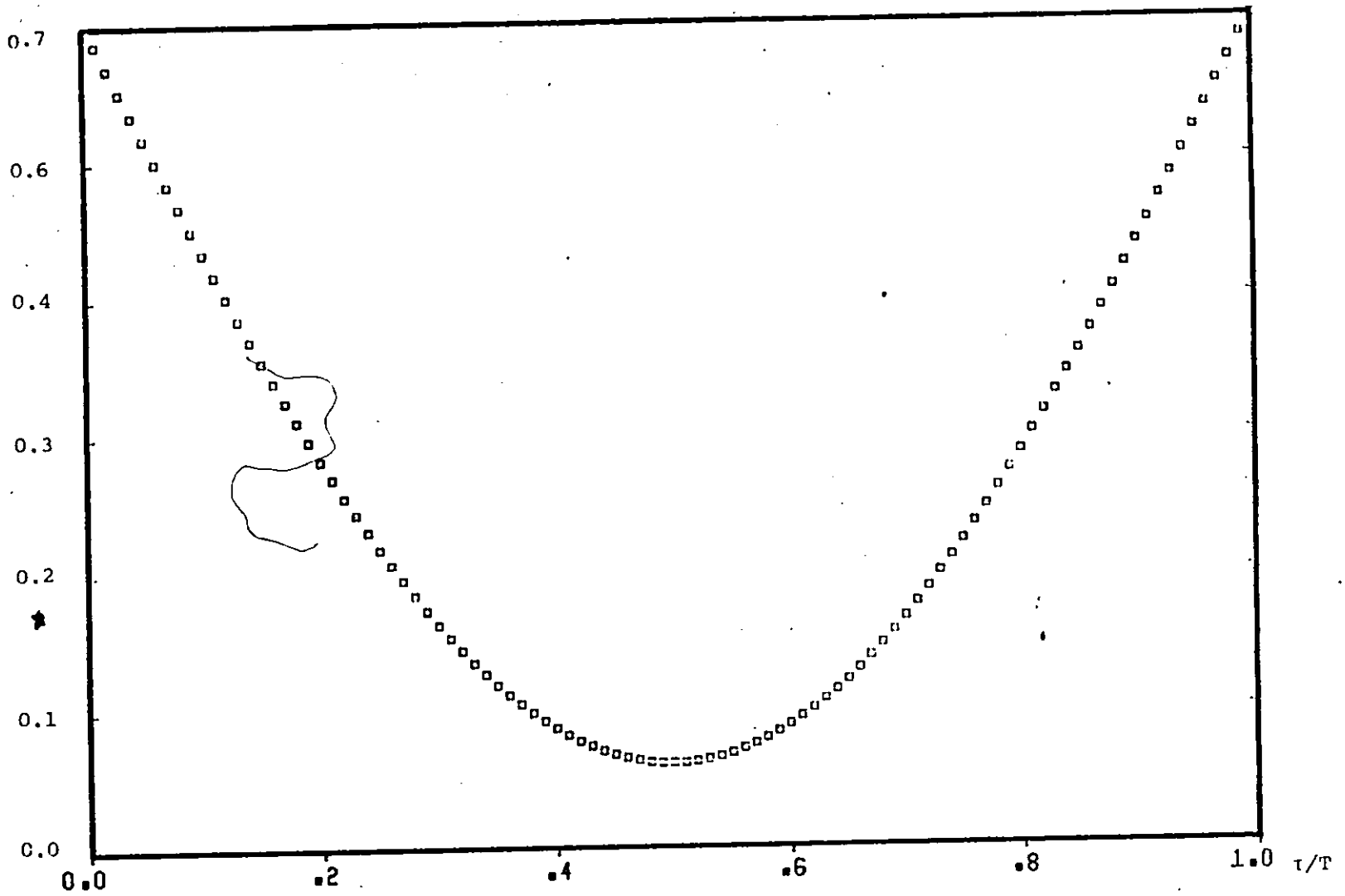


FIGURE 4.5 SAMPLE VARIANCE OF Z_k AS A FUNCTION OF TIMING ERROR τ

decision errors into account is intractable, due to the various reasons mentioned previously. Simulations are again used to solve the problem of evaluation.

In these simulations, the decision feedback decoder is used to estimate the data symbols. In order for the decision feedback decoder to operate correctly, the received signal must be sampled at the correct instants, that is t = 0, T, 2T... etc. In these simulations, the sampling time is shifted by T/2 seconds to achieve the correct time. This will be clearer, when the complete timing recovery circuit is considered in the next section (see Figure 4.8).

By allowing errors to occur, in the decision process, the results for the noise free case are shown in Figures 4.6 and 4.7. The simulated points lie quite close to the ideal characteristics around $\epsilon = 0.5$, where the sampling times are correct. That is, the signal is sampled at the region where the eye is open. Therefore, all decisions are correct, and that is exactly the assumption made in the previous simulations. When the timing phase is off by approximately ±0.1 T, decision errors start to appear, and the timing error signal deviates from the ideal. It is interesting to note that the eye width as defined by Kabal and Pasupathy [17] is 0.18 T for 4-level data. This simulation shows that the eye begins to close at approximately ±0.1 T - that is, at an eye width of approximately 0.2 T. Hence, the results

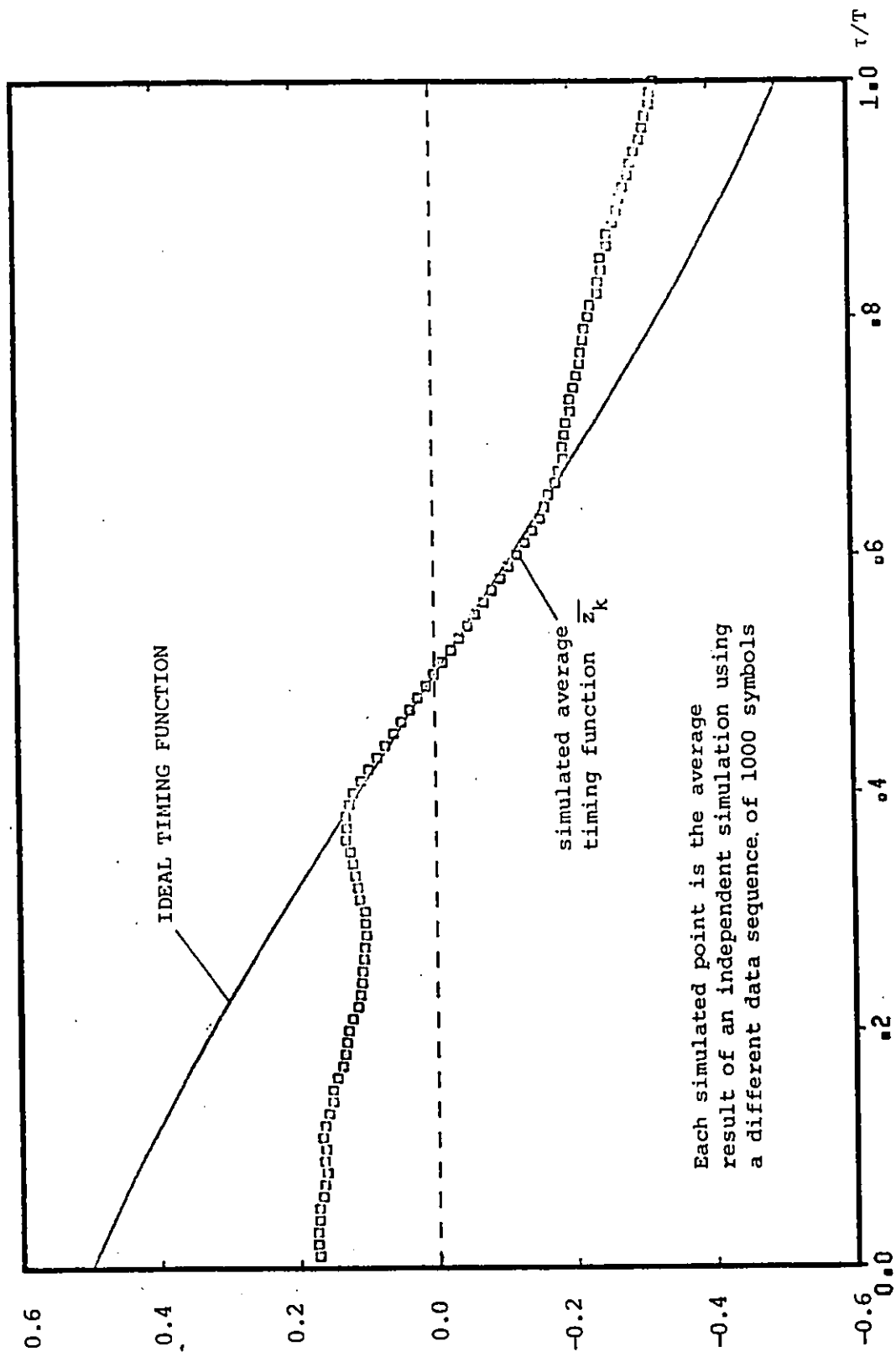


FIGURE 4.6 Simulated Timing Function with Decision Feedback
(4 level data)

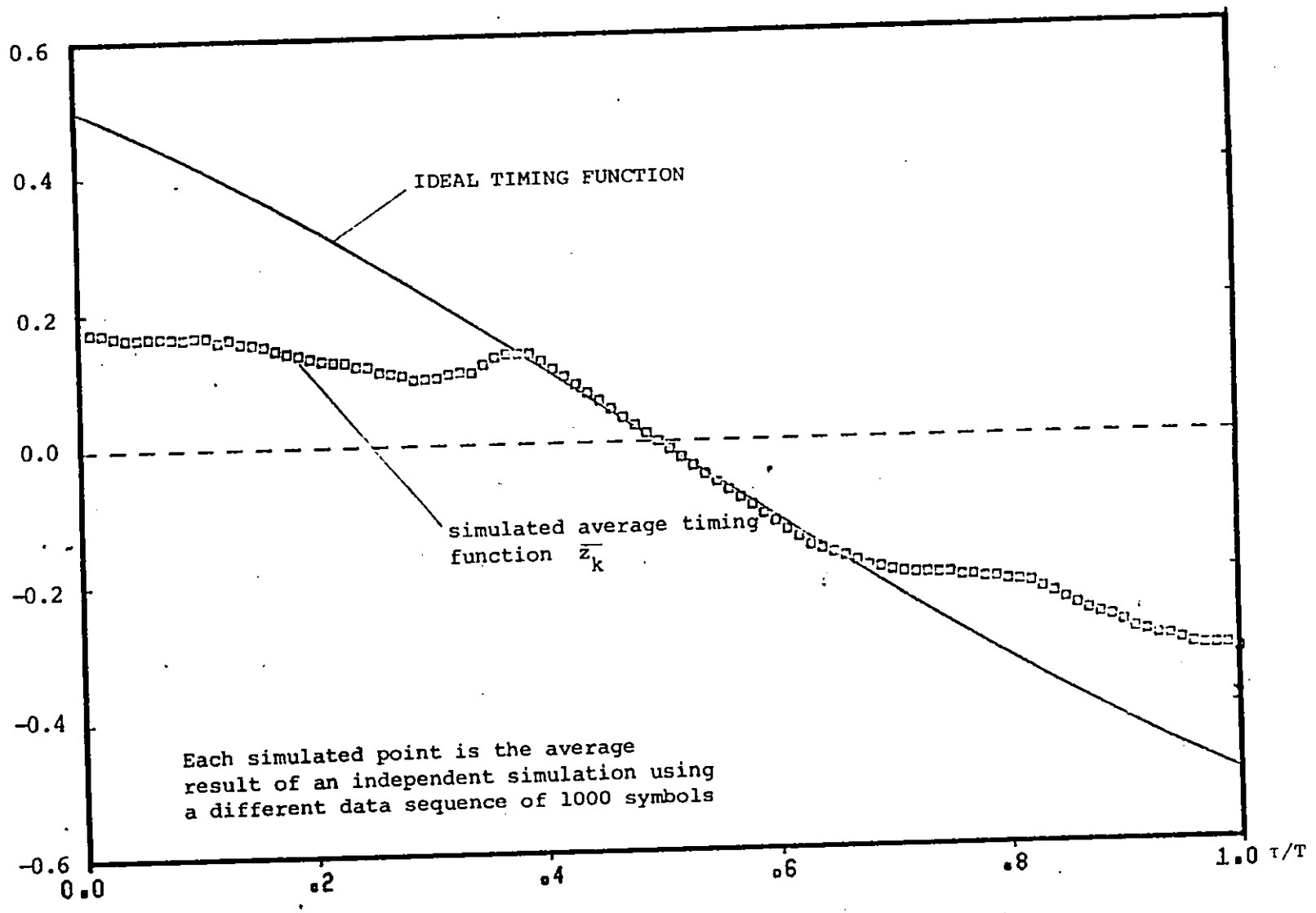


FIGURE 4.7 SIMULATED TIMING FUNCTION WITH DECISION FEEDBACK (4 level data) using different data sequences from figure 4.6

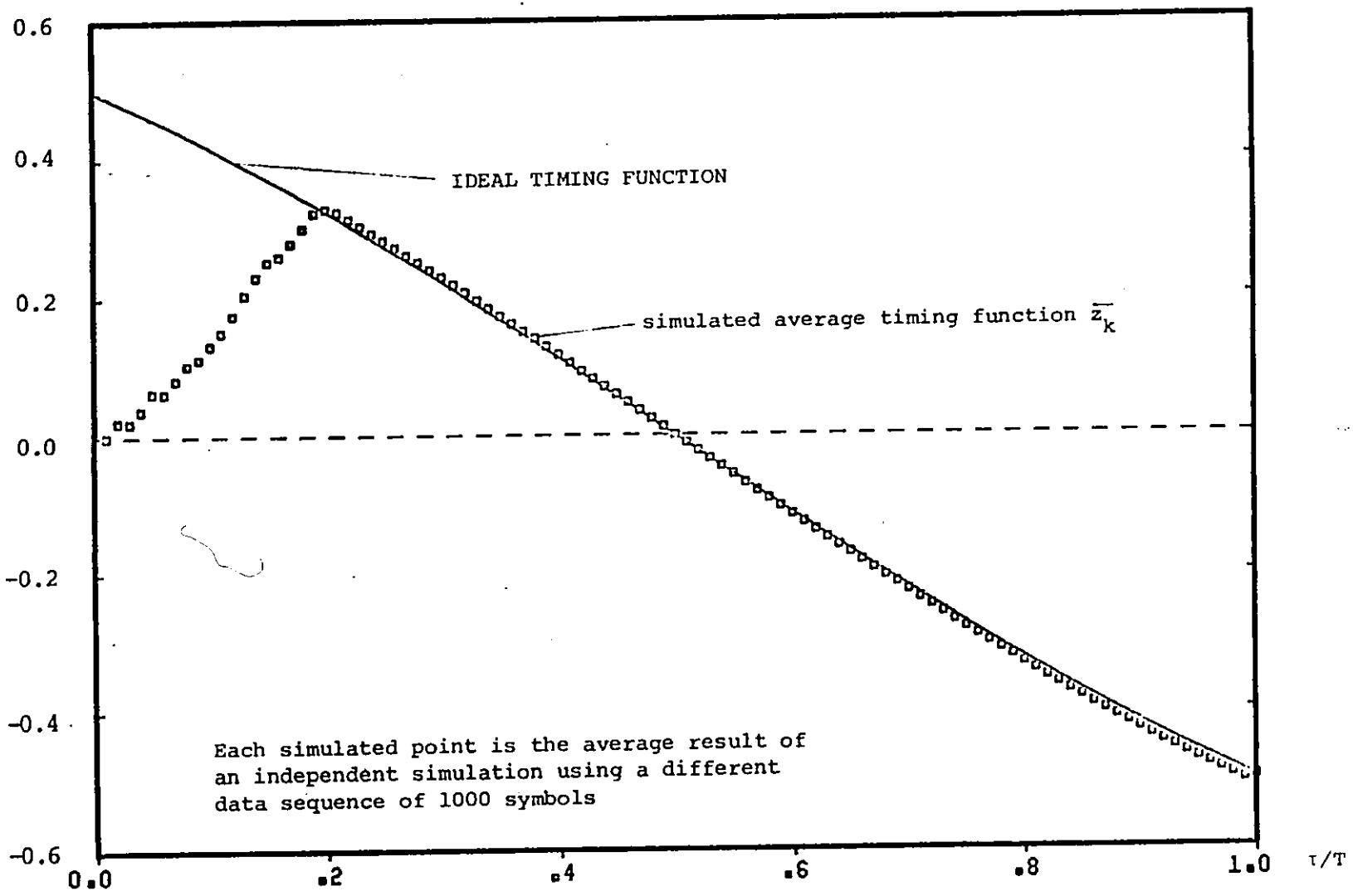


FIGURE 4.8 SIMULATED TIMING FUNCTION (binary data)

on the eye width agree with each other quite closely.

4.1.3 Timing Function for Binary Data

When ideal decisions are fed back to form the timing error signal, simulations have shown that the ideal characteristics can be expected. Simulation results similar to those for the 4-level data case are obtained. When non-ideal decisions are fed back to generate the error signal, deviation from the ideal characteristics (Equation 4.4) again occurs. The results of two of these simulations are shown in Figures 4.8 and 4.9. Unlike the results for 4-level data, these results show great asymmetry about the pulse peak. To explain this asymmetric behaviour, the impulse response of the Duobinary filter has to be considered. There is no symmetry about the sampling instant ($t = 0$), at which decisions are made. The ISI introduced due to timing error is not symmetrical. Therefore, this asymmetric behaviour of the timing function can be expected. This behaviour is also present when the data consists of 4-level symbols. However, its effect is not as pronounced, because an error in the 4-level case need not change the polarity of that decision, whereas the polarity of the decision will be changed from +1 to -1 or vice versa, when the data is binary, and this will cause large deviation in $E[z_k]$ for large timing errors.

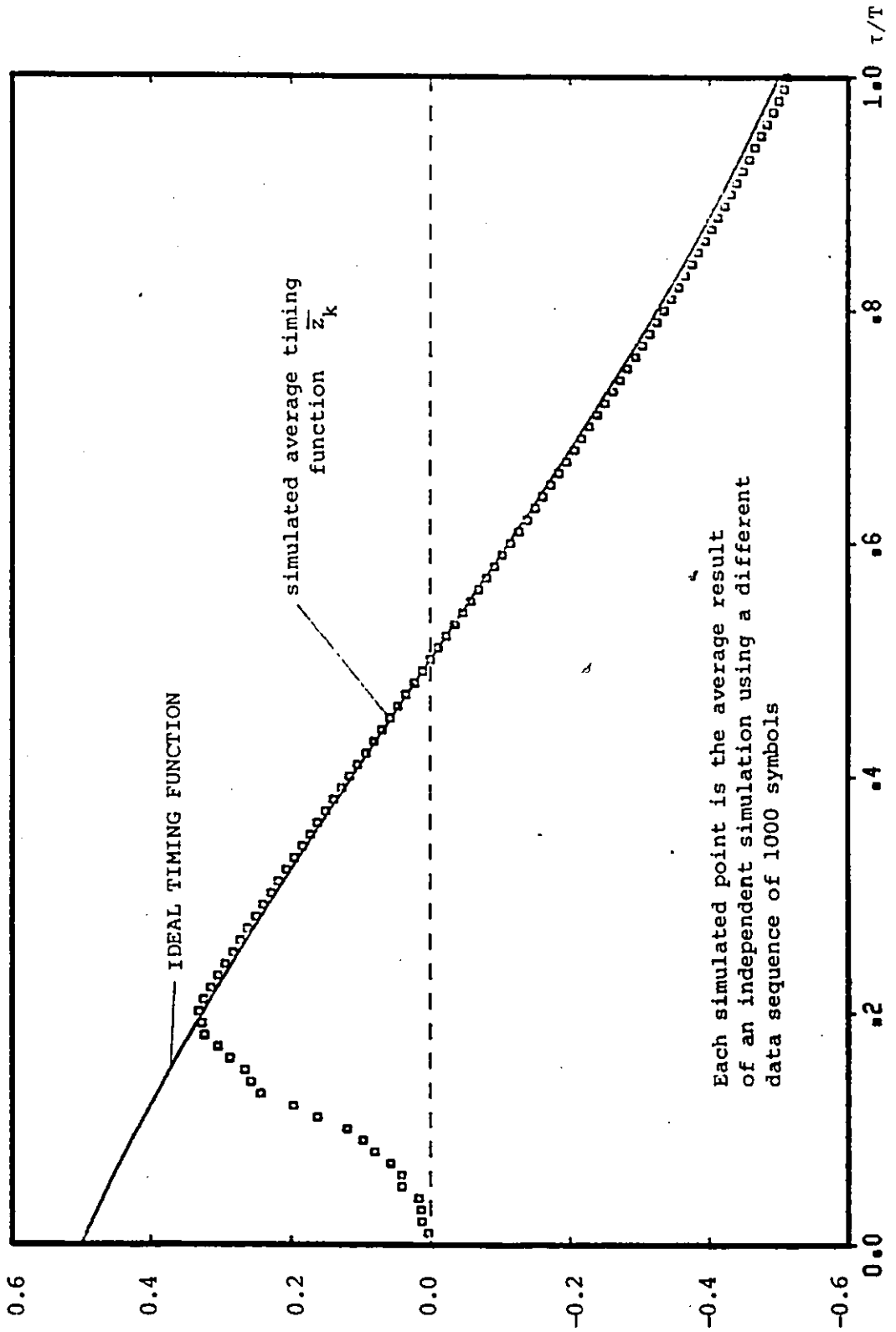


FIGURE 4.9 SIMULATED TIMING FUNCTION (binary data)
using different data sequences from figure 4.8

In Figures 4.8 and 4.9, the error signal at $\epsilon = 0$ is extremely small and this poses a problem during acquisition. If the timing loop is started at $\epsilon = 0$, the loop might stay at $\epsilon = 0$, without locking in, or it might take a long time to lock in. But the probability that the loop will start exactly at $\epsilon = 0$ is theoretically equal to zero. This problem will be explored in a later section.

4.2 LOOP STRUCTURE ANALYSIS

The timing function (4.6) can be used in a loop structure as in Figure 4.10. In general the loop bandwidth will be much smaller than that of the data, therefore, the loop tends to average out the error signal to give the timing function. Using this algorithm, the timing error signal z_k can be generated by two multipliers and an adder. The simplicity of this circuit makes it extremely attractive. The decision feedback decoder already exists in the carrier recovery loop and therefore, a_k and a_{k-1} are readily available. One extra sample and hold circuit and a delay element is required to generate x_k and x_{k-1}. The two sample and hold circuits are operated with a T/2 offset between them. This arrangement is used to take into account the effect of the timing function. In practice, a low pass or loop filter might be used to average out z_k before applying it to control the voltage controlled clock (VCC) or

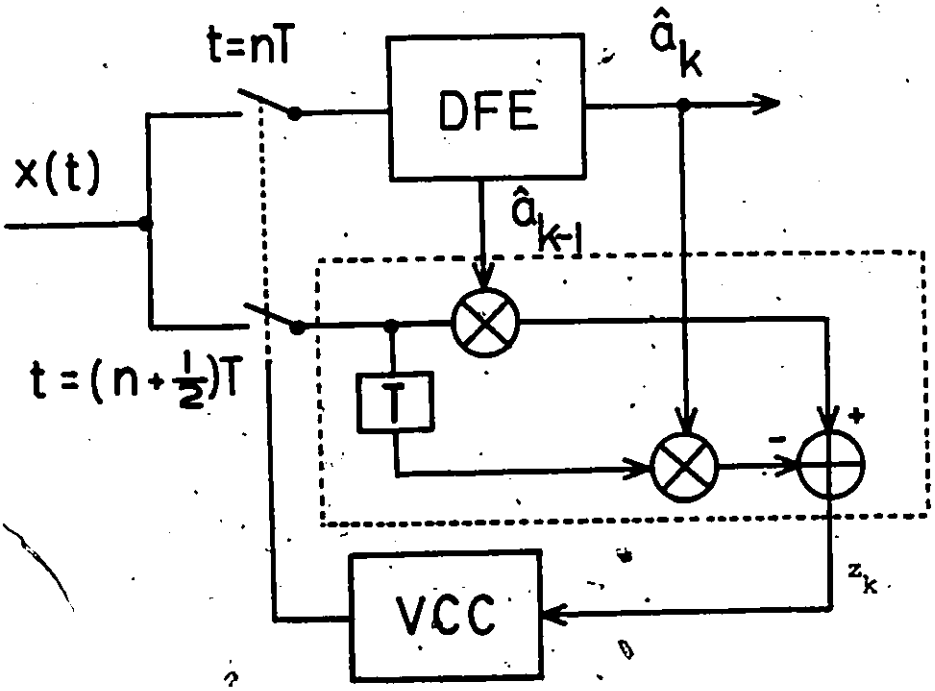


FIGURE 4.10 TIMING RECOVERY LOOP STRUCTURE

timing pulse generator. However, in the present investigation, we shall restrict ourselves to the case of a first order loop (no filter).

The circuit in Figure 4.10 can be represented in block diagram form as in Figure 4.11a, and its phase locked loop equivalent is presented in Figure 4.11b, where the value being tracked is the timing phase τ. The equation of operation of the loop can then be written as:

$$\dot{\tau} = K_o\, e(t) \tag{4.13}$$

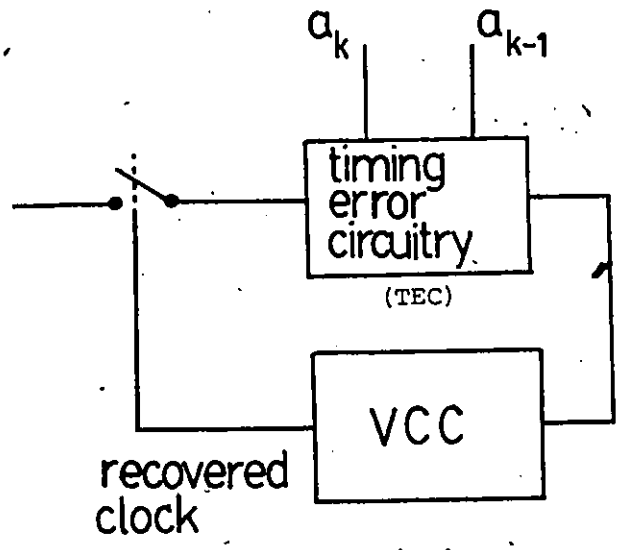
where $e(t) = x_k a_{k-1} - x_{k-1} a_k$

and K_o is the open loop gain.

The factor $1/\{2E[a_k^2]\}$ is included in the loop gain K_o for convenience. The so-called timing function $f(\tau)$ can really be called the phase detector characteristics or average loop S-curve in a stochastic analysis. Simulations have been resorted to further analyze the behaviour of this timing recovery loop.

4.3 SIMULATION ANALYSIS

The circuit in Figure 4.10 is analyzed by means of simulation in the following sections. The previously used truncated pulse shape (Figure 4.3) is again used in these simulations. Here both the steady state performance and the acquisition characteristics are investigated.



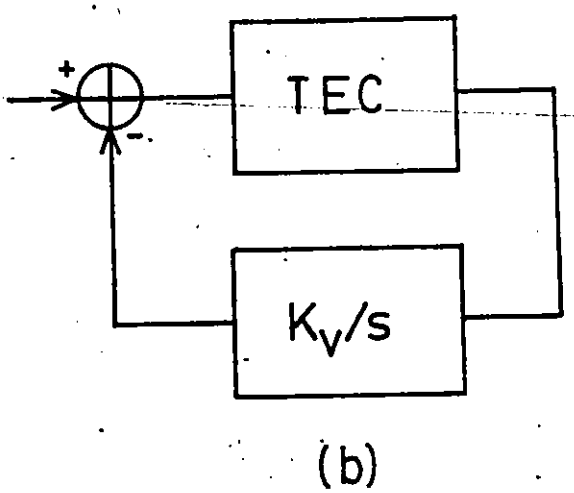


FIGURE 4.11 (a) Block Diagram of Timing Loop
(b) Phase Locked Loop Equivalent

4.3.1 Steady State Analysis

Equation (4.13) shows the operation of the loop. Since the timing phase is updated every baud, the equation of operation can be rewritten as:

$$\frac{\tau_n - \tau_{n-1}}{T} = K_o\, e(t)$$

or
$$\tau_n = \tau_{n-1} + K_o T\, e(t) \qquad (4.14)$$

When expressed in terms of z_k, the sampling phase adjustment equation becomes:

$$\tau_n = \tau_{n-1} + \gamma z_k \qquad (4.15)$$

where $\gamma = K_o T / \{\, 2E[a_k^2] \,\}$

The results obtained from these noise-free simulations are tabulated in Table 4.1, for different values of γ. Each simulation is run using a data sequence of 1000 symbols.

The average sampling point is effectively at t= 0, the desired sampling time. The small bias is probably dependent on the data sequences. If the statistics were taken over a larger sample, say 10000 symbols, the bias would be much smaller. The RMS phase jitter is approximately proportional to the loop gain as Mueller et al have observed [26].

TABLE 4.1 STEADY STATE PERFORMANCE OF TIMING RECOVERY LOOP

γ	AVERAGE LOCK POINT (normalized to T)	RMS JITTER (normalized to T)
0.1	0.002 882	0.028 697 5
0.01	0.000 295	0.002 749 9
	0.000 665	0.002 744 6
	0.000 311	0.002 894 2
0.001	-0.000 055	0.000 286 3
	-0.000 033	0.000 271 9

Next the performance of the loop in additive noise is analyzed. The model simulated is shown in Figure 4.12. Two independent additive white Gaussian noise samples with variance σ^2 are added as shown. The steady state statistics are shown in Table 4.2. The RMS jitter increases as the average SNR decreases. A typical value of the RMS timing jitter for a value of $\gamma = 0.1$ is approximately 3 percent of the symbol period. In practice a much smaller value of loop gain-symbol time product will be used to achieve even better phase jitter performance.

To complete the steady state loop analysis, the steady state timing phase error probability density function (PDF) is compiled for some of the simulations. Histograms showing the PDF of the steady state normalized timing error are presented in Figures 4.13 to 4.15 for different values of Δ. Simulations of the loop for 2-level data have also been done and the PDF's of the steady state timing error for various Δ are shown in Figures 4.16 to 4.17.

4.3.2 Acquisition Performance Analysis

Following the steady state analysis, the acquisition behaviour of the loop is analyzed. At the beginning of each simulation run, the timing phase is set at the maximum possible error, that is $\varepsilon = 1$. The loop mechanism is then allowed to run, and the phase information is acquired after

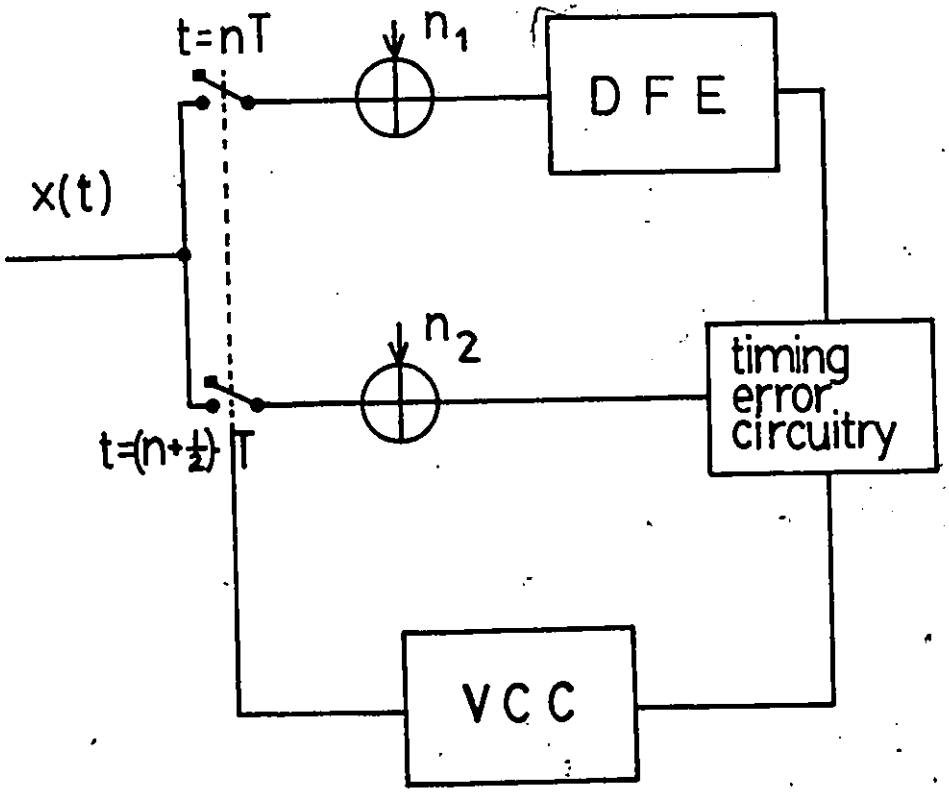


FIGURE 4.12 SIMULATION MODEL OF TIMING RECOVERY LOOP

TABLE 4.2 STEADY STATE STATISTICS OF TIMING RECOVERY LOOP IN THE PRESENCE OF ADDITIVE GAUSSIAN NOISE

Δ	γ	AVERAGE LOCK POINT (normalized to T)	RMS JITTER (normalized to T)
5	0.1	-0.001 008	0.030 596 5
		0.003 069	0.031 776 5
		0.000 160	0.031 660 2
		0.004 038	0.032 951 8
10	0.1	-0.001 888	0.029 807 5
		-0.000 551	0.029 239 9
		0.001 458	0.029 811 0
		0.000 896	0.031 677 6
∞	0.1	-0.001 559	0.028 740 0
		0.000 486	0.028 615 3
		0.000 440	0.028 388 3
		0.000 853	0.030 301 2

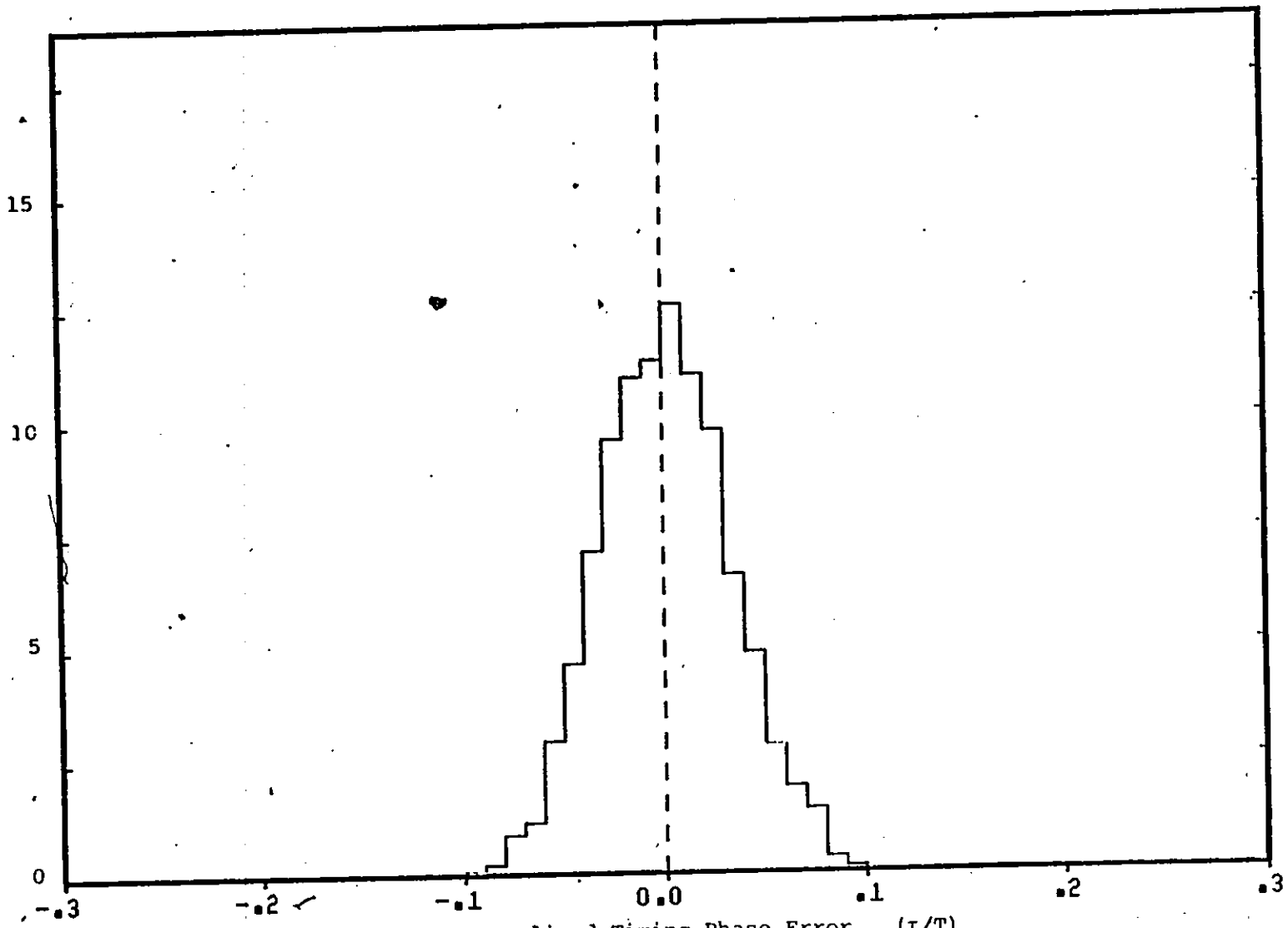


FIGURE 4.13 STEADY STATE PDF FOR TIMING RECOVERY LOOP (Δ=5)

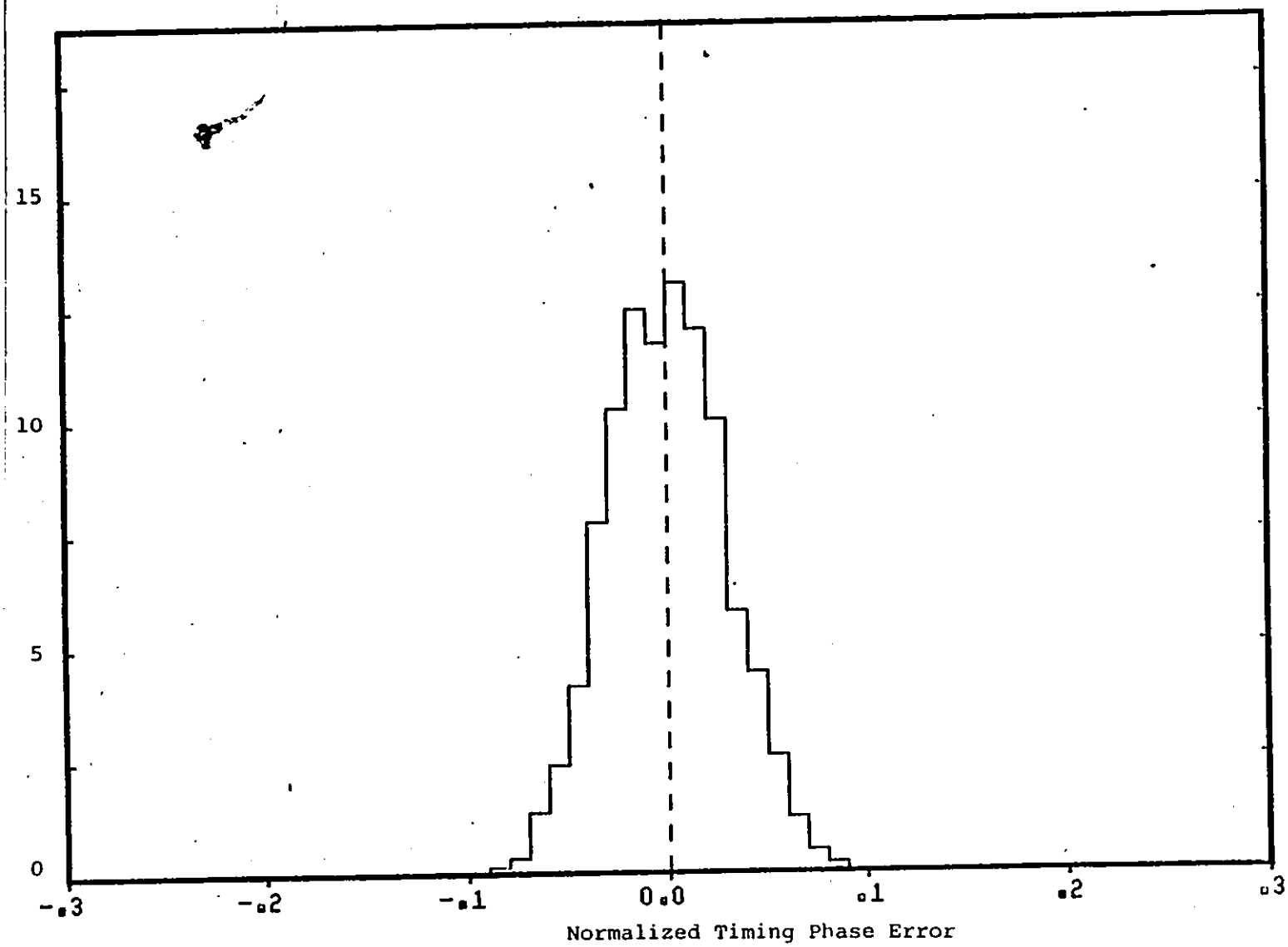


FIGURE 4.14 steady State PDF for Timing Recovery Loop ($\Delta=\infty$)

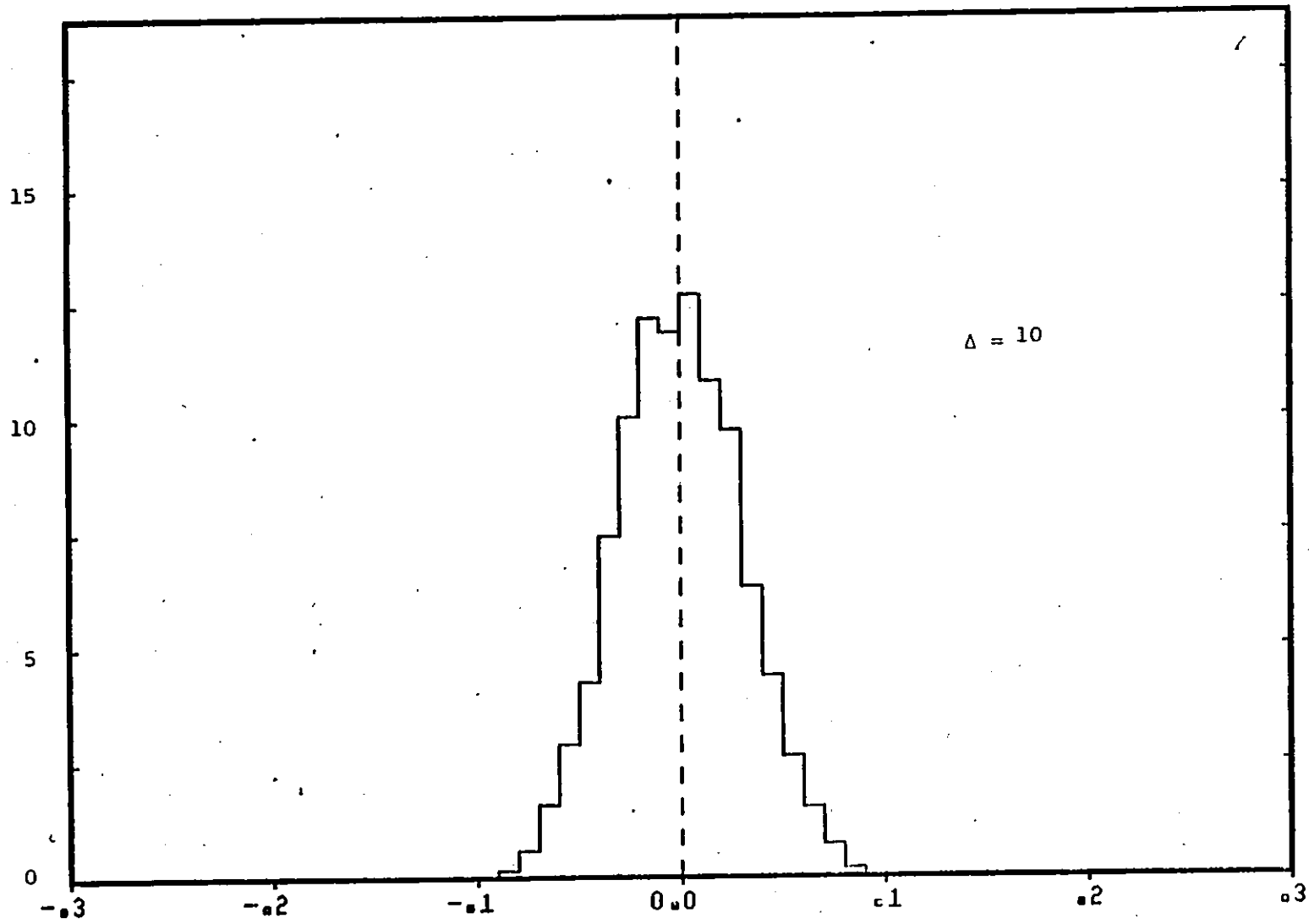


FIGURE 4.15 STEADY STATE PDF FOR TIMING RECOVERY LOOP

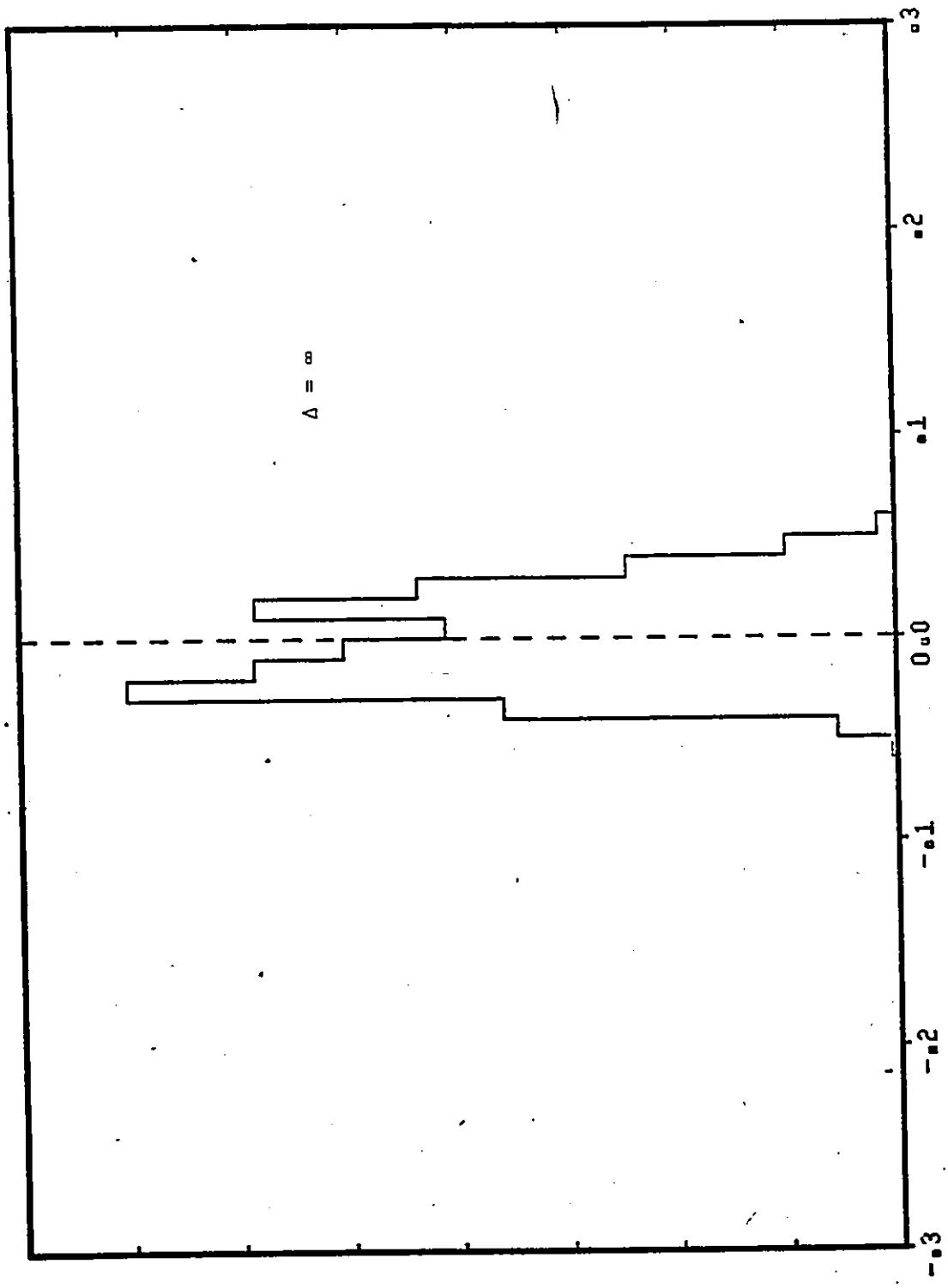


FIGURE 4.16 STEADY STATE PDF OF TIMING PHASE ERROR (binary data)

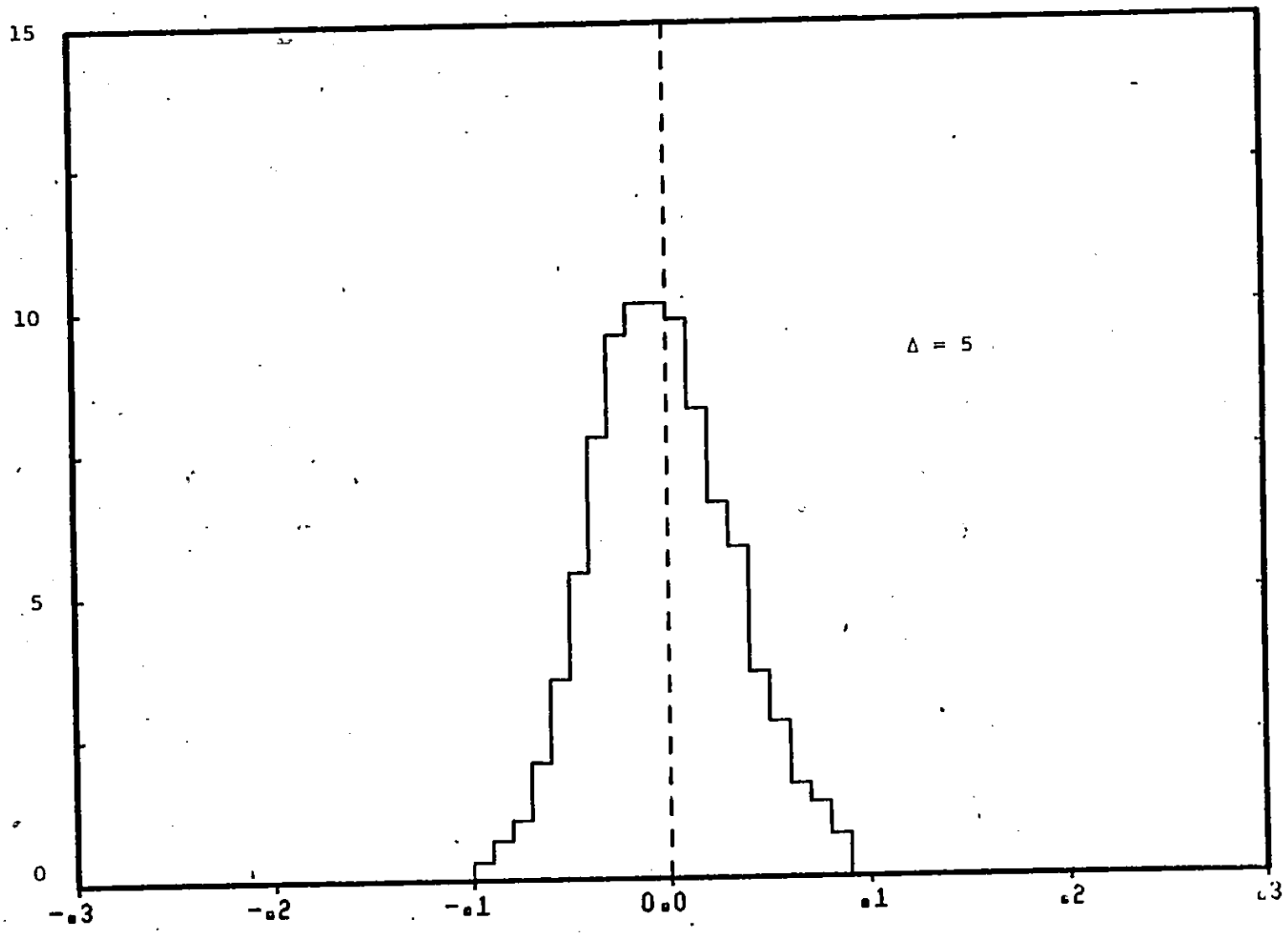


FIGURE 4.17 STEADY STATE PDF OF TIMING RECOVERY LOOP

each adjustment. Fast convergence behaviour is observed. Typically around sixty corrections are required to achieve the mean lock point with $\gamma = 0.1$. Some of these convergence characteristics or acquisition behaviours are shown in Figure 4.18. If a large loop gain is used, the acquisition time is shorter; however, the phase jitter or RMS timing error will be larger. In general, the timing phase will be uniformly distributed over the full period, and therefore the average lock-in time would be less than sixty adjustments. Occasionally, the loop might slip a cycle before locking in. This has been observed to happen when the loop is started at the other extreme, $\varepsilon = 0$. However the acquisition time is almost identical. This is because $\varepsilon = 0$ of one cycle is $\varepsilon = 1$ of the next cycle. If both fast convergence and low jitter is required, a gear-shifting mechanism can be used, so that during acquisition, a high loop gain will be used; after a predetermined time has elapsed, the loop gain can be shifted to a smaller value so as to achieve smaller RMS jitter. Such an arrangement has been simulated, using an initial value of $\gamma = 0.1$ and switching to a value of $\gamma = 0.01$ after 40 adjustments. The lock-in characteristics are shown in Figure 4.19. The lock-in time is comparable with that shown in Figure 4.18, whereas much smaller timing jitter is achieved.

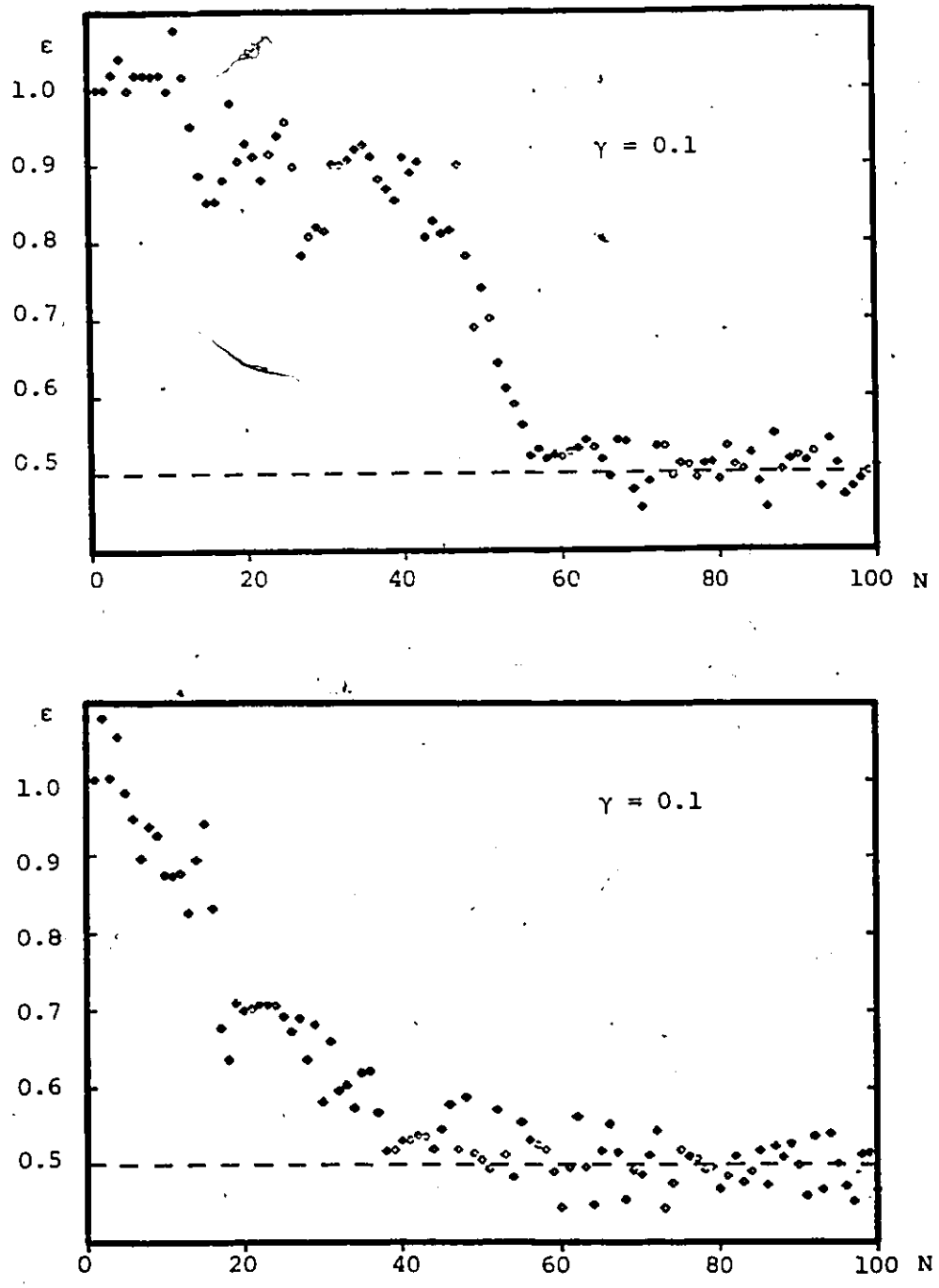


FIGURE 4.18 TYPICAL ACQUISITION BEHAVIOUR OF TIMING RECOVERY LOOP (4-LEVEL DATA) (NOISE FREE) N is the number of adjustments

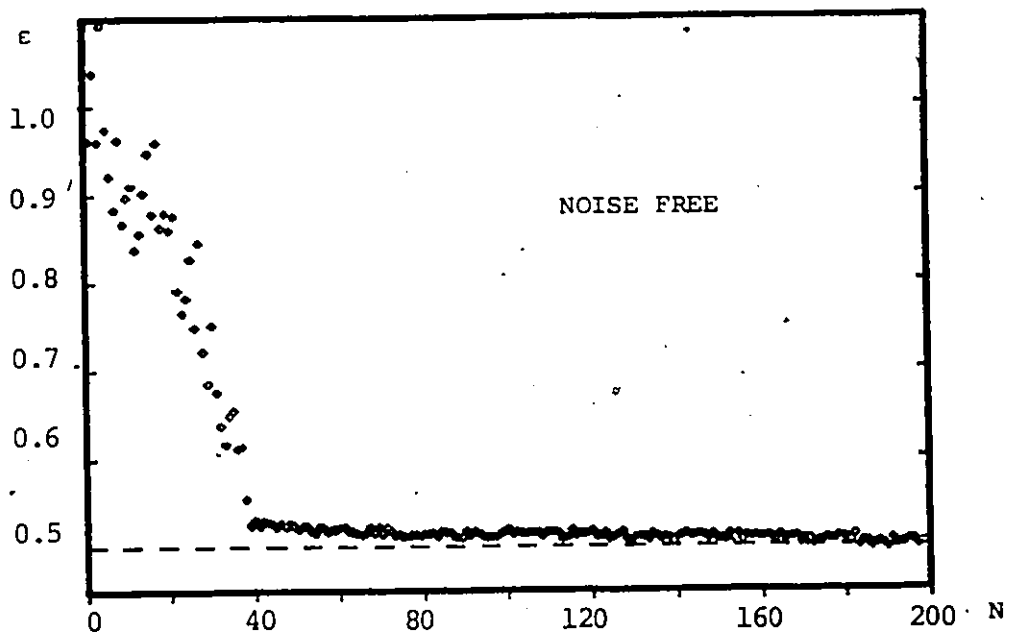
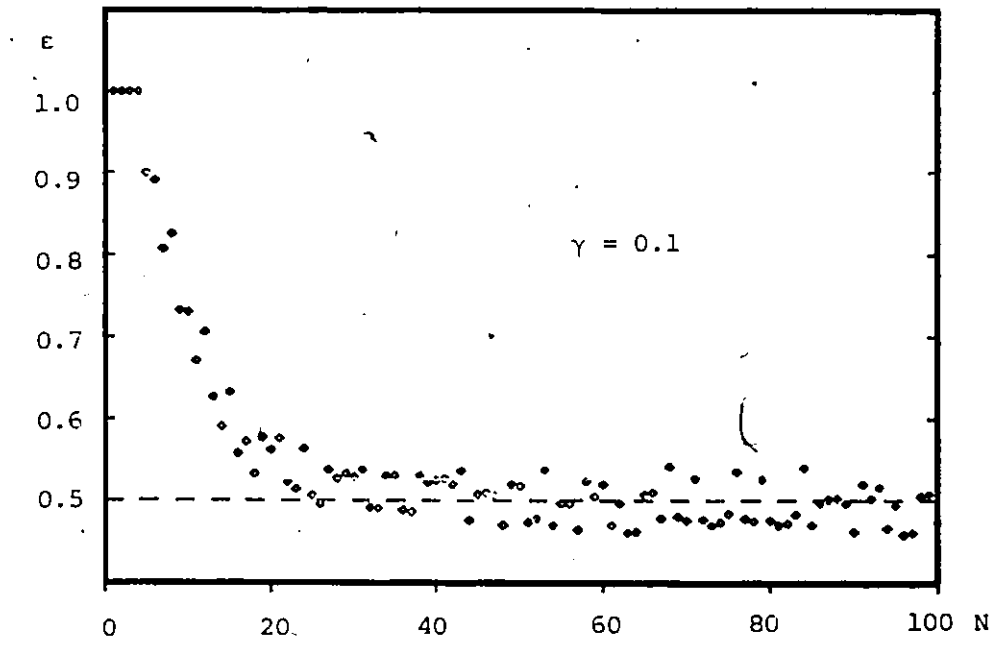


FIGURE 4.19 LOCK-IN CHARACTERISTICS WITH GEAR-SHIFTING TO ACHIEVE LOW RMS JITTER

Simulations have shown that a training sequence is not required for acquisition. Decision directed startup will work quite satisfactorily. However, if a training sequence is used, the timing acquisition period would be even shorter, because with ideal decisions, the error signal is closer to the ideal timing function, and hence provides a large correcting signal to achieve faster lock-in.

The acquisition time is shorter when the data is binary. This is because the eyewidth is larger for binary data, and therefore correct decisions will be obtained even at larger timing error, where the eye is closed for 4-level data. With correct decisions, the large error signal can be used to correct the timing phase error more quickly. Typical acquisition characteristics with binary data are shown in Figure 4.20. Usually, only twenty adjustments are required instead of the fifty to sixty adjustments required when the data is 4-level.

Noise does not seem to affect the acquisition time too much. Figure 4.21 shows the comparison between the acquisition behaviour with and without noise. The same data sequence is used for both simulations, the only difference is that the lower one has Gaussian noise added as in Figure 4.12. The noise tends to increase the phase jitter, and the acquisition time stays approximately unchanged. Therefore, the acquisition time can also be shortened by using binary



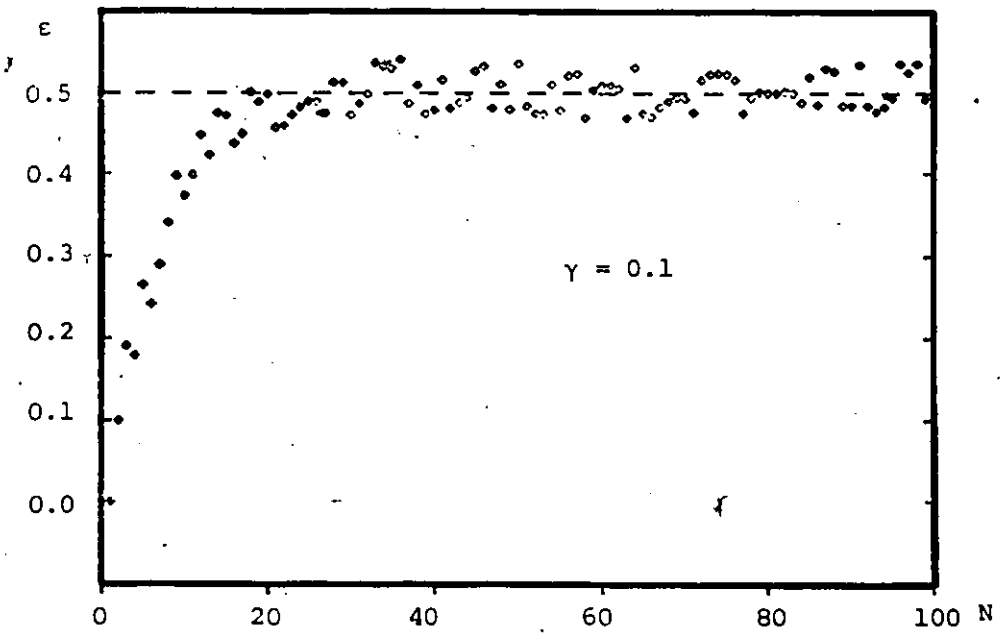
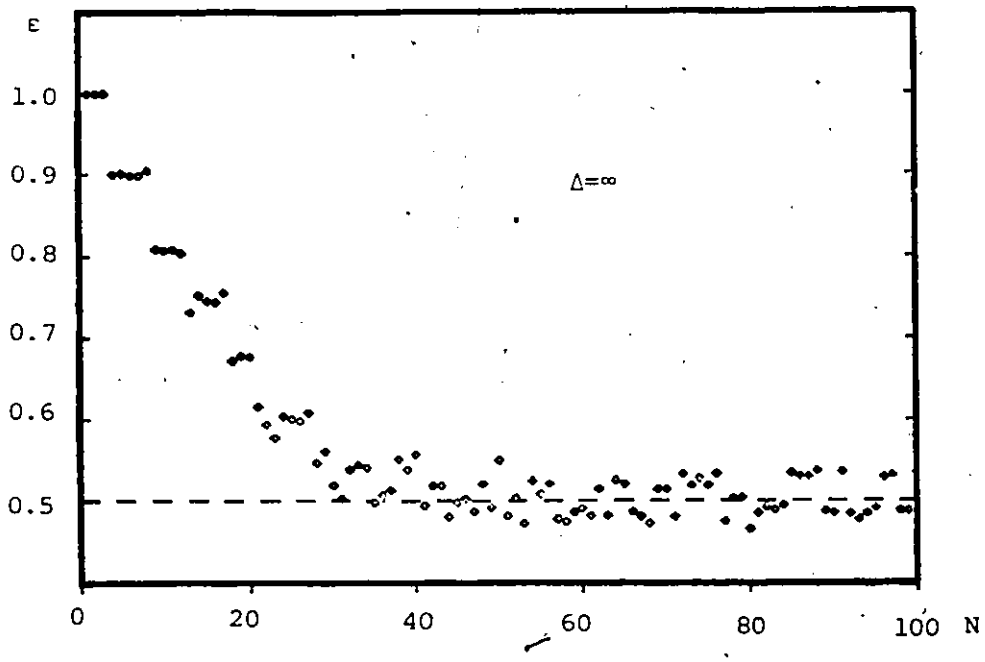


FIGURE 4.20 ACQUISITION CHARACTERISTICS OF TIMING RECOVERY LOOP WITH BINARY DATA



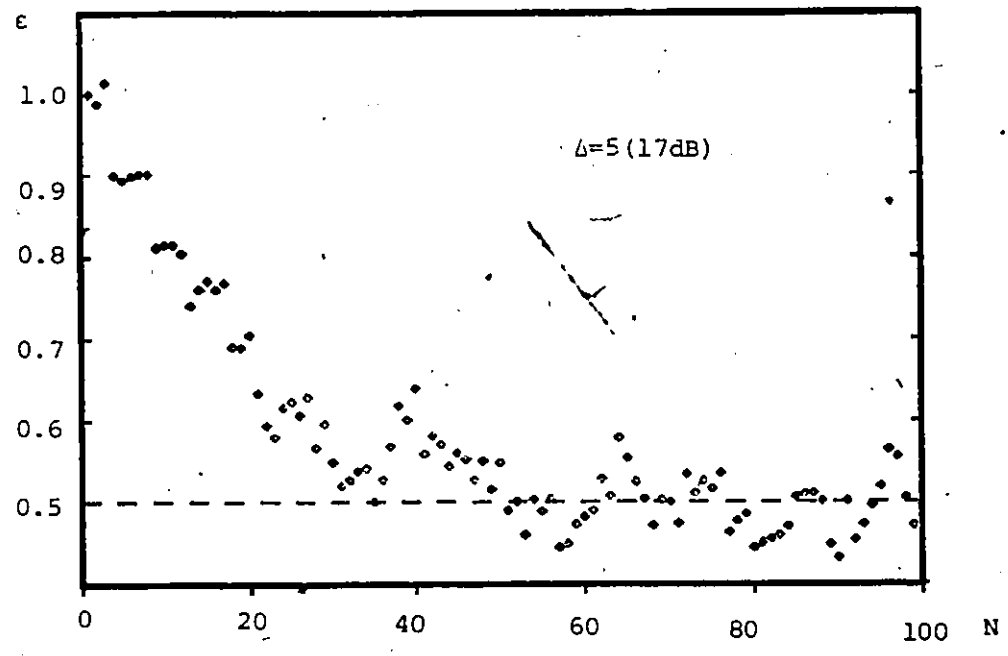


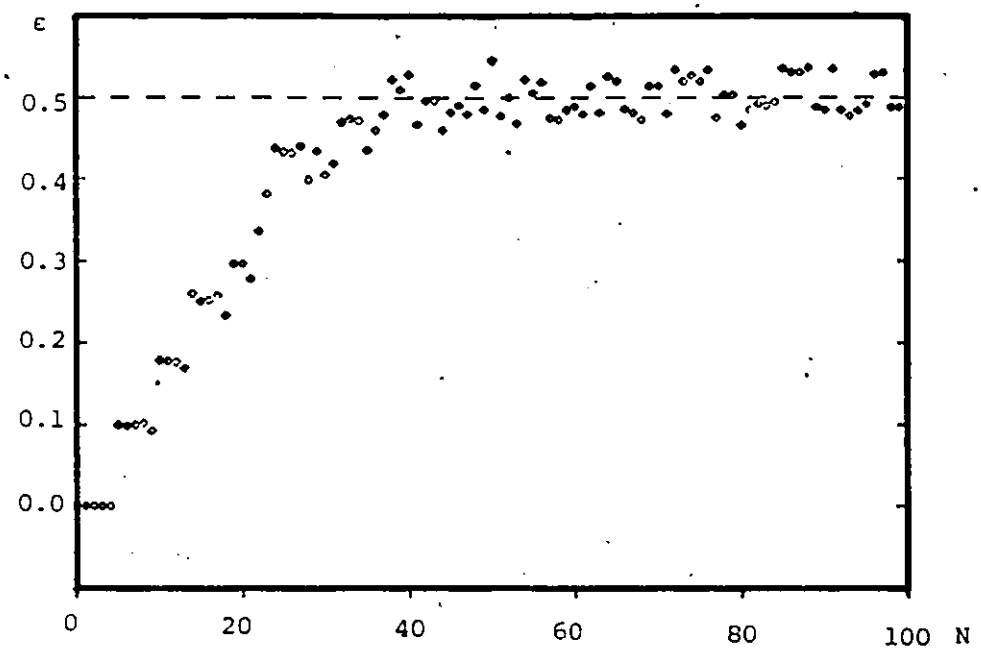
FIGURE 4.21 EFFECT OF NOISE IN ACQUISITION TIME

data to startup and when lock-up is achieved, 4-level data can be transmitted. When an actual circuit is to be designed, tradeoffs between different parameters will have to be considered in detail.

The problem of hang-up at $\epsilon = 0$ is mentioned in section 4.1.3. No hang-up has been observed in simulation even with no noise. Typical lock-in characteristics started at $\epsilon = 0$ are shown in Figure 4.22. The simulated timing error signal, is an average quantity, the actual value will fluctuate about its mean. That is, the error signal at $\epsilon = 0$ need not be identically zero, and hence the loop will move out of this unstable false lock point eventually. In practice, some noise will be present, this again will help to drive the loop out of the unstable lock point.

4.4 CONCLUSION

It has been shown that fast acquisition can be achieved by either starting up with binary data or using a large loop gain. Small timing jitter can be achieved using a small loop gain. Good steady state and acquisition characteristics can both be achieved using this loop structure. Furthermore, this loop will operate with both binary and 4-level data without modification. With this timing recovery loop, the stringent timing requirements of the system can be achieved without difficulty.



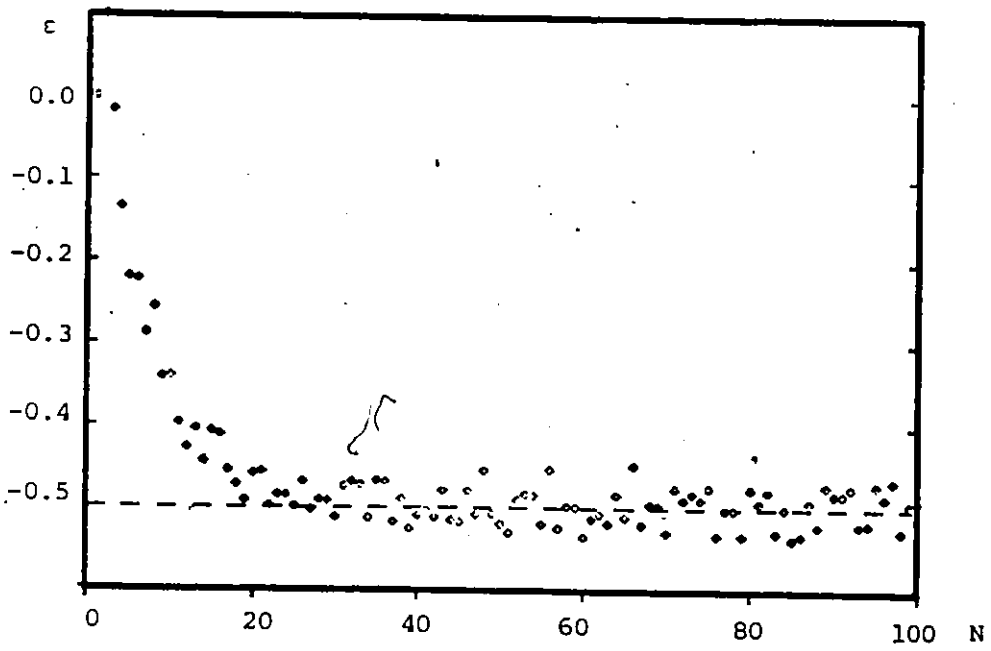


FIGURE 4.22 LOCK-IN CHARACTERISTICS (2-LEVEL DATA) (noise free) starting at $\varepsilon = 0$

CHAPTER 5

CONCLUSION

A feasibility study on a 4 bits/s/Hz digital radio system has been performed. This system uses a 16-QAM signal set, with Partial Response filtering to achieve signalling at the Nyquist rate for a given bandwidth. The problem of power amplification of the Partial Response signal using nonlinear traveling wave tube amplifiers is resolved. The error performance of the system using a decision feedback decoder (DFE) is analyzed by modelling the DFE as a Markov Chain so as to include the error propagation effects. A decision directed carrier recovery loop for this system is analyzed both theoretically, using the Fokker-Planck technique, and by simulation. With a reasonably narrow band loop, an RMS phase error of less than one degree can be achieved. A timing recovery loop for extracting the symbol synchronization information required is designed and analyzed by detailed simulations. Both the steady state performance and acquisition behaviour are excellent.

A complete repeater structure has been presented. All the components required for implementation are available off the shelf. No new technology is required to update the

DRS-8 system from a 2.25 bits/s/Hz system to a 4 bits/s/Hz system. An error rate of 10^{-5} can be achieved with an average signal to noise ratio of approximately 23-25 dB.

With the repeater structure described in this thesis, the system can be easily "switched" from a 4 bits/s/Hz system back to a 2 bits/s/Hz system to achieve an improvement in signal to noise ratio and hence error rate. This feature is highly desirable in situations such as heavy fading or severe multi-path conditions.

At the transmitter, if modulator 2 (Figure 2.6) is turned off, the resultant waveform before the Duobinary shaping filter is a QPSK or a binary QAM signal. At the receiver, the same carrier recovery loop and symbol timing tracking loop can be used without modification. Therefore the complete structure stays unchanged, except for some minor modifications in the serial to parallel converter at the transmitter and receiver. This modification to accommodate the change in data rate can be achieved easily and automatically at the digital logic level.

5.1 SUGGESTIONS FOR FUTURE WORK

Although the analysis has proved that this is a feasible system, there is still a lot of work to be done before the system can be put into operation. System parameters, such as loop bandwidth and filter response, have

to be chosen in an optimal fashion in order for the system to perform closely to the theoretical prediction. An end-to-end simulation of a repeater section should be done to evaluate the effects of actual filters, so as to optimize the system performance. The design of the sharing of pulse shaping between the transmitter and receiver is best performed by using such a simulation.

In this thesis, the timing loop is analyzed via simulations only. A theoretical analysis using Fokker-Planck techniques can be done on the symbol timing recovery loop by making appropriate assumptions to simplify the analysis. The interaction between the carrier recovery and symbol timing loop should be considered. Also, adjacent and co-channel interference effects on this system should be performed. The effects of adjacent channel interference from other modulation schemes such as 8-phase shift keying should be analyzed as well for completeness.

REFERENCES

[1] J.G. Smith, "A Bandwidth Compressive Modulation System", AIAA/CASI 6th Communication Satellite System Conference, Montreal Canada, April 5-8, 1976.

[2] P. Hartmann & J. Crossett, "A 90 Mb/s Digital Transmission System at 11 GHz using 8 PSK Modulation", ICC, Philadelphia, Pennsylvannia, June 14-16, pp. 18-8-18-13.

[3] P. Hartmann & E. Allen, "Design Considerations for a 3 bits/Hertz Digital Radio at 6 GHz", ICC 78, Toronto, Canada, June 4-7, 1978, pp. 33.1.3-33.1.3.

[4] A. Lender, "Seven Level Correlative Digital Transmission Over Radio", ICC 77, Chicago, Illinois, June 12-15, 1977, pp. 18-22-18-26.

[5] I. Godier, "DRS-8 Digital Radio for Long Haul Transmission", ICC 77, Chicago, Illinois, June 12-15,1977, pp. 5.4.102-5.4.105.

[6] B. Godfrey, Q. Chow & F. Halsey, "Practical Considerations in the Design of the DRS-8 System", ICC 77, Chicago, Illinois, June 12-15, 1977, pp. 5.4.106-5.4.110.

[7] S.G. Barber & C.W. Anderson, "Modulation Considerations for the DRS-8 91 Mb/s Digital Radio", ICC 77, Chicago, Illinois, June 12-15, 1977, pp. 5.4.111-5.4.115.

[8] D.P. Taylor & K.S. Yeung, "Conceptual Design of a Four bits/Hz Radio System", ICC 79, Boston, Mass., June 11-14, 1979, pp. 5.8.1-5.8.5.

[9] R.W. Lucky, J. Saltz & E.J. Weldon, Principles of Data Communications, McGraw Hill Book Company, 1968.

[10] A. Lender, "The Duobinary Technique for High-speed Data Transmission", IEEE Trans. on Comm. and Elect., Vol. 82, May 1963, pp. 214-218.

[11] E.R. Kretzmer, "Generalization of a Technique for Binary Data Communications", IEEE Trans. on Comm. Technology, Vol. COM-14, No.1, February 1966, pp. 67-68.

[12] A. Lender, "Correlative Level Coding for Binary Data Communication", IEEE Spectrum, February 1966.

[13] K. Miyauchi et al, "New Techniques for Generating and Detecting Multi-level Signal Formats", IEEE Trans. on Comm., Vol. COM-24, February 1976, pp. 263-267.

[14] W. Weber, P. Stanton & J. Sumida, "A Bandwidth-Compresive Modulation System Using Multi-Amplitude Minimum Shift Keying (MAMSK)" NTC-77, pp. 5.5.1-5.5.7.

[15] W. Wood, "Modulation and Filtering Techniques for 3 bits/Hertz Operation in the 6 GHz Frequency Band", ICC 77, Chicago, Illinois, June 12-15, 1977, pp. 5.3.97-5.3.101.

[16] J.A. Thomas, "A Simple Method of Clock Extraction From a Correlative Pulse Sequence", ICC 78, Toronto, Canada, June 4-7, 1978, pp. 49.3.1-49.3.4.

[17] P. Kabal & S. Pasupathy, "Partial Response Signalling", IEEE Trans. on Comm., Vol. COM-23, No. 9, September 1975, pp. 921-934.

[18] M.K. Simon & J.G. Smith, "Carrier Synchronization and Detection of QASK Signal Set", IEEE Trans on Comm., Vol. COM-22, No. 2, February 1974, pp. 98-106.

[19] M.K. Simon & J.G. Smith, "Offset Quadrature Communications with Decision Feedback Carrier Synchronization", IEEE Trans. on Comm., Vol. COM-22, No. 10, October 1974, pp. 1576-1584.

[20] D.P. Taylor & D. Cheung. "A Decision-Directed Carrier Recovery Loop for Duobinary Encoded Offset QPSK Signals", IEEE Trans. on Comm., Vol. COM-27, No. 2, February 1979.

[21] W.C. Lindsey, Synchronization Systems in Communication and Control, Englewood Cliffs, N.J., Prentice Hall, 1972.

[22] C.W. Anderson & S.G. Barber, "Modulation Considerations for a 91 Mb/s Digital Radio", IEEE Trans. on Comm., Vol. COM-26, No. 5, May 1978, pp. 523-527.

[23] P.A. Wintz & E.J. Luecke, "Performance of Optimum and Sub-Optimal Synchronizers", IEEE Trans. on Comm., Vol. COM-17, No. 3, 1969, pp. 380-389.

[24] H. Sailer, "Timing Recovery in Data Transmission Systems using Multi-level Partial Response Signalling", ICC 75, San Francisco, CA., June 16-18, 1975, pp. 34.24-34.27.

[25] W.C. Linsey & M.K. Simon, Telecommunication Systems Engineering, Prentice Hall, N.J., 1973.

[26] K.H. Mueller & M. Muller, Timing Recovery in Digital Synchronous Data Receivers", IEEE Trans. on Comm., Vol. COM-24, No. 5, May 1976, pp. 516-531.

[27] B.I. Halliwell (Ed.), Advanced Communication Systems, Newnes-Butterworths, 1974, p. 145.

[28] S. Haykin, Communication Systems, John Wiley & Co., 1978.

[29] Eisenberg, "Gated Phase-Locked Loop Study", IEEE Trans. on Aerospace and Electronic Systems, Vol. AES-7, No. 3, May 1971, pp. 469-477.

[30] W.R. Bennett, Introduction to Signal Transmission, McGraw Hill, 1970.

[31] W.R. Bennett, J.R. Davey, Data Transmission, McGraw Hill Book Company, New York, 1965.

[32] B.R. Saltzberg, "Timing Recovery for Synchronous Binary Data Transmission", BSTJ. No. 46 (1967) pp. 593-622.

APPENDIX 1

PROBABILITY OF ERROR FOR DUOBINARY SYSTEMS INCLUDING THE EFFECTS OF ERROR PROPAGATION

The decoder for a class I Partial Response signal is a decision feedback equalizer as shown in Figure A.1.

The current decision is made by first subtracting the previous decision from the present sample and passing the result through a quantizer. Therefore, any errors made in the previous estimate will most likely cause an error to occur in the present decision. This error propagation effect can be analyzed by modelling the system as a Markov Chain as done by Kabal and Pasupathy [17].

Since the encoding rule for Duobinary is:

$$y_n = a_n + a_{n-1} \tag{2.1}$$

r_n the sample with the estimated ISI removed, can be written as:

$$r_n = y_n + \zeta_n - \hat{a}_{n-1} \tag{A.1}$$

where ζ_n is the noise sample at time n.

Rewriting (A.1) we have,

$$r_n = a_n + \zeta_n - e_{n-1} \tag{A.2}$$

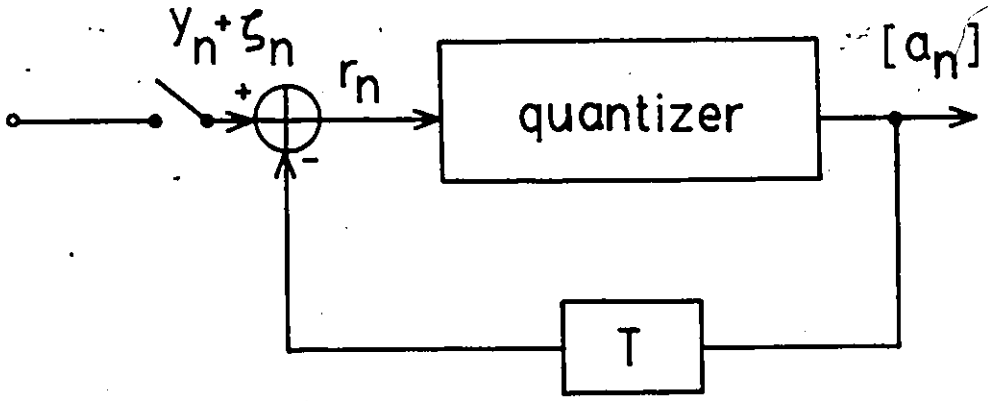


FIGURE A.1 DECODER FOR DUOBINARY ENCODED SIGNALS

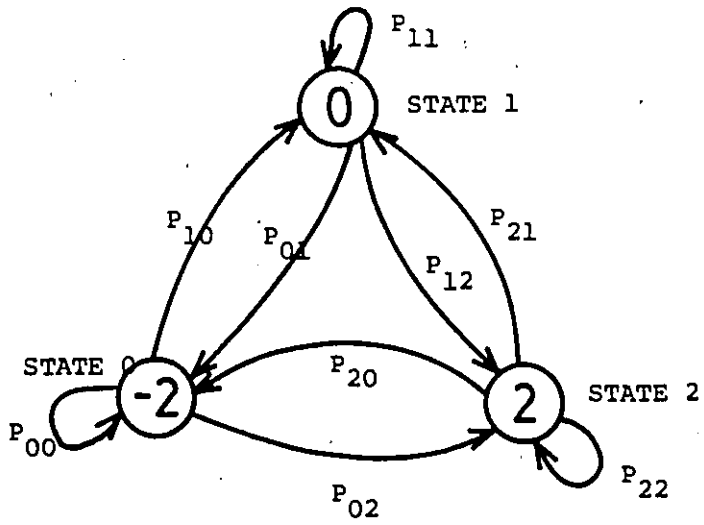


FIGURE A.2 STATE TRANSITION DIAGRAM FOR BINARY DATA

where $e_{n-1} = \hat{a}_{n-1} - a_{n-1}$, the error level at time n-1.

For binary inputs, e_n can take on values of -2, 0, 2, whereas for 4-level inputs, the error levels are -6, -4, -2, 0, 2, 4, 6. From (A.2) it is clear that the present decision will depend on the previous error e_{n-1}.

For binary inputs, we can model the decision feedback equalizer as a 3 state Markov Chain, where each state represents one of the error levels. These will be called error states. This Markov process is best represented by the state transition diagram in Figure A.2.

It is well known that if the process is ergodic, there will be a unique set of stationary probabilities associated with the error states. In this case, the total probability of error can be found, by summing all the stationary probabilities with error state not equal to zero. It can be shown that the decision feedback equalizer is ergodic [17]. Therefore, the probability of error including the effect of error propagation can be evaluated by calculating the stationary probabilities of the error states.

The P_{ij}'s in Figure A.2, are the state transition probabilities. They can be written in a compact matrix form which will be called the state transition matrix (STM).

The elements of the STM can be defined by:

$[STM]_{ij}$ is the transition probability from state j to state i.

Before the STM can be evaluated, all the possible values that each of the error states can take must be known. These are listed in Table A.1 for the 2-level input case.

From Table A.1, we can now write the entries of the state transition matrix explicitly.

$$[STM]_{oj} = P\,[e_n = -2 | e_{n-1} = j]$$
$$= P[a_{n-1} = 1]\; P[(\zeta_n - e_{n-1}) < -1]$$
$$= P[a_{n-1} = 1]\; Q[(1-e_{n-1})/\sigma] \qquad (A.3)$$

$$[STM]_{2j} = P\,[e_n = 2 | e_{n-1} = j]$$
$$= P[a_{n-1} = -1]\; Q[(1+e_{n-1})/\sigma] \qquad (A.4)$$

where σ^2 is the noise variance.

Assuming that the input binary symbols are equiprobable and also knowing that:

$$\sum_i P\,[e_i | e_{n-1}] = 1 \qquad (A.5)$$

we can obtain all the entries of the STM which is given as follows:

Present error state i	Previous error state j = -2	j = 0	j = 2
-2	$\frac{1}{2}\,Q[3/\sigma]$	$\frac{1}{2}\,Q[1/\sigma]$	$\frac{1}{2}\,Q[-1/\sigma]$
0	$1 - \frac{1}{2}\,Q[-1/\sigma] - \frac{1}{2}\,Q[3/\sigma]$	$1 - Q[1/\sigma]$	$1 - \frac{1}{2}\,Q[3/\sigma] - \frac{1}{2}\,Q[-1/\sigma]$
2	$\frac{1}{2}\,Q[-1/\sigma]$	$\frac{1}{2}\,Q[1/\sigma]$	$\frac{1}{2}\,Q[3/\sigma]$

TABLE A.1 POSSIBLE COMBINATIONS FOR ERROR STATES (2-LEVEL)

e_n	$\hat{a}_{n-1}$	a_{n-1}
-2	-1	1
0	-1	-1
	1	1
2	1	-1

The state probability vector at time k, $\underline{S}^k$ is related to $\underline{S}^{k-1}$ by:

$$\underline{S}^k = [STM]\ \underline{S}^{k-1} \tag{A.6}$$

Since this is an ergodic process, $\underline{S}^k$ will tend to a limit $\underline{S}$ as k increases. $\underline{S}$ is the stationary probabilities vector for the error states. This vector is an eigenvector of the STM. That is:

$$\underline{S} = [STM]\ \underline{S} \tag{A.7}$$

However, the sum of all the elements of $\underline{S}$ will be equal to 1. That is, this eigenvector corresponds to an eigenvalue of unity. This probability of error, including the effect of error propagation can then be determined from the stationary probabilities vector $\underline{S}$, using:

$$P_e = \sum_n P\ [e_n \neq 0\ ] \tag{A.8}$$

For binary inputs, the probability of error is then found to be:

$$P_e = Q[1/\sigma]/\ \{1 + Q[1/\sigma]\ - Q[-1/\sigma]/2 - Q[3/\sigma]/2\} \tag{A.9}$$

When the input to the Duobinary encoder consists of 4-level data, that is $a_n = \pm 1, \pm 3$, then e_n can take on seven

values as described earlier. Each of the seven states may be reached from any given state. Therefore, the STM is then a 7x7 matrix. The entries are as follows:

$$\begin{array}{c} \\ -6 \\ -4 \\ -2 \\ 0 \\ 2 \\ 4 \\ 6 \end{array} \begin{bmatrix} P[e_n=-6|e_{n-1}=-6] & \cdots\cdots & P[e_n=-6|e_{n-1}=6] \\ P[e_n=-4|e_{n-1}=-6] & \cdots\cdots & P[e_n=-4|e_{n-1}=6] \\ \vdots & & \vdots \\ \vdots & & \vdots \\ \vdots & & \vdots \\ \vdots & & \vdots \\ P[e_n=6|e_{n-1}=-6] & \cdots\cdots & P[e_n=6|e_{n-1}=6] \end{bmatrix}$$

Similarly, the possible combinations for each error state are listed in Table A.2, when the inputs are 4-level symbols. Again, making the assumption of equiprobable input statistics and with the help of Table A.2, we can then obtain the different transition probabilities as follows:

A For an error of -6 to occur

$a_n = 3$, $\hat{a}_n = -3$

For $\hat{a}_n = -3$, $r_n < -2$

that is $a_n + \zeta_n - e_{n-1} < -2$

or $\zeta_n < -5 + e_{n-1}$

Hence,

$$P[e_n=-6|e_{n-1}] = \frac{1}{4} P[\zeta_n < (-5+e_{n-1})]$$

TABLE A.2 POSSIBLE COMBINATIONS FOR ERROR STATES (4 LEVEL)

e_n	$\hat{a}_{n-1}$	a_{n-1}
-6	-3	3
-4	-3	-1
	1	-3
-2	-3	-1
	-1	1
	1	3
0	-3	-3
	-1	-1
	1	1
	3	3
2	-1	-3
	1	-1
	3	1
4	1	-3
	3	-1
6	3	-3

B For an error of -4 to occur, there are two combinations:

i For $\hat{a}_n$ to be decoded as -3, $r_n < -2$, that is,

$$a_n + \zeta_n - e_{n-1} < -2$$

$$\zeta_n < -3 + e_{n-1}$$

Probability is $\frac{1}{4} P [\zeta_n < -3 + e_{n-1}]$

ii For $\hat{a}_n$ to be decoded as -1, $-2 < r_n < 0$ that is,

$$-2 < a_n + \zeta_n - e_{n-1} < 0$$

$$-5 + e_{n-1} < \zeta_n < -3 + e_{n-1}$$

Probability is

$$\frac{1}{4} P[(-5 + e_{n-1}) < \zeta_n < (-3 + en_{-1})].$$

Hence,

$$P[e_n=-4|e_{n-1}] = \frac{1}{4} P[\zeta_n < (-3 + e_{n-1})] + \frac{1}{4} P[(-5 + e_{n-1}) < \zeta_n < (-3 + e_{n-1})]$$

C For an error of -2 to occur, there are three combinations:

i $a_n = -1$, $\hat{a}_n = -3$, for $\hat{a}_n$ to be decoded as -3, $r_n < -2$, that is,

$$a_n + \zeta_n - e_{n-1} < -2$$

$$\zeta_n < -1 + e_{n-1}$$

Probability is $\frac{1}{4} P [\zeta_n < (-1 + e_{n-1})]$

ii $a_n = 1$, $\hat{a}_n = -1$, for a_n to be decoded as -1,

$-2 < r_n < 0$

that is,

$$-2 < a_n + \zeta_n - e_{n-1} < 0$$

$$-3 + e_{n-1} < \zeta_n < -1 + e_{n-1}$$

Probability is

$$\frac{1}{4} P[(-3 + e_{n-1}) < \zeta_n < (-1 + e_{n-1})]$$

iii $a_n = 3$, $\hat{a}_n = 1$, for $\hat{a}_n$ to be decoded as 1,

$$0 < r_n < 2$$

$$0 < a_n + \zeta_n - e_{n-1} < 2$$

$$-3 + e_{n-1} < \zeta_n < -1 + e_{n-1}$$

Probability is

$$\frac{1}{4} P[(-3 + e_{n-1}) < \zeta_n < (-1 + e_{n-1})]$$

Hence the probability of getting to an error state of -2 is:

$$P[e_n=-2|e_{n-1}] = \frac{1}{4} P[\zeta_n < (-1 + e_{n-1})] + \frac{1}{4} P[(-3 + e_{n-1}) < \zeta_n < (-1 + e_{n-1})] + \frac{1}{4} P[(-3 + e_{n-1}) < \zeta_n < (-1 + e_{n-1})]$$

D Similarly, for an error of +2 to occur,

$$P[e_n=2|e_{n-1}] = \frac{1}{4} P[\zeta_n > (1+e_{n-1})] + \frac{1}{2} P[(1+e_{n-1}) < \zeta_n < (3+e_{n-1})]$$

E For an error of +4 to occur

$$P[e_n=4|e_{n-1}] = \frac{1}{4} P[\zeta_n > (3+e_{n-1})] + \frac{1}{4} P[(3+e_{n-1}) < \zeta_n < (5+e_{n-1})]$$

F For an error of +6 to occur

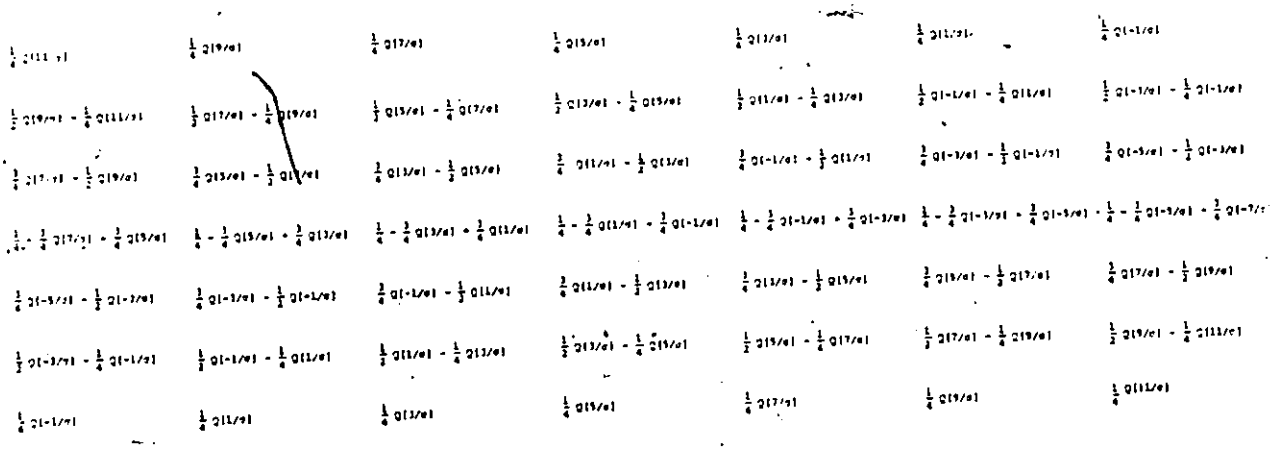
$$P[e_n=6|e_{n-1}] = \frac{1}{4} P[\zeta_n > (5+e_{n-1})]$$

G For an error of 0 to occur

$$P[e_n=0|e_{n-1}] = 1 - \sum_{\ell} P[e_n=\ell|e_{n-1}]$$

where $\ell = 6,4,2,-2,-4,-6$.

When writing the entries in a compact matrix form, it appears as follows:



Before the probability of error can be evaluated, the relationship between the average SNR and the noise variance has to be found. With the particular system model described in section 2.1.3, the noise power at the output of the receiver filter is given by:

$$P_n = (N_o/2) \int_{-\infty}^{\infty} |H_r(f)|^2 \, df$$

$$= (N_o/2) \int_{-1/2T}^{1/2T} df$$

$$= N_o/2T$$

$$= \sigma^2 \qquad (A.10)$$

The signal power at the receiver is given by:

$$P = \frac{E}{T} \int_{-\infty}^{\infty} |H_t(f)|^2 \, df \qquad (A.11)$$

where E is the variance of the transmitted symbols.

If the K level data are statistically equiprobable and are spaced at $\pm d, \pm 3, \ldots \pm(K-1)d$,

$$E = \frac{2}{K} \sum_{i=1}^{K/2} [\, d(2i-1)]^2 = \frac{d^2}{3} (K^2 - 1) \qquad (A.12)$$

Therefore, the average signal power is:

$$P_s = \frac{d}{3T}(K^2-1) \int_{-\infty}^{\infty} |H_t(f)|^2 \, df$$

$$= \frac{d}{3T}(K^2-1) \int_{-1/2T}^{1/2T} 4T^2 \cos^2 \pi fT \, df$$

$$P_s = 2\, d^2(K^2-1)/3 \qquad (A.13)$$

From (A.10) and (A.13), the average SNR may be written as:

$$P_s/P_n = 2\, d^2(K^2-1)/\{3\, \sigma^2\} \qquad (A.14)$$

Similarly, the average SNR for ideal PAM at the receiver can be written as:

$$P_s/P_n = d^2(K^2-1)/\{3\ \sigma^2\} \qquad (A.15)$$

With all this information, the error performance of the Duobinary system, using decision feedback decoder, can be evaluated and the results are shown in Figure 2.8. In practice, a numerical approach has been used to obtain the stationary probability vector, instead of evaluating it as an eigenvector of the STM. An initial state probability vector has been chosen, and this vector is repetitively multiplied by the STM until the stationary probability vector is obtained by applying Equation A.6 repeatedly until the vector converges.