A SIMPLE ADAPTIVE FSK MODEM

BY

ROBERT W. TIERNAY, B. Eng.

A SIMPLE ADAPTIVE FSK MODEM

A Simple Binary Frequency Shift Keying Transmitter and Self-Adapting Receiver for Data Transmission over Dispersive Telephone Channels

by

Robert Walter Tiernay, B. Eng.

A Thesis

Submitted to the Faculty of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Engineering

McMaster University

May 1972

MASTER OF ENGINEERING (1972) McMASTER UNIVERSITY
(Electrical Engineering) Hamilton, Ontario

TITLE: A Simple Adaptive FSK Modem

AUTHOR: Robert Walter Tiernay, B.Eng. (McMaster University)

SUPERVISOR: Professor A.S. Gladwin

NUMBER OF PAGES: 155

ABSTRACT: A simple and relatively inexpensive equalizer has been developed in conjuction with a FSK modulator and demodulator (modem) for the transmission of digital data over time dispersive channels.

The transmitter structure consists of a digital modulator followed by two filters, the first of which is a binary transversal filter (BTF), the second a conventional two-stage R-C low-pass filter.

In the receiver structure a phaselock loop (PLL) is employed as a discriminator. Synchronization is achieved with a digital phaselock loop (DPLL). Considerable attention is placed on the error rate optimization of the receiving structure.

The equalizer uses binary transversal filter comprising a tapped delay line, realized by a long string of metal-oxide semiconductor (MOS) flipflops, with a variable gain circuit associated with each tap in the delay line. The operation of the equalizer is adaptive in that the tap gains are adjusted rom the received data, consequently the

tap gains are continually optimized. The algorithm employed for setting the tap gains is simple and can be quite easily implemented using a combination of analog and digital techniques.

Experimental results are presented which illustrate the inportant features of the system implemented under a variety of operating conditions.

PREFACE

This work has aimed at the continuance of research in the area of adaptive digital communications. The main object of the thesis is to present and describe the adaptive binary frequency shift keying system developed both from a theoretical and practical viewpoint.

A concerted effort was made to develop a system which would prove to be in practice both simple and economical to implement , consequently most of the work was of a practical nature. A working model of the transmitter and adaptive receiver, which incorporated the novel FSK equalization scheme developed, was designed and implemented.

To provide a better understanding of the system portions of the thesis are devoted to presenting background theory, notably the first three chapters which describe the nature of the present switched network and a section of chapter six which is devoted to presenting the error mechanisms found in FSK systems.

ACKNOWLEDGEMENTS

The author would like to express his gratitude to his research director, Dr. A. S. Gladwin, for his assistance and guidance throughout the duration of this work. Thanks also go to Mr. D. Taylor for the many helpful discussions, and to Mr. J. Wolkowsky for his photographic work.

The undertaking of this work was made possible by financial support in the form of a scholarship from McMaster University and a grant from the National Research Council (grant A0902).

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF SYMBOLS

The following list contains those symbols which have been used repeatedly throughout the text. Because of the large number of quantities to be represented, it has been necessary to use some symbols to represent different quantities at different times and places. In every instance, the symbol has been defined where introduced to avoid misinterpretation of its meaning.

Symbol	Quantity	page
a_k	Transmitted pulse of length T	11
c_n	n^{th} tap gain of transversal filter	23
$d(t)$	Input to transversal filter	23
D	Peak distortion	19
e_n	Baseband error at n^{th} sampling instant	25
$\hat{e}_n$	Estimate of baseband error at the n^{th} sampling instant	27
f_n	Baseband output of transversal filter at n^{th} sampling instant	24
$f(\cdot)$	Function of baseband error and input baseband signal	26
g_n	Sampled received baseband waveform (at n^{th} sampling instant)	18
$H(\omega)$	Overall characteristic of transmitting and receiving filters	13
$H_{eq}(\omega)$	Equivalent Nyquist channel	16
I_d	Baseband output of ideal discriminator	59

Symbol	Quantity	page
K_{pd}	Phase detector gain constant	48
K_{vco}	VCO gain constant	50
$L(s)$	Loop filter of PLL	50
m_a	Output of analog FSK modulator	40
m_d	Output of digital FSK modulator	38
M_d	Mean square distortion at baseband	24
M_e	Mean square error at baseband	21
p_n	Input to TF of FSK equalizer at n^{th} sampling instant	80
P_e	Bit probability of error	18
$R(\omega)$	Receiving filter	13
s_n	Ideal received digit (transmitted digit)	21
T	Signalling interval	17
$T(\omega)$	Transmitting filter	13
v_{fpd}	Filtered phase detector output voltage	50
v_{pd}	Phase detector output voltage	48
v_n	Passband output of FSK equalizer at n^{th} sampling instant	80
$w(t)$	Impulse response of TF for impulse input	24
x_n	Sampled baseband impulse response	18
α_n	Phase of output of TF in FSK equalizer (n^{th} sampling instant)	81
$\Delta\omega_h$	Hold-in range of PLL	54
$\Delta\omega_p$	Pull-in range of PLL	54
ζ	Damping ratio of PLL	52
η	Click noise component	59

Symbol	Quantity	page
θ_i	Phase of input signal to PLL	48
θ_{vco}	Phase of VCO (in PLL)	48
κ_n	Amplitude of passband output of TF in FSK equalizer structure	81
λ_k	Discriminator output at the k^{th} sampling instant	65
μ	Non-click noise component of discriminator output	59
ρ	Iteration constant	26
σ	Iteration constant	26
τ	Time constant of lag filter employed in PLL	51
χ	Total noise contribution at output of ideal discriminator	60
ψ_n	Phase of input FSK signal at n^{th} samling instant	81
ω_c	Cut-off frequency of channel	12
ω_c	Carrier frequency of FSK signal	58
ω_d	Carrier deviation of FSK signal	58
ω_n	Natural frequency of PLL	52

CHAPTER ONE

PRELIMINARY DISCUSSION

1-1 The Historical Aspect

In the past decade, the extremely rapid development of digital technology has increased not only the demand of man-machine communications but also that of machine-machine communications.

Through the years, the telephone companies have implemented a very substanial communications network designed primarily for voice communications. This great demand for digital communications will result in the implementation of special networks designed primarily for that purpose. However, as the volume of data traffic increases, economic pressures will undoubtedly require that the digital transmission efficiency of all facilities be increased, and thus we can expect, at least for a few years that the existing network will continue to be utilized for data transmission.

1-2 Basic Statement of the Problem

An investigation of the channels currently being employed in the present voice network reveals that they are characterized by a relatively high signal to noise ratios (typically greater than 25dB), and narrow bandwidths (approximately 2.5 kHz). It is apparent that noise is not the limiting factor, however, the distortion of the trans-

mitted signal as it passes through the channel, created by the non-linear phase-frequency and non-constant amplitude-frequency characteristics which are typical of voice channels in the voice network - is. These two effects create so-called intersymbol interference (ISI), in which succeeding and preceding pulses interfere with the pulse being detected.

One method of alleviating this problem is to employ some fixed network at the receiver which compensates for the distortion incurred in the channel however, since it is often desired to use the switching facilities of the network, the application of this method is somewhat limited, since corrections can only be made on average channel characteristics, and thus all of the distortion can not be removed.

This drawback of "compromise" equalizers had lead to the development of various types of equalizers which can alter their impulse responses in order to minimize the ISI in some sense on the particular channel being used. Most work in this area has centered around linear modulation methods, since in general these methods are employed at high data rates where efficiency is of the greatest importance. However, for low-speed communication (less than 2000 bits per second) FSK is generally employed because of its economical implementation and the somewhat lessened requirement on bandwidth efficiency at these rates. For this reason and in consideration of the fact that the voice

network will still in all likelihood be employed for low speed traffic, an adaptive FSK modem has been implemented, and will be described in later chapters.

CHAPTER TWO

CHANNEL CHARACTERIZATION

2-1 Introduction

In this chapter the major impairments which limit the utilization of the channel for data transmission are described briefly and a simplified channel model is discussed.

2-2 The Transmission Path

A typical channel will normally consist of a so-called "local loop" which is in general a cable pair (or pairs) to the switching center at which point the signal is then filtered and modulated using standard frequency division multiplexing (FDM) techniques. The selection of the FDM channel will of course vary from connection to connection. This FDM signal (along with many others) is then transmitted over wire or microwave facilities, demodulated at the far end and then passed on through another local loop on the receiving end.

2-3 The Carrier Facility

(a) Filtering

The most restrictive section of the path just described is the FDM facility, for here the transmitted signal is bandpass filtered (approximately 300 Hz to 3000 Hz) in

order to ensure that co-channel interference (cross-talk) is minimized. However, the presence of these filters introduces another complication in that the very poor transmission characteristic below 300 Hz precludes the use of any signalling method which has dc, or very low frequency components, consequently some form of additional modulation must be employed (e.g. FSK, PSK, AM etc.).

The channel filters introduce another impairment in the form of time dispersion, which is directly attributable to the non-constant amplitude-frequency response across the band and to the sharp roll-off which the channel filters exhibit, which is usually associated with severe phase-frequency distortion near the band edges. Complete utilization of the channel bandwidth would result in severe ISI due to this latter effect.

(b) Carrier Offset and Phase Jitter

The single sideband suppressed carrier methods commonly employed in FDM systems require that for perfect demodulation that the carrier at the receiving end be of the same relative phase and exactly the same frequency as that at the transmitting end, however, both of these conditions are seldom, if ever met. The instability of the modulating and demodulating carrier supplies create a random jitter in the phase of the received signal which is termed "incidental-FM", since this jitter appears functionally as a low index frequency modulation. The severity of this effect on data

transmission depends on the quality of the equipment used in the FDM system and on the rate and type of data the system employed.

In addition to phase jitter, a finite difference in the average frequency of the modulating and demodulating carrier supplies usually exists. This effect is termed "frequency-offset" and in practice usually has little effect on the efficient use of the channel, since for systems employing coherent demodulation (such as AM) this effect can usually be eliminated, while for the case of incoherent systems (such as the discrimination detection of FSK) it is usually found that as long as the signalling rate is much greater than the frequency offset, that the effect of the offset is minimal.

Technically these two effects make the channel time-varying, however since in practice the effect of the frequency offset can be effectively neglected, as just mentioned, and that of phase jitter usually quite small (this is especially true of newer FDM systems), these two effects are usually considered to be of second order.

2-4 Cable Characteristics

Since most cable pairs are "lossy", the transmission loss is a function of frequency, (increasing with increasing frequency), furthermore, the phase-frequency response becomes distinctly non-linear for long cable lengths. In order to combat the former problem (since the latter has little effect

on voice communications) loading coils (lumped inductances) are introduced periodically on trunking facilities, and on local loops exceeding approximately three miles.

This has two effects, first, the overall amplitude-frequency response becomes irregular across the band and exhibits a sharp cut-off at some upper frequency. The second effect is to introduce "ripples" into the phase-frequency response and make this characteristic of the channel decidedly more non-linear.

2-5 Network Mismatch

The interconnection of the various components of the transmission path requires that several types of matching networks be used which can, when not properly aligned, introduce distortion in the overall characteristic. One of the more common difficulties encountered in this category is that of hybrid mismatch.

A hybrid is basically a four port network designed to interface a two wire circuit (such as a cable pair) to a four wire circuit (such as the FDM facility), the fourth port is used for balancing.

The network is so arranged, so that in perfect balance condition, the signal from the receive side of the four wire circuit passes directly to the two wire circuit, and the signal transmitted from the two wire circuit enters the send side of the four wire circuit. However, since the four wire circuit may connect to a great variety of two wire

circuits because of the switching involved, the perfect balance condition is never achieved because the compromise network used on the fourth port is not optimum for every connection. This results in some of the signal which was transmitted towards the two wire circuit being sent back along the four wire path. If an unbalance condition exists at two (or more) hybrids in the overall connection a portion of the signal can pass completely around the loop formed by the two hybrids and create what is known as "listener echo" - that is a portion of the transmitted signal can arrive a millisecond or more later than the original signal. This effect often creates severe effects in both the phase and amplitude-frequency response of the channel. On longer circuits this effect is reduced by introducing "echo suppressors" in to the four wire circuit, which sense which side of the line is being used and attenuate the other side.

2-6 Noise

Basically the noise encountered in the voice network consists of three components; a guassian background noise which is due to thermal noise in amplifiers, cables, etc. secondly, so-called impulse noise which results from exposure to switching equipment, lightning etc., and which is non-stationary by nature, and lastly "spurious noise" which is usually defined as consisting of discrete tones which are usually introduced by the FDM system or power facilities.

Impulse noise usually plays the major role in limit-

ing the efficient use of the channel, while the background guassian component is, of course, the ultimate limiting factor.

2-7 Non-Linear Effects

Non-linear effects arise from many different sources such as; saturation effects in transformers, non-linear amplifier characteristics, and compandors - which are used on carrier systems to maintain uniform loading by compressing the dynamic range of voice signals at the transmit side and expanding them at the other end. The effect of compandors can be particularly devastating to certain data systems, particularly those which have a high peak to average power ratio (such as AM).

2-8 Channel Model

The model most commonly used of the voice channel, for experimental as well as theoretical work, is that of a stationary dispersive filter whose output is summed with band-limited white noise. The justification for this is three-fold. Firstly, as previously discussed, time-varying effects such as phase jitter and frequency offset usually play only a secondary role in the overall systems performance. Secondly, in most cases the error rates of data transmission systems can usually be rank ordered for all noise environments by comparing performances in the presence of additive guassian noise. This is intuitively understandable, since

in most cases, errors are directly attributable to the size of an individual noise peak, and thus it should not matter to which distribution the noise peak belongs. Thirdly, non-linear effects (excluding compandors) are usually negligible.

Although the ensemble of channels encountered in the voice network is for all practical purposes infinite, approximate bounds can be placed on the characteristic of the dispersive filter used in the model (such as those shown in Appendix A), since in practice, specifications are followed in the implementation and maintenance of the various components of the transmission path.

CHAPTER THREE

LINEAR SYSTEMS

3-1 Introduction

Binary communications through the channel model previously described is considered in this chapter when linear modulation methods are used. Various means of optimizing the system to mitigate the effects of dispersion are considered.

3-2 The Equivalent Baseband Model

In the simplest configuration, illustrated in figure (3.1(a), a serial sequence of binary pulses (a_k) each of length T is presented to the transmitter. These digits are then translated in frequency using a linear modulation technique, transmitted through a channel, demodulated, sampled, sliced and then passed on to the receiving data terminal.

Since the baseband signal linearly modulates the carrier (using AM for example), and assuming that the receiver has knowledge of the correct phase and frequency of the carrier, then the overall system can be reduced to an equivalent baseband system as indicated in figure (3.1(b)). The channel filter here is, of course, continuous down to dc, but is limited to some upper frequency ω_c which is directly related to the passband of the original channel.

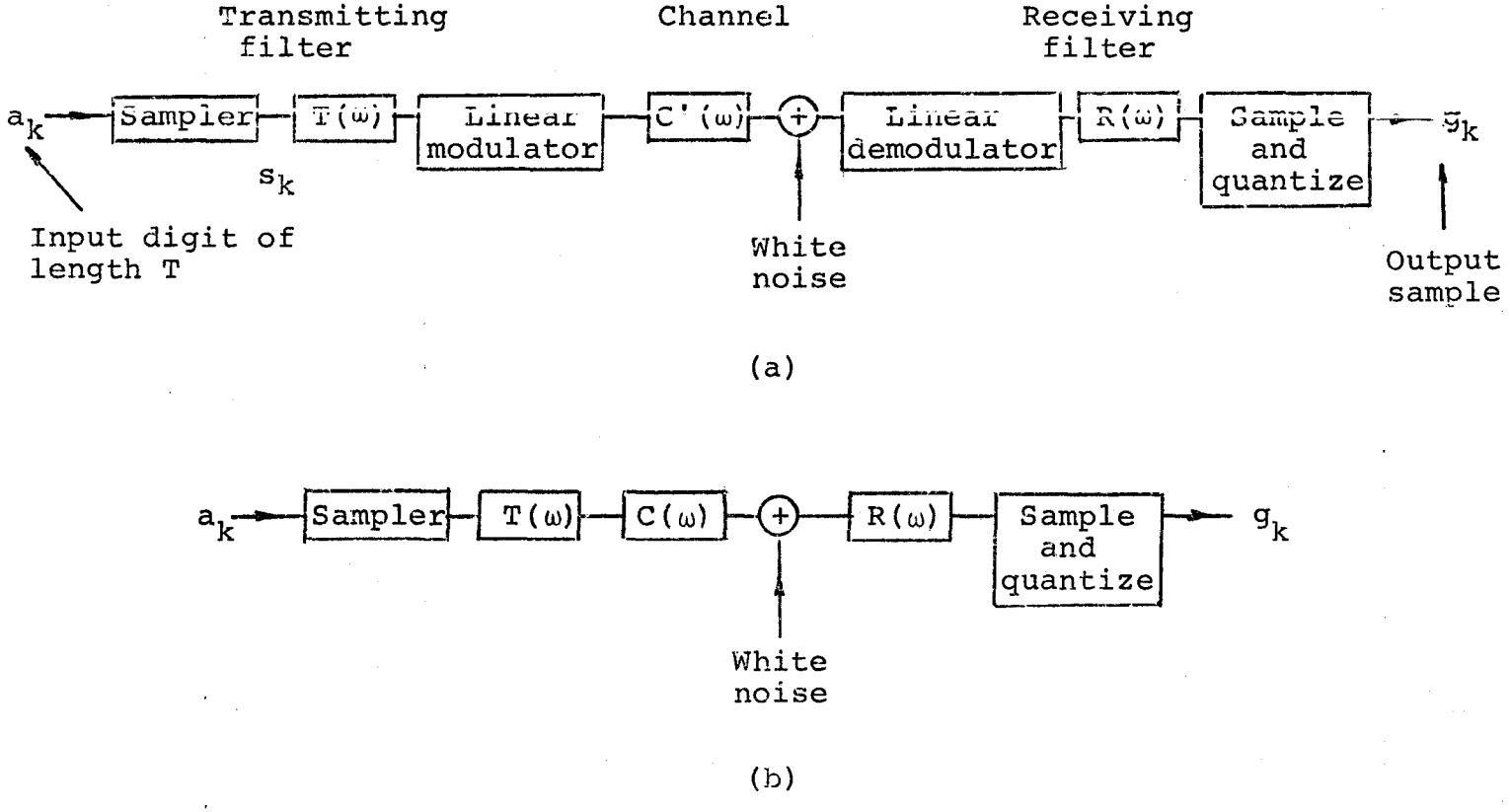


fig.3.1 Channel Models. (a) Passband. (b) Equivalent Baseband Model

3-3 The Effects of Limited Bandwidth

Let us assume for the moment that the channel filter in the reduced model is perfect, in the sense that the phase-frequency response is linear and the amplitude-frequency response is constant up to the cut-off frequency ω_c . We will constrain $\omega_c/2\pi$ to be much greater than the signalling rate 1/T for the time being. Since the channel is now band-limited, the impulse response of the channel is infinite in duration. This fact places a significant restriction on the selection of the overall transfer function $H(\omega)$, where

$$H(\omega) = T(\omega).R(\omega) \qquad (3.1)$$

since the design of this filter must ensure that the overall impulse response is a maximum at the desired sampling point, yet is zero at all adjacent sampling points which are a multiple of T seconds away, even though the channel is band-limited in some sense. If this last constraint is ignored, some of the transmitted pulses will interfere with each other and thus intersymbol interference will be present.

From the sampling theorem it is known that 2/T independent samples per second can be transmitted through $H(\omega)$ if this transfer function is of the form of an ideal low-pass filter with cut-off frequency $2\pi/T$. The impulse response of the ideal low-pass filter is, of course, the sinc function which conforms to the earlier requirement

made (that is, that the response be zero at all adjacent sampling points, even though the channel is band-limited). However, the use of the ideal low-pass filter characteristic would be undesirable since;

(1) The ideal low-pass filter is physically unrealizable.

(2) Any deviation in the pulse rate, filter cut-off frequency, or sampling instant would produce immediate failure since the overlapping tails of the sinc function represent a divergent series which can add up to large values.

With a more gradual roll-off of the transmitting and receiving filters, the oscillatory nature of the impulse response can be reduced and the above problems lessened. The design of the overall filter ($H(\omega)$) with this characteristic can be achieved with the use of Nyquist's theorem on vestigal symmetry.

3-4 The Nyquist Criterion

In a classic paper, Nyquist derived the restrictions on the transfer function $H(\omega)$ which result in the absence of ISI. To see how these restrictions arise, consider figure (3.2) when a "one" (of length T') is applied to the system.

Effectively an impulse is applied to $H(\omega)$. The Fourier transform of the sampled impulse response of $H(\omega)$ is

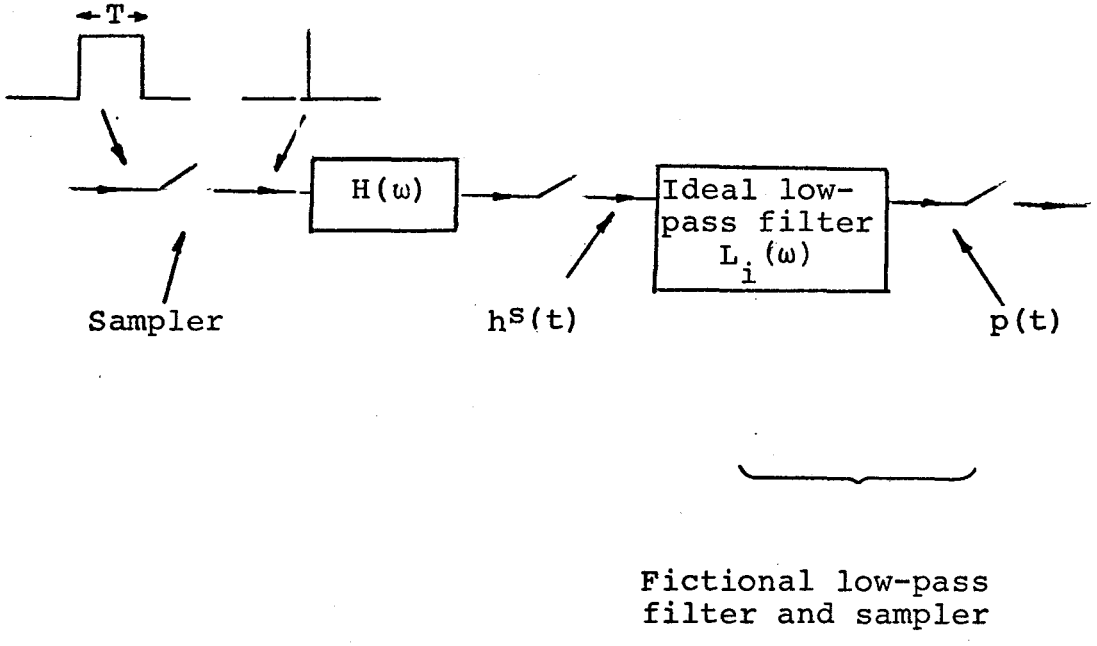


fig.3.2 Model for Developing the Nyquist Criterion

given by

$$F(h^s(t)) = \sum_k H(\omega+\frac{2\pi k}{T}) \tag{3.2}$$

and thus

$$\begin{aligned} F(p(t)) &= \sum_k H(\omega+\frac{2\pi k}{T}) && |\omega| \leq \frac{\pi}{T} \\ &= 0 && |\omega| > \frac{\pi}{T} \end{aligned} \tag{3.3}$$

Denoting

$$H_{eq}(\omega) = F(p(t)) \tag{3.4}$$

where $H_{eq}(\omega)$ is termed the "equivalent Nyquist channel". If $H_{eq}(\omega)$ has the characteristics of an ideal low-pass filter then the overall characteristic is that of an ideal low-pass filter. Furthermore, since the sampled output of the fictional low-pass filter $L_i(\omega)$ is the same as the sampled output of $H(\omega)$ (within a proportionality constant) by the sampling theorum, we can conclude that the condition for no ISI is that $H_{eq}(\omega)$ have the characteristics of an ideal low-pass filter - that is

$$\begin{aligned} |H_{eq}(\omega)| &= b && |\omega| \leq \pi \\ \frac{d}{d\omega} \angle H_{eq}(\omega) &= c. && |\omega| \leq \pi \end{aligned} \tag{3.5}$$

where b and c are constants.

From equation (3.5), it is apparent that there is an infinite class of filter characteristics which will not generate ISI even though the channel in which the system is operating is band-limited. One well known characteristic is that of the raised-cosine class which is real and is at once easily recognizable as one of the class of filters

which do not generate ISI, since the overlap created by sampling (equation (3.2)) produces the desired ideal low-pass characteristic in the desired range ($|\omega| \leq \frac{\pi}{T}$).

3-5 Optimum Reception for the Noisy Ideal Channel

The optimum Baysian receiver structure for a single pulse in a white noise background is, of course, a matched filter! Thus, we require for the reception of one pulse that

$$R(\omega) = T^*(\omega) \tag{3.6}$$

where the asterisk denotes complex conjugation. Once again we have tacitly assumed that $T(\omega)$ and $R(\omega)$ are well within the cut-off frequency ω_c.

If we assume that $T(\omega)$ is real (for the sake of simplicity) then

$$T(\omega) = T^*(\omega) \tag{3.7}$$

and thus for optimum reception

$$R(\omega) = T(\omega) = \sqrt{H(\omega)} \tag{3.8}$$

Furthermore, we can constrain $H(\omega)$ such that it be of the Nyquist class for no intersymbol interference, and thus ensure that every pulse is sampled independently of all others.[2] Thus the optimum receiver for symbol by symbol detection may be quite easily implemented by ensuring that the transmitting and receiving filters are described by equation (3.8).

3-6 The Effects of Significant Distortion

If we remove the constraint imposed earlier that the channel is ideal up to the cut-off frequency ω_c, then in general ISI will be introduced because the overall impulse response will not be of the Nyquist class. That is, denoting x(t) as the impulse response of the overall system, then

$$x_m \neq 0 \text{ for } m \neq 0 \tag{3.9}$$

where x_m is the sampled impulse response (asssuming a reference delay of zero). The sampled output of $R(\omega)$ at sampling time k can be written,

$$g_k = \sum_n s_k x_{k-n} + n_k \tag{3.10}$$

where the first term here represents a convolution of the input sequence and the overall impulse response. We can rewrite equation (3.10) as,†

$$g_k = s_k x_0 + \sum_n' s_k x_{k-n} + n_k \tag{3.11}$$

The second term here is, of course, the unnormalized instantaneous intersymbol interference, which we will denote as I_k. The probabilty of error at time k can be written (recalling that the amplitude of s_k is ± 1),

$$\begin{aligned} P_e = {} & \Pr\{(n_k+I_k) < -x_0\} \text{ given } s_k = +1 \\ & + \Pr\{(n_k+I_k) > x_0\} \text{ given } s_k = -1 \end{aligned} \tag{3.12}$$

from which it is apparent that the presense of ISI may have

† The notation $\sum'$ denotes summation over all n except the zeroth term.

a serious detrimental effect on the system's performance.

3-7 Performance Indices

(a) The Eye Pattern

In practice, the distribution I_k+n_k is readily observable by using an oscilloscope to sweep out the received baseband signal at a horizontal sweep rate of 1/T (or submultiple of 1/T). The resulting trace is called an "eye pattern" because of its resemblance in form to the human eye. The eye pattern is often used as an indicator of the system's performance, since information such as, timing sensitivity, noise margin, etc., can be obtained directly from this trace as shown in figure (3.3).

(b) Peak Distortion

One parameter which is directly obtainable from the eye pattern (in the absence of noise) is the peak distortion D. The peak distortion is the maximum normalized value of the instantaneous ISI, that is

$$D = (\max I_k)/x_0 \tag{3.13}$$

The maximum value of I_k occurs when a sequence of digits is sent such that the worst possible values of the impulse response add up at one time. Thus the worst case occurring will produce

$$I_k = \sum_n' |x_n| \tag{3.14}$$

consequently from equation (3.13),

$$D = 1/x_0 [\sum_n' |x_n|] \tag{3.15}$$

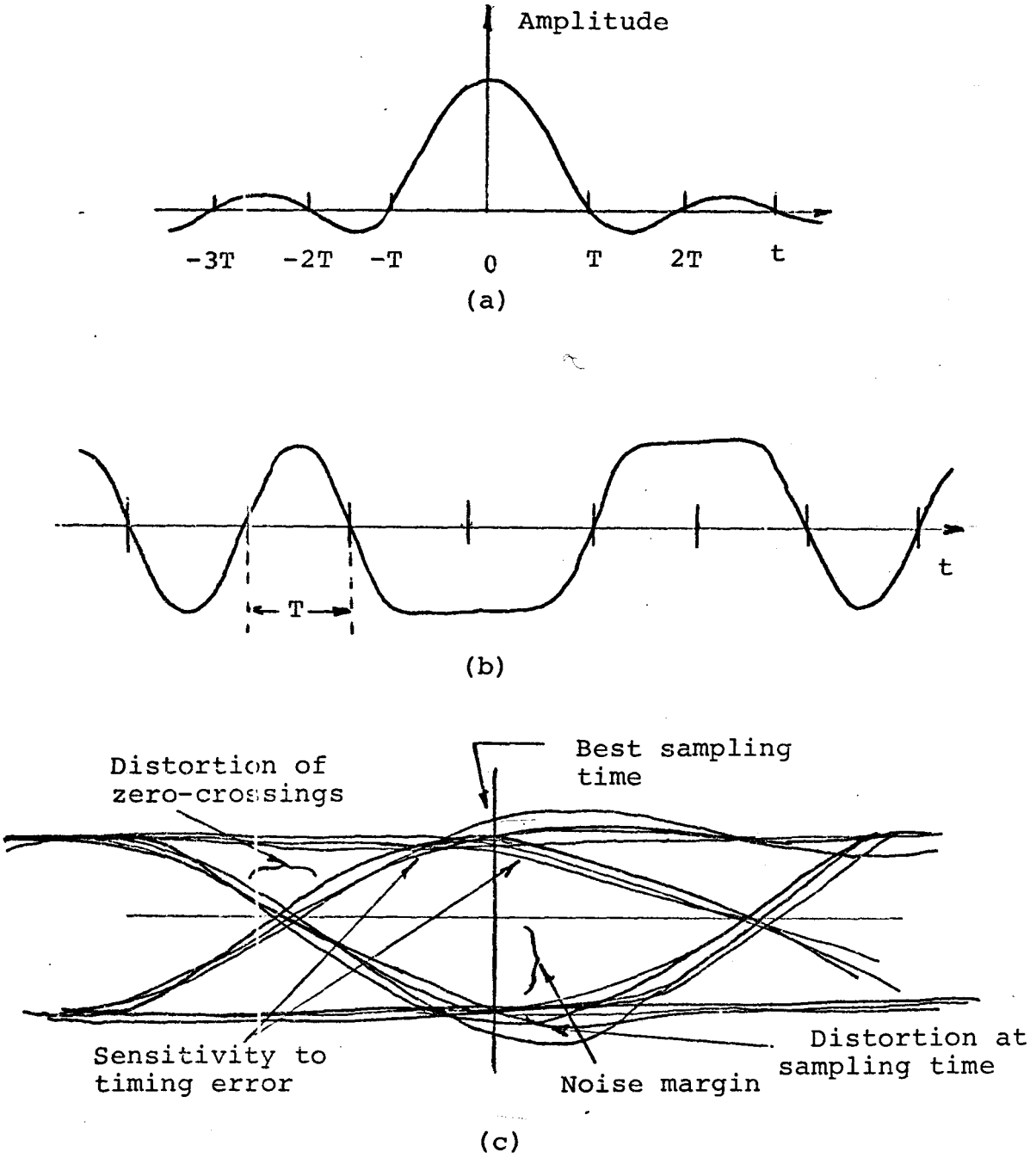


fig.3.3 Concept of the Eye Pattern. (a) Nyquist Pulse (b) Typical Undistorted Pulse Train. (c) Distorted Eye Pattern

(c) Mean Square Error

The mean square error is often used as an index of the system's performance because of its mathematical tractibility. The mean square error is defined as

$$M_e = \overline{(g_n - s_n)^2} \tag{3.16}$$

where s_n is the ideal received signal (i.e., the transmitted sequence). A bound can be placed on the probability of error from M_e which is usually somewhat tighter than that obtained by using D.

3-8 The Optimum Receiver for the Dispersive Channel

Intuitively, one would think that if D or M_e could be set to zero by the use of an external compensating network that the probabilty of error would be miminized, however, in fact the optimum Baysian receiver can be shown to be of the form of a bank of W filters (where WT is the effective channel memory time) followed by a non-linear sampled data system with memory.[3] The complexity of the sampled data system makes it unfeasible to implement this receiver, except in cases where the increase in performance is essential

If the receiver is constrained to be linear, and the data sequence limited in time to (2M+1)T, the optimum form of the receiver structure is that shown in figure (3.4). The structure can be seen to be that of a matched filter, followed by a transversal filter. In this latter filter,

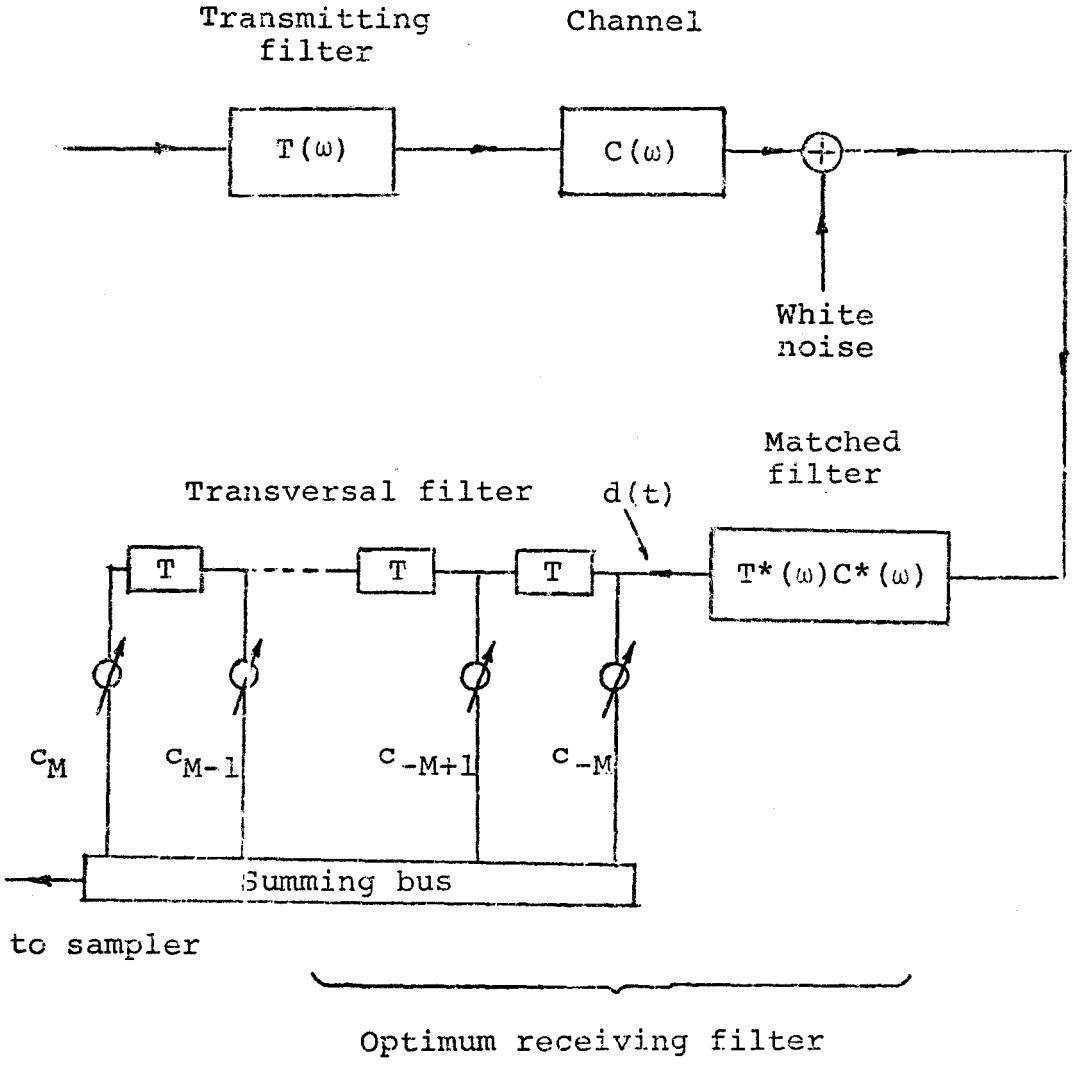


fig.3.4 Structure of the Optimum Linear Receiver

the output signal is obtained by taking a linear combination of delayed versions of the input signal, that is,

$$w(t) = \sum_{n=-M}^{M} c_n \cdot d(t-nT) \tag{3.17}$$

where the coefficients (c_n) may be positive or negative, d(t) is the input to the filter, and w(t) is the output. The frequency response of the TF is given by,

$$W(\omega) = \sum_{n=-M}^{M} c_n \cdot \exp(-jn\omega T) \tag{3.18}$$

3-9 Sub-Optimal Equalization - the Conventional Equalizer

The structure described above would be difficult to realize practically, since $C^*(\omega)$ must be found and also the 2M+1 coefficients of $W(\omega)$. This would require that measurements be made of the channels characteristics and the solving of M simultaneous non-linear equations to obtain the 2M+1 coefficients (c_n) which minimize the probabilty of error.[2] Since in general M is large, it is apparent that some alternate means is required to achieve the desired reduction in the probabilty of error.

One question arises at this point; since the ensemble of channels to be considered has a relatively high signal to noise ratio, with distortion playing the most important role in limiting the efficiency, why not use just the transversal filter to compensate for the overall distortion - with some penalty being incurred, of course, for removing the matched filter? This is , in fact what is done in designing the

majority of equalizers for telephone work.

After removing the restriction of optimality, the problem that now arises is how to determine the 2M+1 tap coefficients. One obvious method would be to use M_e or D as the criterion and set the coefficients accordingly.

3-10 Minimization of the Mean Square Distortion

In the absence of noise the resulting mean square error is referred to as the "mean square distortion" since the error is then only due to the channel characteristics. Thus we define

$$M_d = \overline{(f_n - s_n)^2} \tag{3.19}$$

where it is understood that the ampltude of the sample f_n is not noise dependent.

Let us now assume that f_n represents the sampled output of a transversal filter whose input is the received baseband sequence (g). The output of the TF at time n is from equation (3.17),

$$f_n = \sum_{j=-M}^{M} c_j \cdot g_{n-j} \tag{3.20}$$

and thus from equation (3.19)

$$M_d = \overline{\left(\sum_{j=-M}^{M} c_j \cdot g_{n-j} - s_n \right)^2} \tag{3.21}$$

From the above equation it is apparent that the mean square distortion is a second order function of the tap coefficients which suggests that in order to minimize M we consider the gradient with respect to the tap gains. The k^{th} element

of the gradient is

$$\frac{\partial M_d}{\partial c_k} = \overline{2\left(f_n \cdot \frac{\partial f_n}{\partial c_k} - s_n \cdot \frac{\partial f_n}{\partial c_k}\right)} \tag{3.22}$$

$$= \overline{2 \cdot e_n g_{n-k}} \tag{3.23}$$

where e_n is the error at the n^{th} sampling instant. For a minimum we require

$$\overline{e_n \cdot g_{n-k}} = 0 \tag{3.24}$$

That is, the optimum settings of the tap gains of the equalizer will force the cross-correlation between the input signal and the output error to be zero at each tap within the equalizer's structure.

3-11 Implementation of the Mean Square Distortion Minimizing Equalizer

Essentially, there are two means of implementing the above minimization technique. The first is called "preset equalization", in which the gradient is measured by transmitting a series of test pulses widely spaced in time and evaluating the cross-correlation at each tap.[4] The gains of the taps are moved in a direction opposite to the gradient, that is, towards the minimum. In the simplest case the sign of the gradient is determined after a suitable averaging period, and the equalizer taps are moved in the opposite direction to the gradient by a fixed increment. This system has two disadvantages, first the data stream must be held back for a period usually in the order of a

thousand bauds, so that equalization can take place; secondly, if the channel changes even slightly, the tap settings are no longer optimal in the mean square sense.

The second method is termed "adaptive equalization", in which the data stream is not held back, but rather is used in the minimization process. The method of steepest descent is again used to obtain the minimum. That is, one begins with an initial guess as to the optimal settings of the tap weights. Then, the unbiased estimate of the gradient ($e_n \cdot g_{n-k}$, for the k^{th} tap) determines the direction of movement, which is again in the opposite direction. More explicity, for the j^{th} tap in the $(p+1)^{th}$ iteration,

$$c_j^{p+1} = c_j^{p} - f(e_n \cdot g_{n-j}) \qquad (3.25)$$

where, in this case

$$f(e_n \cdot g_{n-j}) = \sigma \cdot e_n \cdot g_{n-j} \qquad (3.26)$$

There is a wide range of allowable functions f for which the minimization process will still take place. Perhaps one of the most simple functions in this class[5] is

$$f(e_n \cdot g_{n-j}) = \rho \cdot \mathrm{sgn}(e_n \cdot g_{n-j}) \qquad (3.27)$$

where ρ (and σ) is termed the iteration constant.

Equation (3.27) superficially appears easy to implement digitally, however it is apparent that a problem exists in determining e_n, in that the transmitted sequence s_n is not known at the receiver. However, the sequence delivered at the output of the threshold detector is in practice almost the same as that transmitted, provided the channel is

not severely distorted. Thus for purposes of adaptive equalization, the data delivered at the output of the equalizer can be employed as if it were the correct transmitted sequence. If the decisions are sufficiently good (say $P_e < 10^{-3}$), then in general, the equalizer can make a sufficiently accurate estimate of the gradient and thus move in a direction to reduce the MSE. This reduction improves the error rate and correspondingly, better estimates of the gradient can be made. However, the reverse can occur and the equalizer can lose complete acquisition of the channel. To be strictly correct, we should then rewrite e_n in equations (3.26) and (3.27) as $\hat{e}_n$, where $\hat{e}_n$ is the estimate of the error obtained by using the output data sequence. The resultant structure (algorithm described by equations (3.25) and (3.27)) is illustrated in figure (3.5). The integrators here serve to store the previous value of the tap gains. Since the output of the binary multiplier is ± 1, at each iteration the tap gains will go either up or down one step. The size of the step is determined by the gain of the integrators.

It is generally observed that most decision-directed equalizers suffer from an initial hesitation in the equalization process, however once the approximately correct equalization is acquired, the gradient estimates become quite accurate and the equalizer tracks easily - unless, of course, the channel should change faster than the time

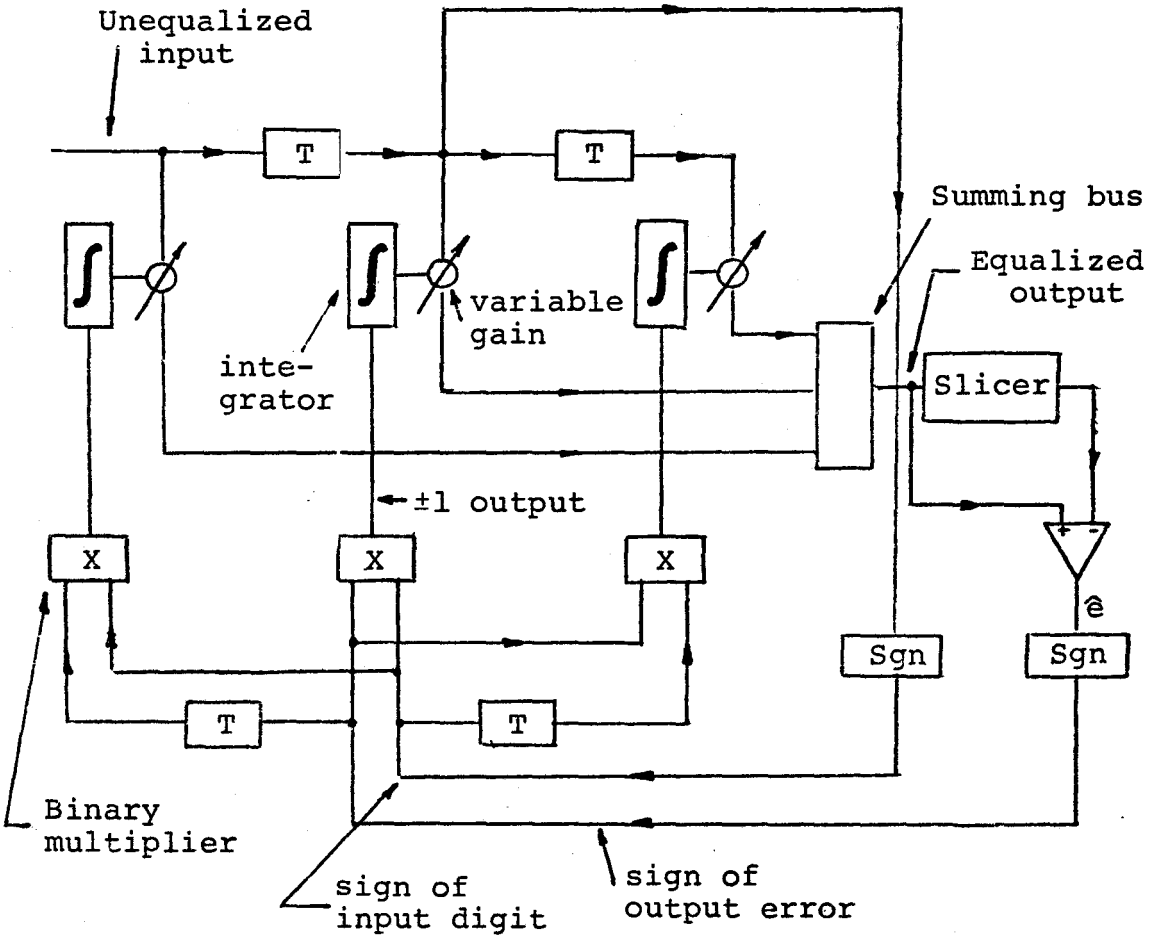


fig.3.5 MSE Minimizing Equalizer Structure

constant of the equalizer. In order to combat the problem of hesitation, preset equalization is often employed until the eye pattern "opens up). This technique is frequently used for high speed transmission where the initial distortion may be significant.

3-12 Minimization of the Total Mean Square Error

It can be shown that the algorithms described earlier which minimized the MSE in the absence of noise need not be modified when noise is introduced into the system.[6] The system still minimizes the total MSE - however, two significant differences in performance are introduced.

Firstly, the selection of the iteration constants (σ,ρ) becomes critical, whereas in the noiseless case the two constants could be made large in order to minimize the convergence time, the effect of noise will be to cause the tap gain vector to wander in a random fashion about its optimum setting. The smaller the iteration constants, the smaller the residual mean square error created by this effect will be. A decrease in the iteration constants has a disadvantage however, in that the convergence time will be increased. This problem is sometimes circumvented by employing a variable step size during the convergence process (decreasing with time), but this method has the disadvantage of being somewhat more costly to implement and furthermore tends to delay reconvergence if the character-

istics of the channel should change suddenly.

Thus in the presence of noise, the optimization of the equalizer's structure tends to be a more critical compromise between residual MSE and convergence time than in the noiseless case. Although somewhat crude estimates of the effects of varying the iteration constants can be made,[6] the more usual approach taken is to perform computer simulations and experimental investigations to determine the optimum choice for the particular application.

The second and perhaps the more important effect of noise in the system is that of noise enhancement. This latter effect can be seen from the following argument. Let us again assume that the overall shaping is Nyquist and furthermore that

$$R(\omega) = T(\omega) = \sqrt{H(\omega)} \tag{3.28}$$

where again $T(\omega)$ and $R(\omega)$ are assumed real. The noise variance is given by

$$P = \frac{N_0}{2\pi} \cdot \int_{-\infty}^{\infty} H(\omega) \cdot |W(\omega)|^2 d\omega \tag{3.29}$$

where N_0 is the noise spectral density (white noise assumed). From Parseval's theorem, equation (3.29) can be written.

$$P = N_0 \cdot \int_{\infty}^{\infty} |w(t)|^2 dt \tag{3.30}$$

which from equation 3.18 gives for the variance at the sampling times

$$P = N_0 \cdot \sum_{n=-N}^{N} c_n^2 \qquad (3.31)$$

Thus, it is apparent from equation (3.31) that in mitigating the effects of dispersion, that the effect of noise is enhanced by the sum of the square of the tap gains.

3-13 Limitations of the MSE Minimizing Equalizer

(a) Physical Realizability

As the number of taps is increased, it is apparent from equation (3.18) that apart from a finite delay, the characteristic of the equalizer becomes that of a Fourier series with real coefficients, and thus can approximate virtually any physically realizable function in the equivalent Nyquist channel. However, since the minimization process is in some sense trying to achieve the inverse of the channel, it is apparent that the presence of a null cannot be corrected for, since the equalizer will try to achieve an infinite gain at the null frequency - which in practice would be impossible to achieve.

(b) Tapped Delay Line Length

The length of the tapped delay line in the structure of the equalizer effectively determines the degree of correction which can be achieved for a given distorted channel. The shorter the TDL becomes, the larger the contribution to the residual MSE due to truncation of the series mentioned above becomes. This can be more readily appreciated if one considers the effect of truncation in the time

domain. As the TDL is made shorter, cross-correlations outside of the equalizer's structure become significant, but cannot be removed. The length of the TDL is thus determined by the signalling rate and the type of channel to be encountered, and thus will vary depending on the particular application.

(c) Non-Random Data

It has been tacitly assumed up to this point that the data stream is random in nature, however, one may well wonder what the equalizer will do if non-random data such as all 1's or 0's is sent. Clearly, it is impossible for the equalizer to make an identification of the gradient. Fortunately, it can be shown that when non-random data is sent that the settings of the tap gains tend to oscillate about the settings obtained by true random data until the repetitive sequence ends and meaningful data is once again transmitted.[2]

3.14 Minimization of Peak Distortion

When peak distortion is used as the criterion in the equalization process, an algorithm similar in form to that obtained for the MSE criterion, can be developed to set the tap gains (the so-called zero-forcing algorithm).[6,7] However, unlike the MSE equalizer, which was found to have a quadratic error surface (in non-decision-directed mode), the peak distortion minimizing algorithm has much more

restrictive convergence properties. In fact, convergence is only guaranteed if the initial distortion is less than unity (i.e., the eye pattern is open initially).[2] Furthermore, this algorithm tends to be much slower in converging. For these two reasons, most recent implementations have been of the MSE minimizing type.

CHAPTER FOUR

THE ADAPTIVE FSK MODEM

4-1 Introduction

One may well question the use of FSK as a modulation method, since other methods, for example VSB-AM, require less bandwidth and perform somewhat better as far as error rates are concerned. The main reason for employing FSK is one of economics. The FSK system is very simple to implement and provides good, stable operation over a wide variety of channels. This combination of simplicity and performance, has made FSK modulation the most popular choice for low speed transmission. However, for higher transmission rates, the increased efficiency required, often warrants the increased complexity of other systems (high data rates generally require multi-level signalling and thus more susceptible to noise).

Another important advantage of the FSK system is that a hard limiter is usually employed in the receiver-thus the signal at the output of the limiter is in a binary form and further processing can be performed quite readily using the binary format, in contrast to linear systems which require A/D convertors.

4-2 Simplified Description of the Modem

A block diagram of the modem is shown in figure (4.1)

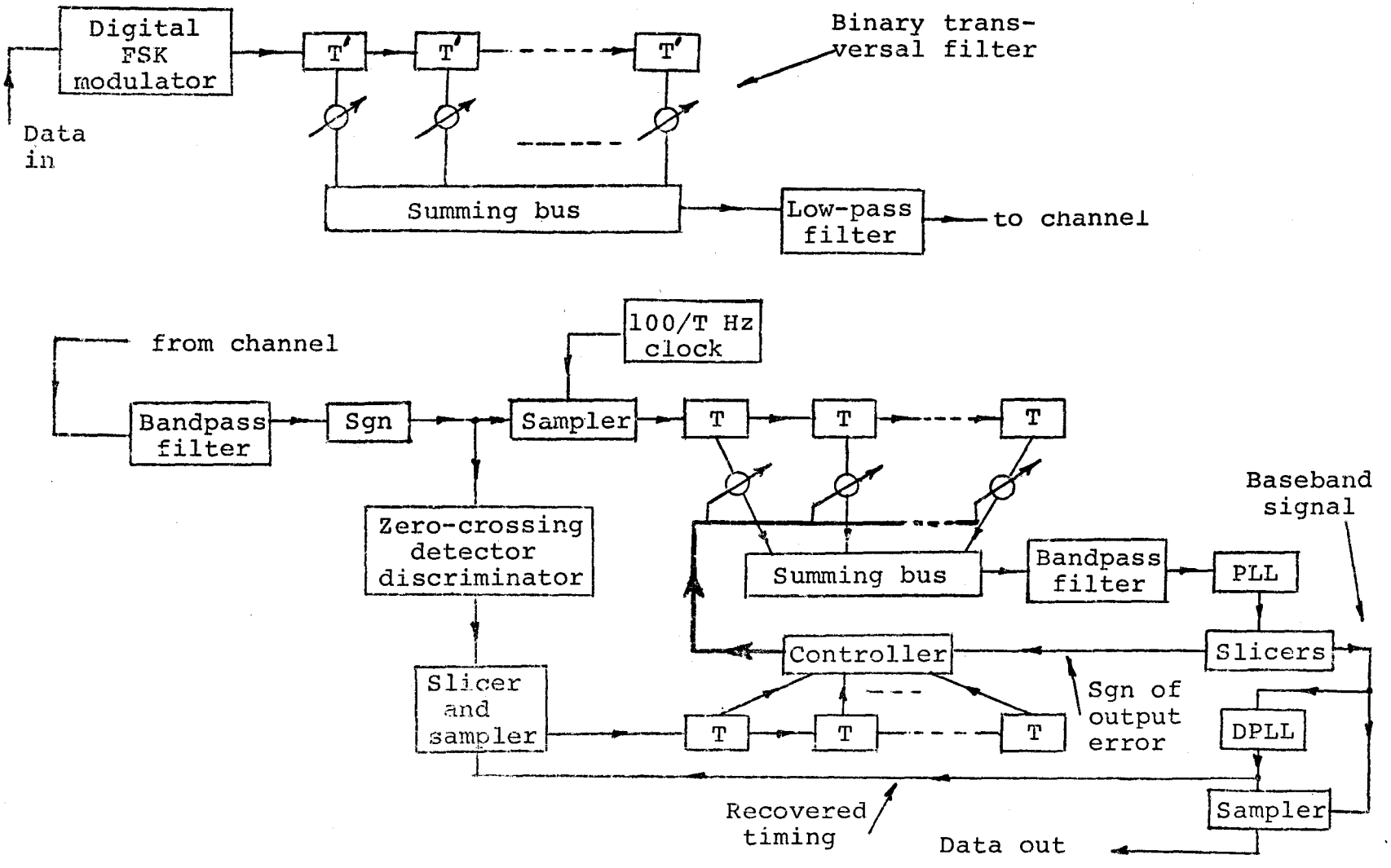


fig.4.1 Simplified Diagram of Modem Structure

a) Modulator

The input binary data enters in a serial form at the rate of 1200 bits per second. Depending on the sign of the input digit, a 1200 or 2400 Hz square wave is selected. This switching is performed coherently. The new binary pulse train is then filtered by a binary transversal filter (BTF), and then by a conventional low-pass filter. The output of the low-pass filter is then sent into the channel.

b) Receiver Structure

The received signal is initially band-pass filtered to remove noise outside the spectrum of the line signal. The output of the band-pass filter is then applied to a zero-crossings detector type discriminator (ZCDD) - the output of which is hard limited to obtain the sign of the input digit. This latter signal is then sampled and passed through a delay network consisting of a series of flip-flops, whose outputs are connected to the controller.

The hard limited FSK signal is also applied to a sampler, which samples this limited signal at 100 times the bit rate (100/T) Hz. This latter pulse train is then passed through a BTF. The taps of the BTF are adjusted by the controller to reduce the ISI. The output of the BTF is connected to a phaselock loop (PLL) discriminator. The baseband output of the PLL is sliced to obtain (a) the sign of the output symbol and (b) the sign of the error at the sampling time. The former signal is applied to a digital

phaselock loop (DPLL) which recovers the timing reference, while the latter quantity is delivered to the controller.

In the following chapters, each of the sections described briefly above is considered in more detail.

CHAPTER FIVE

THE MODULATOR

5-1 Introduction

This chapter is devoted to an analysis of the modulator employed in the modem. It should be pointed out at this time that no attempt was made to employ a bandwidth reduction technique, such as that suggested by de Buda,[1] nor was the joint optimization of the transmitting and receiving filters considered - other than in the terms described in this and the next chapter.

5-2 Description of the Digital Modulator

As illustrated in figure (5.1), the modulator consists of two square waves which can be switched digitally. In the example shown, the output of the digital modulator will be a 2400 Hz square wave if the input data is a one (positive logic convention), similarly if the input is a zero, the output will be a 1200 Hz square wave. The data must be synchronized to the two square wave sources. This is accomplished by sending the 1200 Hz square wave to the transmitting data equipment which then sends the data back synchronized to this clock. The output of the digital modulator can be expressed in Boolean form as,

$$m_d(t) = a_k \cdot q_1(t) + \bar{a}_k \cdot q_2(t) \qquad (5.1)$$

where: a_k is the k^{th} input digit of length T, which

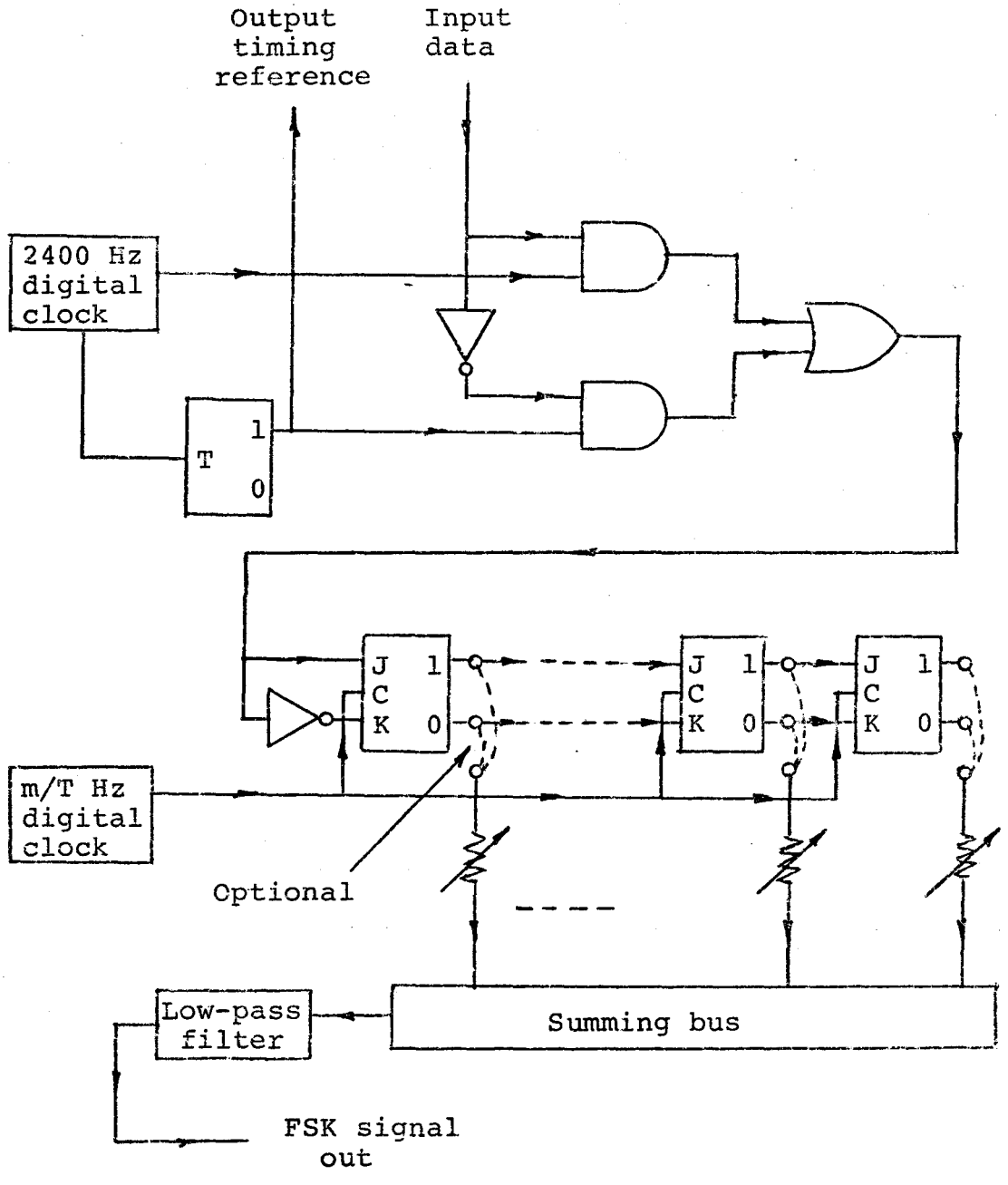


fig.5.1 Modulator

is +1, or 0 for this analysis

$q_1(t)$ is the 2400 Hz square wave.

$q_2(t)$ is the 1200 Hz square wave.

$m_d(t)$ is the output of the digital modulator

It is apparent from equation(5.1)that the modulation process will generate many harmonics. The use of a square wave carrier will give rise to sidebands of the carrier harmonics which are unlimited in bandwidth because of the square wave modulation. However, it will be shown that because the carrier stands in particular relation to the bit rate that this so-called aliasing error has the character of linear distortion and thus can be corrected by linear means.

5-3 Pulse Response of the Synchronous Analog FSK Modulator

In this section the pulse response of the analog FSK modulator is considered. The switching, which in this case occurs between two sinusoidal oscillators, is again assumed to be synchronous with the carrier as in the previously described digital case. If we consider a pulse of ω_1(again of length T), then as indicated in figure (5.2), the output of the modulator can be decomposed into three terms.

$$m_a(t) = \sin(2\omega_1 t) - \sin(2\omega_1 t)\cdot p_T(t) - \sin(\omega_1 t)\cdot p_T(t) \tag{5.2}$$

where

$$p_T(t) = 1 \qquad -T/2 \leq t \leq T/2$$

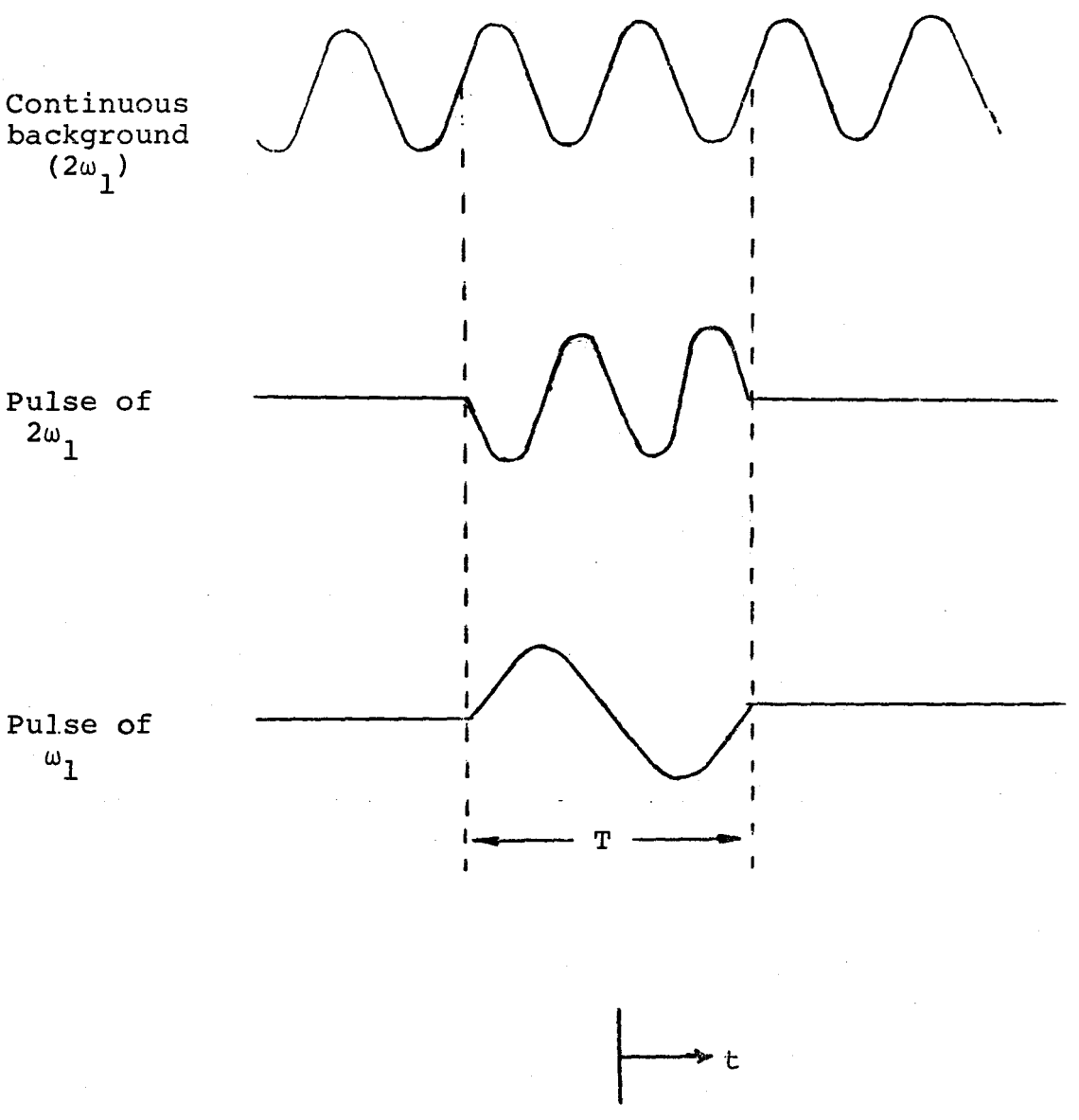


fig.5.2 Decompostion of a Binary FSK Wave for a Pulse of ω_1

$$= 0 \qquad \text{elsewhere} \tag{5.3}$$

The Fourier transform of $m_a(t)$ is

$$M_a(\omega) = -j\pi\delta(\omega-2\omega_1) + j\left\{\frac{\sin(\omega-2\omega_1)\cdot T/2}{\omega-2\omega_1} + \frac{\sin(\omega-\omega_1)T/2}{\omega-\omega_1}\right\}$$

(positive frequency contribution shown)

(5.4)

Since $\omega_1 = 2\pi/T$ and $\omega_2 = 4\pi/T$ (5.5)

then equation (5.4) reduces to

$$M_a(\omega) = -j\pi\delta(\omega-2\omega_1) + j\sin(\omega T/2)\left\{\frac{1}{\omega-2\omega_1} + \frac{1}{\omega-\omega_1}\right\} \tag{5.6}$$

The first term here is, of course, a spectral line at $2\omega_1$ which represents the background tone. The second and third terms are the spectra of interest - the so-called envelope spectrum, which does not change, regardless of whether a one or zero is keyed.

5-4 Envelope Spectrum of the Digital Modulator

Using a Fourier series representation of the square wave carriers, then from equation (5.1),

$$m_d(t) = (a_k A + a_k A) + \frac{4Aa_k}{\pi}\left(\sin(\omega_1 t) + \frac{\sin(3\omega_1 t)}{3} + \ldots\right) + \frac{4Aa_k}{\pi}\left(\sin(2\omega_1 t) + \frac{\sin(6\omega_1 t)}{3} + \ldots\right) \tag{5.7}$$

where: 2A is the amplitude of the carriers

a_k is still interpreted in the sense

$a_k = 1$ if the input digit is +1

$a_k = 0$ if the input digit is -1

The first term in equation (5.7) represents a dc component

of A, and can be removed by a simple R-C high-pass filter and thus is of no importance.

The pulse response (ignoring the gain constant and dc component) can be written

$$M_d(\omega) = -j\pi \sum_n \frac{1}{n}\delta(\omega-2n\omega_1) + j\sin(\omega T/2) \sum_n \frac{1}{n}\left(\frac{1}{\omega-2n\omega_1} + \frac{1}{\omega-n\omega_1}\right)$$

$$n = 1, 3, 5, \ldots\ldots$$

(positive frequency contribution shown) (5.8)

It is apparent from equation (5.8), that the discrete spectral lines will lie well outside the effective passband of most channels (excluding the two at 1200 and 2400 Hz). In addition, in the vicinity of the carriers, (the 1200 and 2400 Hz tones), both the envelope spectra (i.e., the digital and analog spectra) are continuous. Thus, in effect, only the shape of the envelope spectrum of the digital modulator has been affected in this area, or in other words, there exists some correction filter, such that in the vicinity of the two main carriers

$$M_a(\omega) = C_c(\omega) \cdot M_d(\omega) \qquad (5.9)$$

where $C_c(\omega)$ is the correction filter. Thus, the aliasing error takes on the character of linear distortion and can be corrected by linear means. Van Gerwin, and van der Wolf show that this is true for the general case, if the carrier is some multiple of half the bit rate.[2] Furthermore, they describe a simple method which can be employed to determine $C_c(\omega)$. However, since in this case it was decided to

use the transmitting filter to remove any residual distortion in the overall system, little would be gained from the knowledge of its characteristics.

5-5 Post Modulation Filtering Requirements

From equation (5.8) it is apparent that the output of the digital modulator will contain odd harmonics of the carrier frequencies. Since signal levels in the FDM system are set prior to filtering and modulation, these harmonics will reduce the effective signal to noise raio at the receiving end. Furthermore, cross-modualation requirements usually dictate that the harmonic content of all signals be low outside the 3 kHz band. Thus, in order to comply with these requirements, the digitally modulated signal must be filtered. The restriction here is, of course, that the post-modulation filtering must not introduce significant inter-symbol interference.

The structure selected for the post-modulation filter was that of a binary transversal filter followed by a two section R-C low-pass network.

5-6 The Binary Transversal Filter

In chapter three, we found that the transversal filter proved to be a particularly useful structure for minimizing ISI, however, because of its nature, it is also particularly well-suited to performing filtering on digital signals.

Recalling the definition of the transversal filter,

$$w(t) = \sum_{n=-M}^{M} c_n \cdot d(t-nT) \qquad (5.10)$$

where w(t) is the output of the TF; d(t-nT) represents a delayed version of the input signal d(t); and c_n is the coefficient of the n^{th} tap (which again may be positive or negative). The delays can be obtained by the use of analog (LCR) delay networks - however, in this case we can take advantage of the fact that the digitally modulated signal is synchronous. That is, transitions of the digital FSK signal can take place only at the transitions of the 2400 Hz clock. Thus, the digital FSK waveform can be delayed without distortion by means of a synchronous shift register. Furthermore, the tap gains can be derived by a simple resistive network, as indicated in figure (5.1).

Because of the periodicity ($2\pi m/T$) of the BTF's frequency response, the proper choice of the delay time (i.e., the shift frequency) becomes a critical factor. We could, of course, select the delay time to be T, however, when the shift frequency equals the bit rate of the data signal, the period of the transfer function becomes too small to achieve the desired reject bands (indicated in figure (5.3). Furthermore, the requirement on the analog filter following the BTF becomes very stringent, if this condition exists. Therefore, the shift frequency is always chosen to be some multiple of the bit rate. The higher the

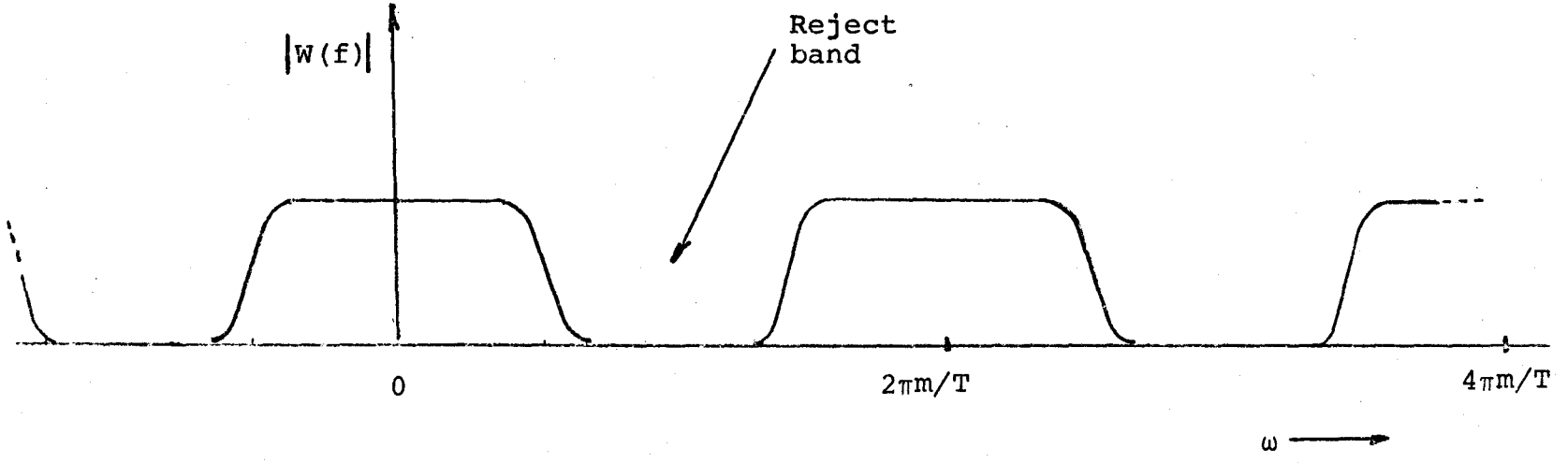


fig.5.3 Amplitude Characteristic of BTF

shift frequency, the wider the reject band becomes, however, correspondingly, the longer (and consequently more expensive) the shift register becomes in order to achieve the same accuracy in the transfer function. If the clock frequency is sufficiently high, one capacitor is all that is required to remove the unwanted sidebands.

Since dual-packaged integrated flip-flops were employed, it was felt that the selection of the shift frequency of 8/T Hz (m=8), a 10 tap BTF, and a two stage low-pass filter (3 dB point at approximately 2.7 kHz), represented the most economical solution to the design of the post-modulation filter. As previously indicated, the tap gains were left as variables, in order to achieve flexibility in the overall implementation since residual ISI could be removed simply by adjusting the taps on the transmitting BTF.

CHAPTER SIX

THE PHASELOCK LOOP AND ZERO-CROSSING DETECTOR DISCRIMINATORS

6.1 Introduction

In this chapter, the two discriminators employed in the receiver structure are discussed. The PLL discriminator is described in detail, since its characteristics determine the performance of the receiver directly.

6-2 The Phaselock Loop - Basic Description

As indicated in figure (6.1), the PLL consists of three basic components,

(a) a phase detector (PD)

(b) a low-pass filter (L)

(c) a voltage-controlled oscillator (VCO)

The phase detector compares the phase of the input signal with the phase of the VCO. The output of the PD is a measure of the phase difference between its two inputs - we will for the moment assume that the relationship is linear although in most practical cases, it is not. Thus

$$v_{pd} = K_{pd}(\theta_i - \theta_{vco}) \qquad (6.1)$$

where θ_i is the phase of the input signal

θ_{vco} is the phase of the VCO

v_{pd} is the output of the PD

K_{pd} is the phase detector gain constant

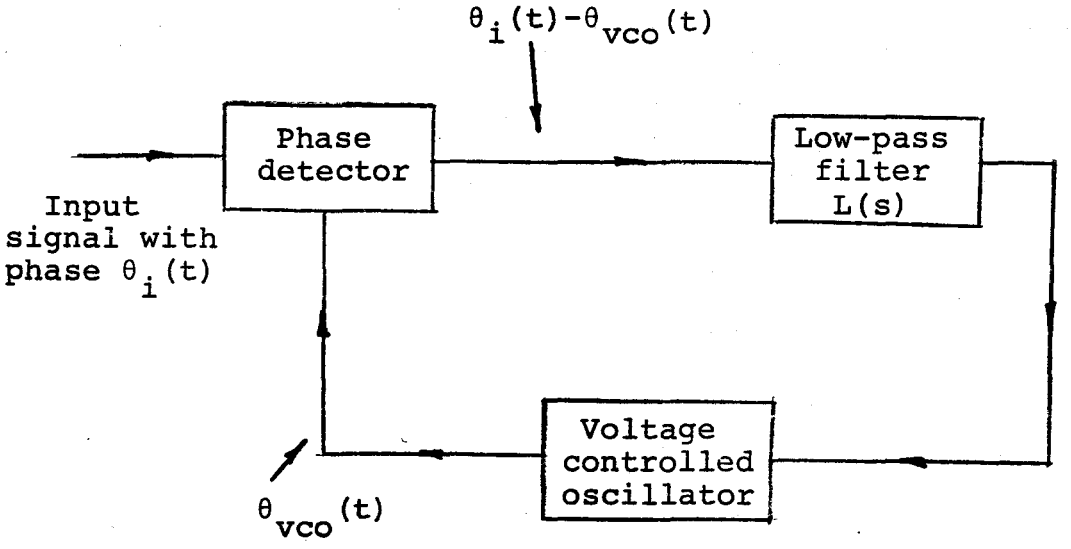


fig.6.1 Phaselock Loop - Linear Model

having the dimensions of volts/radian

The voltage v_{pd} is then filtered by the low-pass filter, the characteristics of which determine both the static and dynamic performances of the PLL. The filtered error voltage is then applied to the VCO in such a way that the instantaneous deviation in frequency from the idle frequency is given by,

$$\Delta\omega = K_{vco} \cdot v_{fpd} \tag{6.2}$$

where K_{vco} is the VCO gain constant in radians/second/volt. Equation (6.2) may be rewritten

$$\frac{d\theta}{dt}_{vco} = K_{vco} \cdot v_{fpd} \tag{6.3}$$

and using Laplace transforms

$$\theta_{vco}(s) = \frac{1}{s} \cdot (K_{vco} \cdot V_{fpd}(s)) \tag{6.4}$$

Thus, the phase of the VCO output is proportional to the integral of the input control voltage.

The remaining loop equations can now be written

$$V_{pd}(s) = K_{pd}(\theta_i(s) - \theta_{vco}(s)) = K_{pd}\theta_e(s) \tag{6.5}$$

$$V_{fpd}(s) = L(s)V_{pd}(s) \tag{6.6}$$

with the resulting transfer functions

$$\frac{\theta_{vco}(s)}{\theta_i(s)} = \frac{K_{vco}K_{pd}L(s)}{s + K_{vco}K_{pd}L(s)} \tag{6.7}$$

$$\frac{\theta_e(s)}{\theta_i(s)} = \frac{s}{s + K_{vco}K_{pd}L(s)} \tag{6.8}$$

6-3 Selection of the Loop Filter L(s)

In practice, high order loop filters are not used

in PLL applications because:

(a) The increase in performance gained when high order filters are employed, is usually not significant and consequently, the increased complexity is not warranted.

(b) The introduction of a high order filter often creates stability problems.

The most widely used loop filters are, in fact, of the first order. Of this class, the lead-lag filter is perhaps the most commonly used, since in most PLL applications, the loop bandwidth required is usually only a small fraction of the VCO's center frequency, and consequently in order to maintain control over all parameters of the PLL, and to improve the transient response, this type of filter must be employed. However, since in this application, the bandwidth requirement is large (approximately 66% of the VCO center frequency), the simple lag filter is all that is required. Thus, in the following development we will restrict ourselves to this type of filter.

For the lag filter, L(s) is given by

$$L(s) = \frac{1}{s\tau + 1} \tag{6.9}$$

and consequently, from equation (6.7), the overall transfer function is

$$\frac{\theta_{vco}(s)}{\theta_i(s)} = \frac{K_{vco}K_{pd}/\tau}{s^2 + s/\tau + K_{vco}/\tau} \tag{6.10}$$

$$= \frac{\omega_n^2}{s^2 + 2s\zeta\omega_n + \omega_n^2} \qquad (6.11)$$

where $$\omega_n = (K_{vco} \cdot K_{pd} / \tau)^{.5} \qquad (6.12)$$

$$\zeta = .5(K_{vco} \cdot K_{pd} \cdot \tau)^{-.5} \qquad (6.13)$$

It is apparent from equation (6.11) that the three loop parameters (the overall gain $K_{vco} \cdot K_{pd}$, ω_n, and ζ) cannot be chosen independently - however, since as previously mentioned, the bandwidth requirement is large, a compromise should be attainable to ensure good transient response and tracking characteristics. However, if it were necessary to have large gain and small bandwidth it is obvious, that the loop would be badly underdamped and transient response would be poor.

6-4 The PLL as a Discriminator

Phaselock loops are widely used as frequency discriminators for analog transmission because of their improved noise threshold over conventional discriminators.[1] However, in this application, the PLL is used primarily for ease of instrumenation (PLL are available in integrated circuit form). The PLL also offers somewhat better performance than simple discriminators such as the zero-crossing detector type.

In order to indicate how the PLL can be used as a discriminator, consider the filtered error voltage, which from equations (6.5) and (6.6) can be written

$$V_{fpd}(s) = K_{pd} \cdot L(s) \cdot \Theta_e(s) \qquad (6.14)$$

Using the above equation, and equations (7.8) and (7.9)

$$V_{fpd}(s) = (1/K_{vco}) \cdot s\theta_i(s) \cdot \frac{\omega_n}{s^2 + 2s\zeta\omega_n + \omega_n^2} \qquad (6.15)$$

Examining equation (6.15); the first term is simply a gain constant, the second term represents the discriminated frequency modulation of the input signal, and the third term is of form of a second order low-pass filter. It should be noted at this point that in practice if a poor choice of this latter transfer function is made, effective correction can not be achieved by introducing another filter at a later stage, since in most realizations of PLLs, the phase detector characteristic is not linear.

6-5 Tracking and Acquisition Characteristics

It has been tacitly assumed up to this point that the phase detector characteristic is linear, however, as just indicated, this assumption is not in general true. Practical detectors are usually of the multiplier or chopper type, which have sinusoidal transfer functions, that is, of form,

$$v_{pd} = K_{pd} \cdot \sin(\theta_i - \theta_{vco}) \qquad (6.16)$$

The introduction of this non-linearity into the loop has many reprecussions. Perhaps the most serious effect is that of constraining the static and dynamic performances.

Since the sine function cannot exceed unity in magnitude, the loop will lose lock if the static phase error should exceed $\pm \pi/2$ radians. Beyond this point, the VCO

will begin to move in the wrong direction. This effect in fact limits the static tracking range ("hold-in range") to[1]

$$\Delta\omega_h = K_{vco} \cdot K_{pd} \tag{6.17}$$

That is, the maximum difference in frequency, from the rest frequency, (which in this case is in the center of the spectrum of the transmitted signal - 1800 Hz) for which the loop can maintain lock, is given by the dc loop gain.

The "pull-in" range is defined as the maximum difference in frequency from the rest frequency, for which the loop can achieve lock in the steady state sense. For the lag filter, the pull-in range can be shown to be given approximately by,[2]

$$\Delta\omega_p \simeq K_{vco} \cdot K_{pd} \cdot |L(j\Delta\omega_p)| \tag{6.18}$$

$$\simeq \omega_n \tag{6.19}$$

Since the magnitude of the low-pass filter response is always less than unity, it is apparent from equations (6.17) and (6.18) that the capture range will always be smaller than the hold-in range.

If the frequency difference between the input and the VCO is less than the loop bandwidth, the PLL will lock up almost instantaneously without skipping.[1] However, there are some loop types in which the VCO frequency "walks" slowly towards the input frequency even though the input frequency is much larger than the loop bandwidth. Since in this application the loop must regain lock quickly once it is lost, otherwise a long string of errors could occur,

we must then ensure that the loop can lock quickly on to either of the two transmitted frequencies, and thus,

$$2\Delta\omega_p \geq 2\pi \cdot 1200 \tag{6.20}$$

In order to speed up the acquisition process, the loop bandwidth can be increased. This will also increase the capture range and has the serious disadvantage of increasing the amount of noise admitted to the system.

Before leaving this section, it should be pointed out that, if the loop is near lock, it will behave in accordance with the previously developed linear analysis, since for small phase error angles, the sinusoidal transfer function of the phase detector is approximately linear.

6-6 Noise Bandwidth

In the absence of an input signal, it would be desirable to minimize the noise received at the output of the low-pass filter- or equivalently the phase jitter of the VCO. If we again assume a linear phase detector characteristic, then from equation (7.11), the normalized noise spectral density of the VCO's output (assuming a white noise input) can be written,

$$N_{vco}(\omega) = \left| \frac{\omega_n^2}{(j\omega)^2 + 2j\omega\zeta\omega_n + \omega_n^2} \right| \tag{6.21}$$

It follows that the effective noise bandwidth is given by

$$B_{eff} = \int_0^\infty N_{vco}(\omega)\, d\omega \tag{6.22}$$

which reduces to[3]

$$B_{eff} = 2\pi \cdot \frac{\omega_n}{8\zeta} \tag{6.23}$$

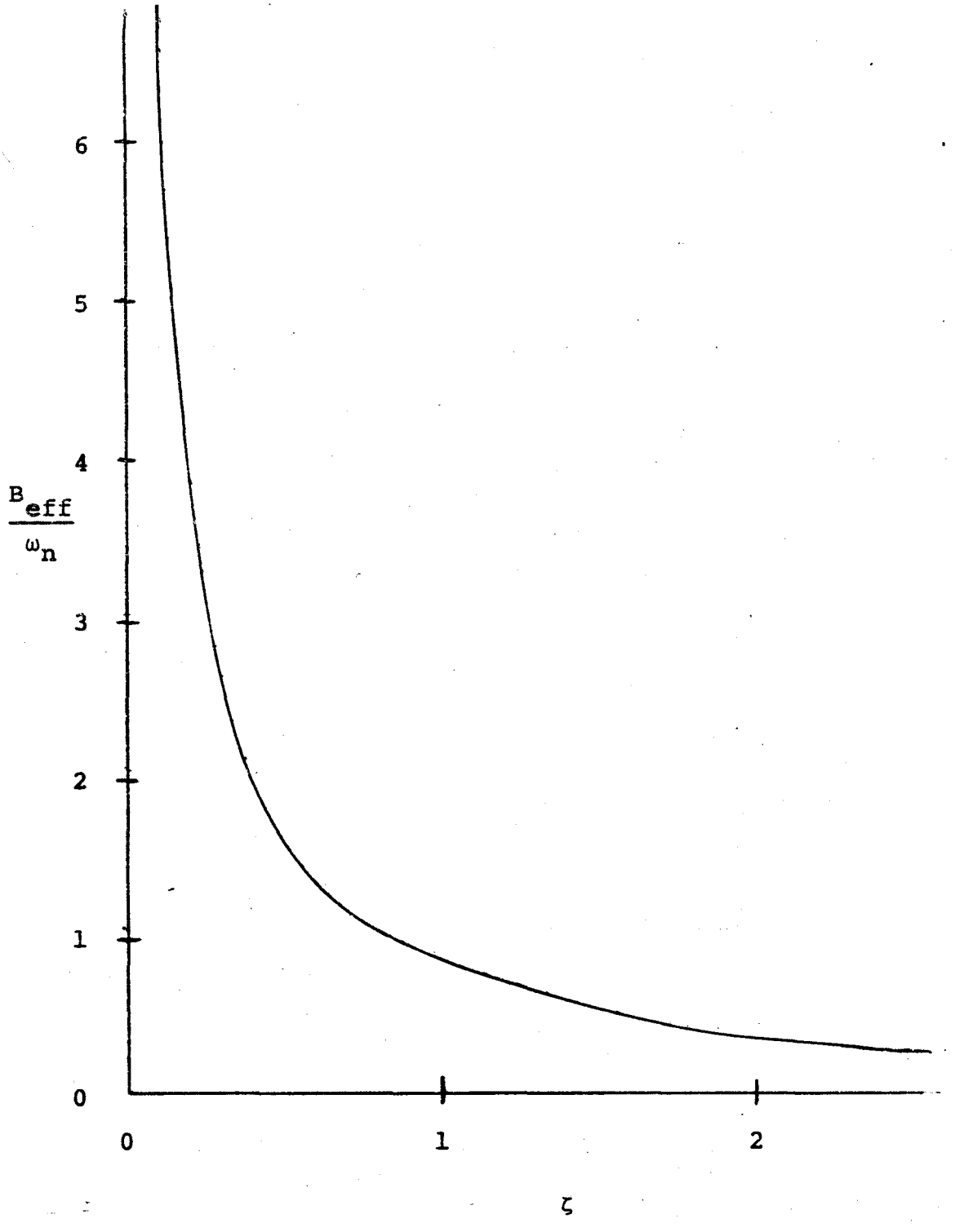


fig.6.2 Normalized Loop-Noise Bandwidth as a Function of Damping for a Lag Loop Filter

It is apparent from figure (6.2) (B_{eff}/ω_n as a function of ζ), that as ζ decreases below .5, that the phase jitter of the VCO, or equivalently, the noise observed at the output of the low-pass filter becomes excessive.

6-7 Optimization of the Loop Parameters

As in most PLL applications, the problem of optimizing the performance of the loop in the presence of noise involves two contradictory factors. Ideally, it would be desirable to make the loop bandwidth as large as possible to obtain good tracking and acquisition characteristics, while on the other hand, we would like to make the loop bandwidth as small as possible, in order to minimize the effects of noise.

One obvious compromise is to make the loop bandwidth only as large as needed to accomodate the input signal, that is design the loop so that it can quickly achieve lock, and maintain lock, when the input frequency is either 1200 or 2400 Hz. This, of course represents the absolute minimum bandwidth. For this, we require

$$\begin{aligned} \omega_n &\simeq B_{eff} \\ &\simeq \Delta\omega_p \end{aligned} \geq 2\pi \cdot 600 \text{ rads} \qquad (6.24)$$

and furthermore ensure that (equation (6.23)

$$\zeta \simeq .75 \qquad (6.25)$$

These values are only nominal, because as will be discussed later, the pre- and post-detection processing plays an important part in the determination of the above

parameters. The only valid criterion which can be used in the optimization of these parameters is, of course, the bit error rate.

6-8 Mechanisms of Error Production

The problem of determinimg the performance of FSK systems which employ discrimination detection is not, in general, mathematically tractable - however, Mazo and Salz[4] have analysed the binary FSK case, using a simple model which yields a relatively accurate insight into the basic mechanisms of error production for these types of systems.

In their model:

(a) They assume that the overall system is wideband so that distortionless transmission can be considered, that is, the received waveform in the absence of noise can be described as,

$$v(t) = A\cos(\omega_c t + a_k \omega_d t + q(a)) \qquad (6.26)$$

where: $v(t)$ is the received signal during the k^{th} signalling interval

$q(a)$ is a constant whose value depends on past symbols

ω_c is the carrier frequency

ω_d is the carrier deviation

a_k is the symbol received in the prescribed interval having a value of ± 1

(b) The discriminator operates on v(t) plus a zero-

mean Gaussian noise process n(t), whose spectrum is assumed symetrical about ω_c.

(c) The post-detection filter was assumed to be of the integrate and dump type.

Using this model, they then showed that the sampled output of the post-modulation filter can be represented as

$$I_d = (\omega_c + a_k\omega_d)\cdot T + \eta + \mu \qquad (6.27)$$

where the first term represents the demodulated signal, the second a "click" noise component and the third a non-click noise component.

The click noise component is a discrete random variable which can only take on the values $\pm 2n\pi$, n = 0,1,.... It can be shown that the aiding clicks can be ignored and that the opposing clicks have a Poisson distribution. The probability of obtaining M clicks in T seconds is given by

$$P(M_T) \simeq \frac{\exp(-R_-T)\cdot(R_-\cdot T)^M}{M!} \qquad (6.28)$$

where R_- is the rate of occurance of the opposing clicks and is given by

$$R_- \simeq (\omega_d/2\pi)\cdot\exp(-S/N) \qquad (6.29)$$

where S/N is the signal to noise ratio.

The net effect of the click process is to pull the output of the integrate and dump filter towards ω_c/T, regardless of whether a 1 or -1 was transmitted.

The non-click component of equation (6.27) is a continuous and bounded ($|\mu| \leq \pi$) random process described by

$$p(\mu) = p(\nu) * p(-\nu) \tag{6.30}$$

where

$$p(\nu) = \sqrt{(S/\pi N)} \cdot \cos(\nu) \cdot \exp(-S/N) \cdot \sin^2(\nu) \cdot \mathrm{erf}[\sqrt{(S/N)} \cdot \cos(\nu)] + \frac{\exp(-S/N)}{\pi} \tag{6.31}$$

and the asterisk denotes convolution.

A qualitative sketch of the distribution of the total noise χ, (where $\chi = \mu + \eta$) is illustrated in figure (6.3). Notice that the probability density consists of a central lobe about zero extending to $\pm\pi$ on each side (the form is given by equation (6.30), plus identically shaped lobes displaced by integral multiples of 2π in the negative direction. These displaced lobes are weighted by the probabilty of obtaining the appropriate number of clicks in the interval T.

6-9 Probability of Error

For our system, the modulation parameters are

$$\begin{aligned} \omega_c &= 2\pi \cdot 1800 \\ \omega_d &= 2\pi \cdot 600 \end{aligned} \tag{6.32}$$

Consequently, the noise-free output of the discriminator is, from equation (6.27)

$$I_d = a_k \pi + 3\pi \tag{6.33}$$

When a_k=1, a 2400 Hz sinusoid is received, and from equation (6.33), the output of the integrate and dump filter would be 4π. However, if noise is present and $\chi \leq \pi$, then an error is made. Thus, the probability of error is given by

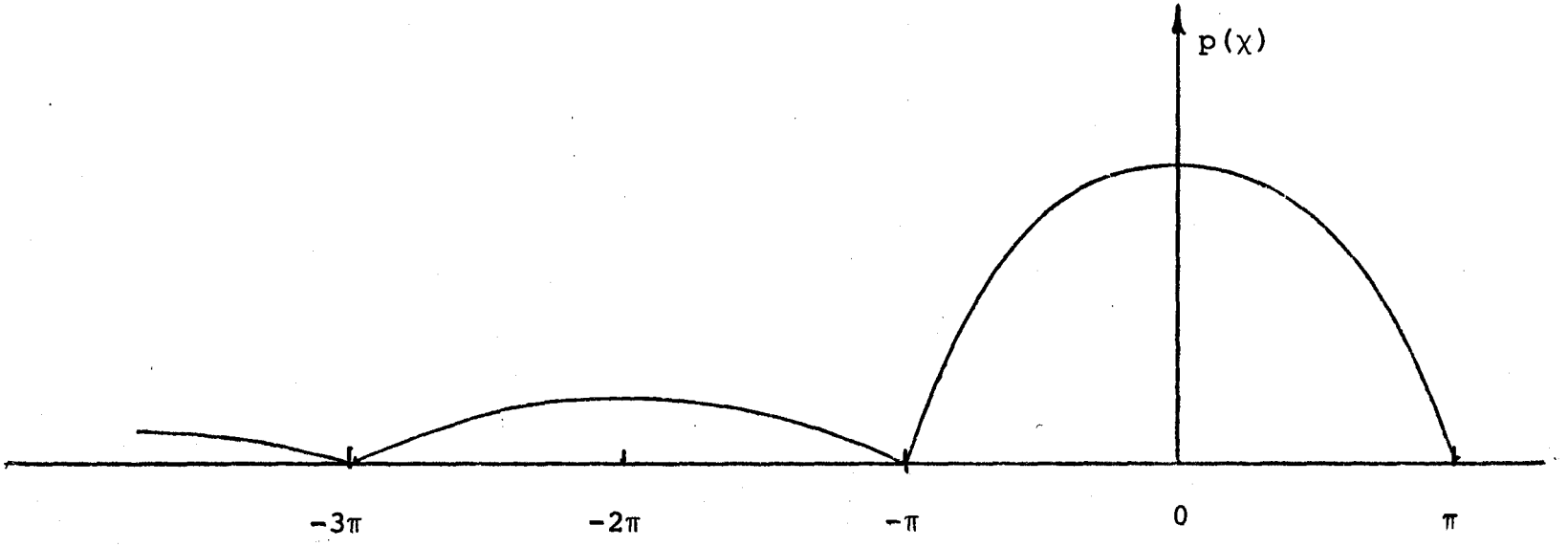


fig.6.3 Qualitative Sketch of $p(\chi)$ for $a_k = +1$

$$P_e^1 = P(\chi < -\pi) \tag{6.34}$$

or equivalently,

$$P_e^1 = 1 - P(-\pi < \chi < \pi)$$

$$= 1 - \Pr(0 \text{ clicks in the interval } T) \tag{6.35}$$

From equation (6.28), the probability of obtaining 0 clicks in the interval T is

$$P(0_T) \approx \exp(-R_- \cdot T) \tag{6.36}$$

which from equation (6.29) gives

$$P(0_T) \approx \exp\left(-\frac{\exp(-S/N)}{2}\right) \tag{6.37}$$

If the signal to noise ratio is large ($S/N \gg 1$), then equation equation (6.37) reduces approximately to

$$P(0_T) \approx 1 - \frac{\exp(-S/N)}{2} \tag{6.38}$$

Consequently from equation (6.35)

$$P_e^1 \approx \frac{\exp(-S/N)}{2} \tag{6.39}$$

We have, of course, assumed in the preceding derivation assumed that the digit +1 was transmitted. The transmission of a_k=-1 would give rise to the same expression for the bit probability of error, assuming that the decision threshold still remains set at 3π, thus by symetry

$$P_e = P_e^1 = P_e^0 \approx \frac{\exp(-S/N)}{2} \tag{6.40}$$

6-10 Comparison with the Optimum Incoherent Orthogonal Receiver

In the preceding derivation, the two transmitted

signals were, in fact, orthogonal, that is,

$$\int_0^T s_1(t)\cdot s_2(t)dt = 0 \tag{6.41}$$

where:

$$\begin{aligned} s_1(t) &= A\cos(\omega_1 t) \\ s_2(t) &= A\cos(2\omega_1) \end{aligned} \tag{6.42}$$

The probability of error for the optimum incoherent orthogonal receiver is given by,[5]

$$P_e = (1/2)\cdot\exp[-(S/N)'/2] \tag{6.43}$$

Superficially, it would appear that discrimination detection is superior (equation (6.40)), however, it must be remembered that the signal to noise ratio of the optimum incoherent receiver is given by,

$$(S/N)' = A^2/2\sigma_n^2 \tag{6.44}$$

where

$$\sigma_n^2 = \gamma\cdot(1/T) \tag{6.45}$$

and γ is the noise power spectral density (one-sided). It is apparent from the above equations that the effective bandwidth of the optimum receiver is 1/T. However, for the discriminator, Mazo and Salz assumed that the input bandwidth of the pre-discrimination bandpass filter was large enough to ensure that the signal was not distorted in the discrimination process. The result is, that the input bandwidth of the discriminator will be much larger than 1/T, consequently the discriminator will admit too much noise, producing clicks which cannot be removed by the post-detection

filter. As a result, the performance of the discriminator is inferior to that of the optimum receiver - as it must be.

It should be pointed out that had we selected a smaller frequency deviation, the click process would become less important as a mechanism of error production. This is evident from figure (6.3); for example, if $\omega_d T = \pi/2$, (a narrow band system), it is apparent that the continuous component of χ, rather than the click component, will play the more important role in determining the error rate.

6-11 Performance of the Discriminator in the Dispersive Channel

The preceding analysis of Mazo and Salz has been somewhat unrealistic, in that real channels are band-limiting, furthermore phase distortion is usually present. These latter effects complicate the detection process in two ways.

Firstly, the envelope of the carrier is no longer constant. Instead, it becomes a function of time (A(t)). As a result, the effective signal power available at the input to the discriminator in the k^{th} signalling interval changes from $A^2/2$, to

$$P_k = (1/2T) \cdot \int_0^T A_k^2(t)\,dt = \iota_k \cdot (A^2/2) \qquad (6.46)$$

where ι_k will vary from interval to interval.

The second effect is to alter the output of the integrate and dump filter such that it is no longer $a_k \omega_d T$, but rather $\delta_k a_k \omega_d T$, where the bias term of ω_c has been

ignored, and λ_k varies from interval to interval.

Following the method of Bobilin and Lindenlaub[6], we use the previously described analysis for the non-distorting wideband case to write the probability of error in the k^{th} interval as

$$P_e(k) = \Pr(\text{number of clicks is one or greater}) + \Pr(\text{non-click component is less than } -\lambda_k) \cdot \Pr(\text{no clicks occur in this interval}) \quad (6.47)$$

where again we have assumed that $a_k=1$ is received, and

$$\lambda_k = \delta_k a_k \omega_d T \quad (6.48)$$

more explicitly,

$$P_e(k) = P(M_T>1 \mid \lambda_k, P_k) + P(\mu<-\lambda_k \mid P_k) \cdot P(M_T=0 \mid \lambda_k, P_k) \quad (6.49)$$

From equation (6.49), it is apparent that the effect of the ISI is to create a variable pedestal for the noise. The effective signal to noise ratio at the input, as well as the demodulated output level will vary from interval to interval.

In the wideband case, the non-click component had no effect on the error rate, however, because of this varying pedestal effect, it now enters into the error mechanism. It should be expected from equation (6.49) however, that since it was found in the distortionless case that the dominant error mechanism was that of clicks, that this component will be even more emphasized when distortion is present and thus that it will still play the dominant role.

The usefullness of equation (6.49) is somewhat limited, in that, in order to estimate the average probability of error for a given channel, $P_e(k)$ must be evaluated for many long sequences of pseudo-random data.

6-12 Pre-Detection Filtering

Bennett and Davey[7] have derived asymtotic expressions for the error rate of the conventional discriminator for the general case when ISI is present. However, the expressions derived have the same limitation as that of equation (6.49). If however, one assumes that no ISI is present, the expressions become useful for determining the optimum bandwidth of the pre-detection filter.[8] The equations (after reduction) are

$$P_e \sim \begin{cases} \exp(-c^2) & a > 1 \\ \exp(-a^2c^2) & a \leq 1 \end{cases} \qquad c^2 >> 1 \quad (6.50)$$

where the parameter c^2 can be interpreted physically as the signal to noise ratio (within a proportionality constant), and the parameter a can be interpreted physically as the bandwidth ratio (the ratio of the peak frequency deviation to the noise bandwidth of the input bandpass filter).

It is apparent from equation (6.50), that as the parameter a becomes less than unity the probability of error increases - however, increasing a beyond unity will produce no significant improvement in performance. Physically this is understandable, since if a is less than unity it means that the front-end filter is broader than necessary to ac-

comodate the received signal, and consequently, excess noise is admitted to the detector (in this case the PLL) with a resulting increase in the error rate. As a is increased beyond unity, the effective noise power delivered to the detector is reduced, however, the signal power delivered to the detector will also be reduced, thus accounting for the lack of improvement beyond unity. It should be noted that this latter result is somewhat misleading in that as a becomes very large, ISI must occur because of the severe bandlimiting and consequently the assumtion of no ISI becomes invalidated.

In summary, one may conclude that for a binary FSK system, the front-end filter should have no wider a bandwidth than is required to accomodate the transmitted signal. For the modulation parameters of our system (equation 6.32), this would require that the bandwidth selected should be in the order of $2\pi \cdot 1200$ radians.

The introduction of this pre-detection filter will affect the earlier selection of the loop parameters somewhat, since the reduction of the noise bandwidth at the input of the PLL will allow for an increase in the loop bandwidth in order to enhance the tracking and acquisition characteristics, without sacrificing noise immunity.

6-13 Pre-Detection Processing

As can be observed from figure (4.1), the PLL is effectively preceded by a sampled bandpass limiter. The

sampling rate was set at 100/T, as it was conjectured (and later verified experimentally) that this high rate of sampling would not introduce any significant degradation into the system, and furthermore that the effects of sampling could be ignored.

Ordinary (non-sampled) bandpass limiters are often employed in tandem with a PLL in practice, primarily for two reasons; firstly because the power delivered to the PLL becomes relatively constant with the consequence that the PLL becomes input level insensitive, and secondly because the addition of the bandpass limiter causes the system to adapt to varying signal to noise ratio conditions.

This latter effect occurs because the effective signal voltage delivered to the PD of the PLL is reduced as the noise increases according to the "limiter signal suppression factor" ξ , where ξ is given by[1]

$$\xi = \left(\frac{S/N}{4/\pi + S/N}\right) 1/2 \qquad (6.51)$$

This effect is advantageous in that although we have assumed thus far that the phase detector is ideal (i.e. independent of input level), this is in general not true of the practical phase detectors (multipliers, choppers etc.) employed in practice. In fact, in most of these detectors the output of the PD is directly proportional to the level of the input signal. Thus in all the preceding material, whenever the loop gain $K_{pd}K_{vco}$ appears, it must be multiplied

by ξ if a limiter is added. With this suppression factor present, it is quite easy to see why the limiter is a desirable component of the receiver structure. For example, as the SNR decreases, then for a loop with a lag-type filter, the natural frequency ω_n decreases and the damping ζ increases - both of which will aid the performance of the PLL as the noise power increases. For the analog case, the limiter enhances the signal to noise ratio by 3 dB, while for very low signal to noise ratios, the degradation is only in the order of 1 dB.[1]

6-14 Post-Detection Filtering

A simple Sallen-Key low-pass filter was employed at the output of the loop filter in order to suppress higher order harmonics, further reduce the noise bandwidth and aid in the overall Nyquist shaping.

6-15 The Zero-Crossing Detector Discriminator

The zero-crossing detector discriminator (ZCDD) is quite widely used in commercial modems because of its ease of implementation. The idea behind the discriminator is that a measure of the instantaneous frequency can be obtained by counting the number of zero-crossings per unit time. The method is simple, economical and particularly attractive when the frequency shift is large in comparison with the bit rate, since then the output becomes almost directly proportional to the instantaneous frequency. Filtering consider-

ations limit the usefulness of the ZCDD to systems in which the bit rate is less than, or equal to the lowest shift frequency of the FSK signal.[7]

The basic configuration is illustrated in figure (6.4). The input signal is limited, differentiated, rectified, and a pulse of l_z and amplitude h_z generated for every zero-crossing. The pulse train is then low-pass filtered, with the resultant output being an approximation to the instantaneous frequency.

6-16 Approximate Analysis of the ZCDD

If $\omega_c t+\theta_m(t)$ represents the instantaneous phase of the received FSK waveform, then $A\cos(\omega_c t+\theta_m(t))$, the received waveform goes to zero for

$$\omega_c t + \theta_m(t) = 0,\pi,2\pi,... \qquad (6.52)$$

Consequently, a pulse will be generated at times t_n, such that

$$\omega_c t_n + \theta_m(t_n) = m\pi \qquad (6.53)$$

where $m = 1,2,3,...$

The $(n+1)^{th}$ pulse occurs for

$$\omega_c(t_{n+1}-t_n) - \theta_m(t_{n+1}) = (m+1)\pi \qquad (6.54)$$

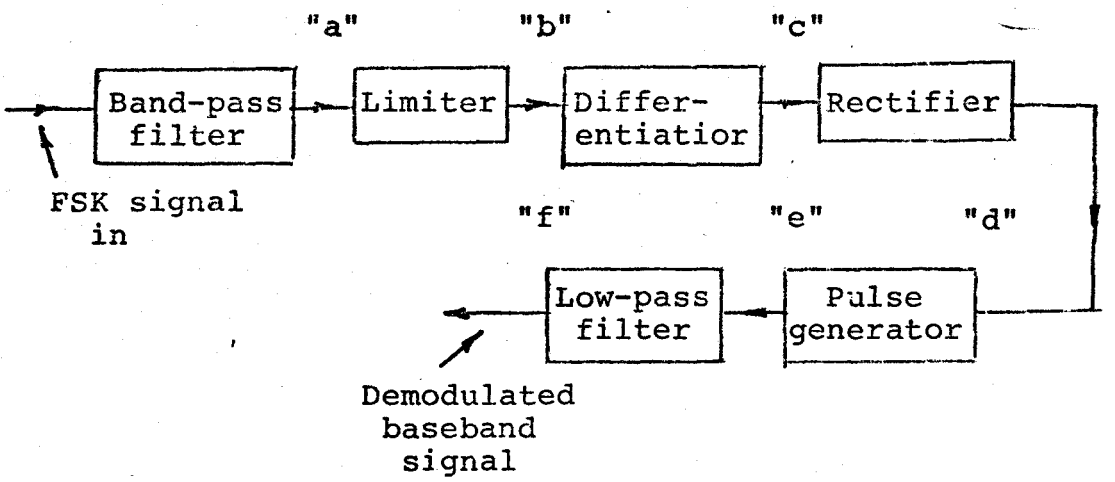
The phase difference in the interval between t_n and t_{n+1} will be π radians, and thus from equations (6.53) and (6.54)

$$\omega_c(t_{n+1}-t_n) + \theta_m(t_{n+1}) - \theta_m(t_n) = \pi \qquad (6.55)$$

Expanding $\theta_m(t_{n+1})-\theta_m(t_n)$ in a Taylor's series and assuming that $\theta_m(t_{n+1})-\theta_m(t_n)$ is small, then

$$\theta_m(t_{n+1}) \simeq \theta_m(t_n) + \dot{\theta}_m(t_n)\cdot(t_n-t_{n+1}) \qquad (6.56)$$

which from equation (6.55) gives



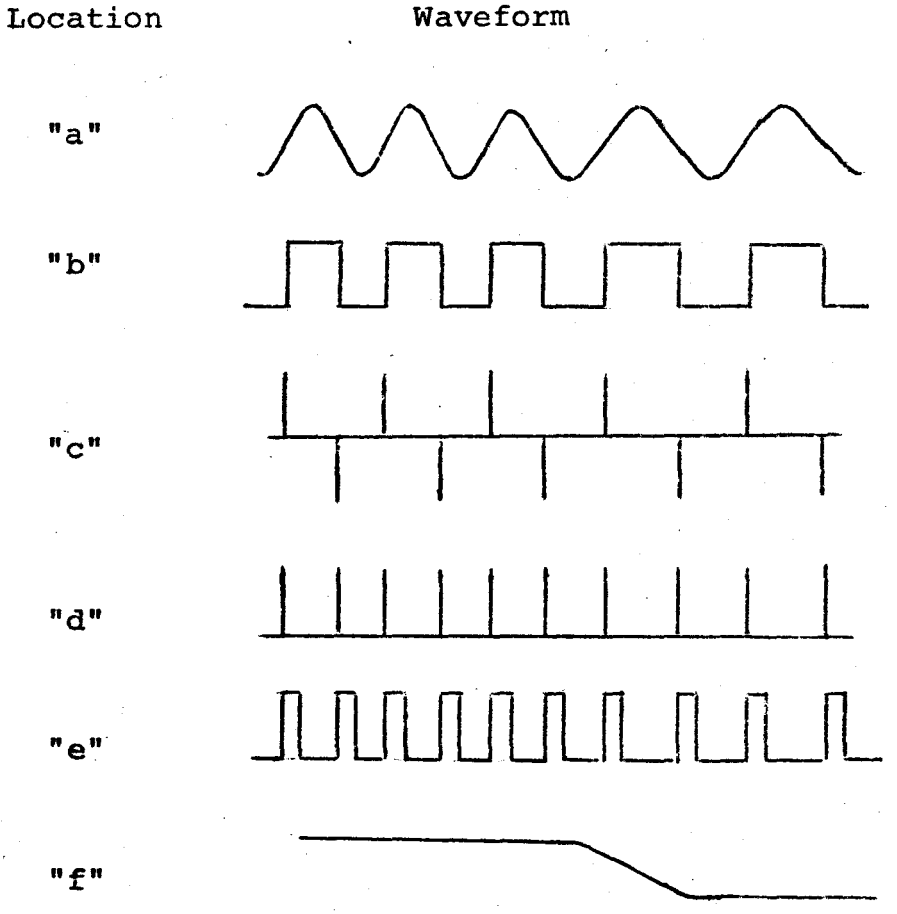


fig.6.4 Zero-Crossing Detector Discriminator

$$t_{n+1} - t_n \simeq \frac{\pi}{\omega_c + \dot{\theta}_m(t_n)} \tag{6.57}$$

If the low-pass filter is designed such that it produces the average dc value of the pulse train in the interval between the beginning of successive pulses, then the output of the filter will be

$$V_{zcd} = \frac{l_z h_z}{t_{n+1} - t_n} \tag{6.58}$$

and thus from equation (6.57)

$$V_{zcd} \simeq (l_z h_z/\pi) \cdot (\omega_c + \dot{\theta}_m(t_n)) \tag{6.59}$$

Thus, the short term average value of the low-pass filter gives approximately the derivative of the phase of the input signal. It is obvious that the accuracy is limited by the approximation made in equation (6.56). As $t_{n+1} - t_n$ increases in magnitude in comparison with $\dot{\theta}_m(t_n)$, the output of the low-pass filter represents less and less the derivative of the phase.

A direct analysis of the ZCDD is extremely difficult as far as error rate and distortion analyses are concerned, and no theory is at present available which can adequately descibe its performance.[7] An exact analysis would be of little value in this case anyway, since as will be descibed later, the performance requirements on the ZCDD are not particularly stringent because of the function this discriminator performs in the overall system.

CHAPTER SEVEN

SYNCHRONIZATION

7-1 Introduction

In order to reconstruct the transmitted sequence at the receiver correctly, a timing reference is required. That is, the received data pulses must be sampled at their maximum values (at the widest opening of the "eye pattern") and not at the transition points. However, since this timing information is not generally available at the receiver directly, it must be extracted from the received signal. This is usually accomplished by employing one of the following two methods.

One method is to send two auxiliary tones outside the band occupied by the spectrum of the transmitted signal which are so related, that when they are modulated together the difference frequency produces the desired timing reference while eliminating the carrier offset problem. However in the presence of phase distortion, there will in general be a discrepancy between the ideal timing reference and that obtained by this method.

The other, more commonly used method, is that of employing a PLL to extract the timing information contained in the data transitions. This information is obtained by generating a series of unipolar pulses which mark the trans-

ition times. The spectrum of this pulse train has a discrete component at the bit rate frequency which the loop can be locked on to.[1]

In the implementation of this system, a first order digital phaselock loop (DPLL) was employed to extract the timing information, the schematic of which is shown in figure (7.1).

7-2 Operation of the DPLL for Timing Extraction

The DPLL operates in a manner similar to the analog PLL in this application. The input signal is processed in exactly the same manner as previously described for the analog PLL; that is, the hard-limited output of the discriminator is differentiated and rectified - thus producing a pulse for every zero-crossing. The phase of the output of the divide by N flip-flop chain is compared in phase to the incoming transitions using logic elements to obtain sign $\phi_e(k)$, where

$$\phi_e(k) = \phi_i(k) - \phi_c(k) \qquad (7.1)$$

and $\phi_i(k)$ is the phase of the input at the k^{th} transition

$\phi_c(k)$ is the phase of the output of the divide by N flip-flop chain at the k^{th} transition

where $\text{sgn}(\phi_e(k)) = 0$ if no transition is made.

If $\text{sgn}(\phi_e(k))$ is positive, this signifies that the data is leading the clock in phase, which in the analog case would result in an increase in frequency of the VCO (i.e., the receive clock). In the DPLL, this increase in frequen-

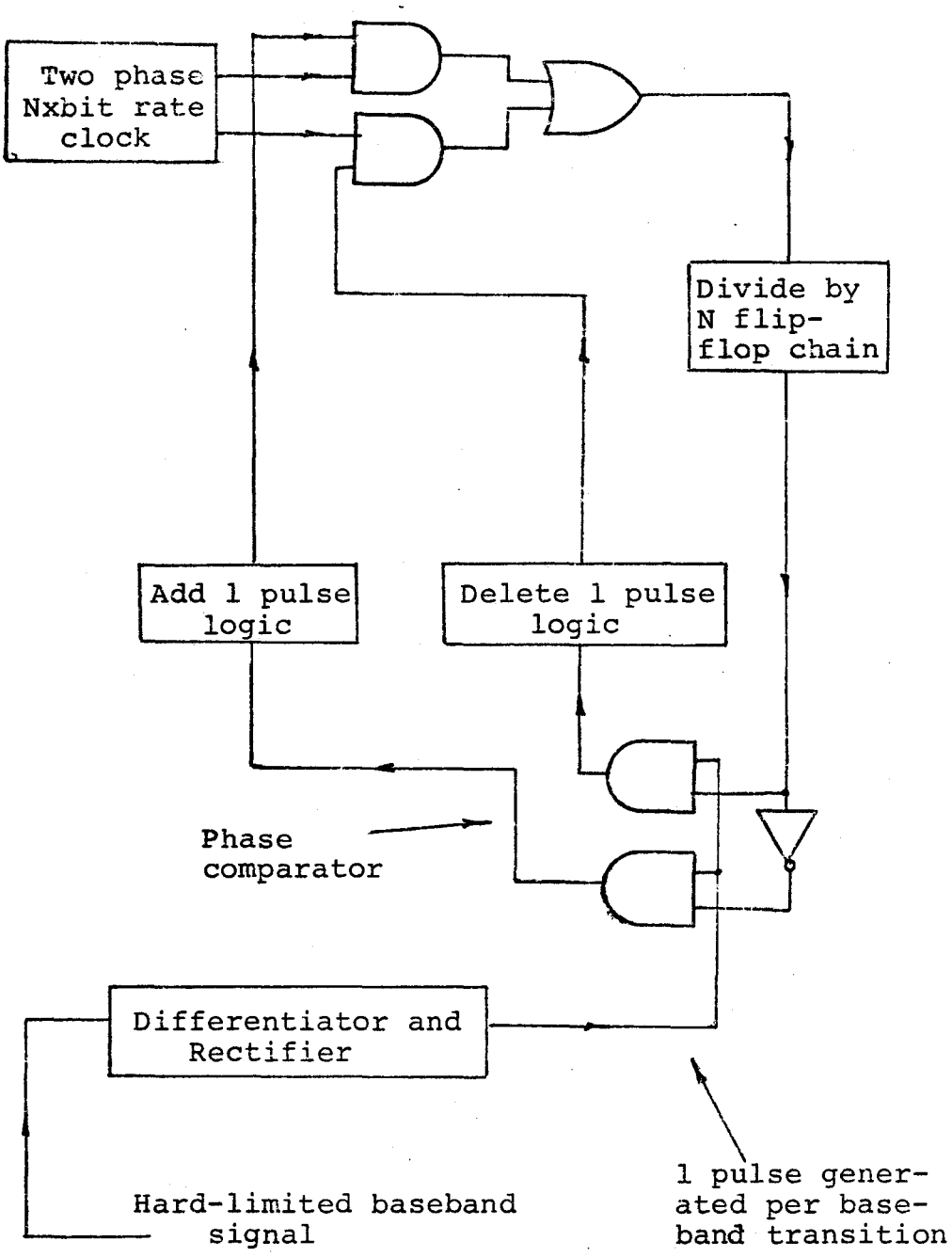


fig.7.1 Digital Phaselock Loop

cy is accomplished by injecting one pulse in the Nxl/T clock train. When this modified pulse train is passed through the divide by N flip-flop chain, the output is advanced in phase by 1/N cycles. A similar procedure is used if the data lags the clock, in which a pulse is deleted from the high speed pulse train. However, if no transitions are made (that is, a continuous one or zero is sent), then the DPLL runs at its idle frequency 1/T (Hz).

7-3 Locking and Acquisition Characteristics

The only parameter which can be varied in the DPLL is, of course, N. As N increases,

(1) The lock-in range decreases, which can be readily seen if one considers the response of the DPLL to the particular sequence 101010... . The maximum correction that the DPLL can achieve is one deletion or addition per transition and thus the lock-in range is ±1/(NT) Hz. Thus, it is apparent that as N increases the maximum allowable offset in frequency (which may exist because the transmit clock may be supplied by the data terminal, and thus is not directly under the designer's control) is decreased.

(2) The acquisition time increases; for example, if the worst case condition should occur, where the data and clock are 180° out of phase, the minimum acquisition time (which will occur for the 1010...

pattern) is NT/2 seconds.

(3) The phase jitter decreases, since effectively the averaging time is increased. This is a desirable feature since phase jitter in the receive clock results in a decrease in the effective signal to noise ratio because of the non-optimum sampling times. Increasing N corresponds to decreasing the loop bandwidth in the analog PLL. In a high noise or ISI environment, it may be desirable to make N as large as possible.

7-4 Data Dependence

One of the main disadvantages of the DPLL which is apparent from the foregoing discussion, is that its operational characteristics are data pattern sensitive. For example if long sequences of ones and zeros are present in the receive data the clock will drift off from the optimum sampling time at a rate determined by the difference in frequency between the transmit and receive clocks. For this reason, in practice so-called "scrambling" patterns are modulo 2 added to the transmitted data to ensure that the data has sufficient transitions to extract the required phasing information.

7-5 Selection of N

The value of N selected was 64. This value is commonly used in the implementation of modems designed for the voice network as it is found in practice to yield the

best compromise between the previously mentioned factors of averaging time, lock-in range, and acquisition time.[2]

CHAPTER EIGHT

THE CONTROLLER

8-1 Introduction

In this chapter the algorithm employed in the equalizer is considered.

8-2 Development of the Equalizer's Structure

Initially, the mean square error minimizing algorithm described earlier in chapter three was implemented at baseband (i.e. after discrimination). It was found that the equalizer converged, both in the absence and presence of noise, to settings which apparently removed a great deal of the ISI when distortion was present. However, in the presence of noise, even though the eye pattern "opened", the error rate remained essentially unchanged (improvements of 10 to 20% maximum were observed). This interesting result can be understood if one recalls the nature of the error mechanism described earlier in chapter six. There we found that the non-linear mixture of the noise and signal in the discriminator tended to increase the number of clicks produced when ISI was introduced into the system. Once a click is produced, because of its large amplitude, little can be done to mitigate it by post-detection filtering. However, it is possible that the effects of the non-click noise can be alleviated because its distribution is more gaussian in

nature (which would tend to explain the small improvement observed, since the effect of this component was earlier shown to be small).

Because of this severe limitation of the conventional baseband equalizer, an alternate structure was required. The obvious solution was to place the equalizer in front of the discriminator. With this modification made, the equalizer can alleviate the effects of ISI at passband before detection, and thus reduce the sensitivity to clicks.

8-3 Analysis of the Algorithm

To see how the algorithm arises, we follow the same method employed in chapter three for the linear MSD (mean square distortion) minimizing algorithm.

Denoting v_n as the output of the equalizer at passband at baseband sampling time n, then†

$$v_n = \sum_j'' c_j p_{n-j} \tag{8.1}$$

where again c_j is the j^{th} tap gain, and p_k is the passband input to the equalizer at time kT. Since the input to the equalizer is bandpass limited (note that the post-limiter bandpass filter is located after the TF as shown in figure (4.1) - however, since the summing operation is linear we can consider the input to be in a filtered form), we

† The notation $\sum_j''$ indicates summation over all 2M+1 taps such that $\sum_j'' = \sum_{j=-M}^{M}$.

can describe p_{n-j} as (approximately)

$$p_{n-j} \simeq \cos(\psi_{n-j}) \tag{8.2}$$

where ψ_{n-j} is the phase of the input FSK signal at sampling time (n-j)T, and the gain constant of the bandpass limiter is assumed to be unity. Equation (8.1) can be rewritten

$$v_n \simeq \kappa_n \cos(\alpha_n) \tag{8.3}$$

where

$$\kappa_n = [(\sum_j'' c_j \sin(\psi_{n-j}))^2 + (\sum_j'' c_j \cos(\psi_{n-j}))^2]^{1/2} \tag{8.4}$$

$$\alpha_n = \arctan[(\sum_j'' c_j \sin(\psi_{n-j}))/(\sum_j'' c_j \cos(\psi_{n-j}))] \tag{8.5}$$

If we now assume that the action of the post-detection low-pass filter is only to remove higher order harmonics, then the output of the ideal discriminator ($\dot{\alpha}_n$) is (from equation (8.5))

$$\dot{\alpha}_n = (1/\kappa_n) \cdot (\sum_j'' c_j \dot{\psi}_{n-j} \cos(\alpha_n - \psi_{n-j}) \tag{8.6}$$

If the effects of ISI are small, then at the sampling times ψ_{n-j} and α_n will be approximately positive integer multiples of 2π, thus equation (8.6) reduces approximately to

$$\dot{\alpha}_n \simeq (1/\kappa_n) \cdot \sum_j'' c_j \dot{\psi}_{n-j} \tag{8.7}$$

As in the linear case, the mean square distortion is defined (we have again assumed the noiseless case)

$$M_d = \overline{(\dot{\alpha}_n - s_n)^2} \tag{8.8}$$

where s_n contains a positive bias term due to the fact that

the discriminator output has a constant bias term of ω_c.

The k^{th} component of the gradient (with respect to $\underline{c}$) is given by

$$\frac{\partial M_d}{\partial c_k} = 2 \cdot \left(\frac{\partial \dot{\alpha}_n}{\partial c_k} \cdot (\dot{\alpha}_n - s_n) \right) \tag{8.9}$$

From equations (8.7) and (8.4) (after reduction)

$$\frac{\partial \dot{\alpha}_n}{\partial c_k} \simeq (1/\kappa_n) \cdot \dot{\psi}_{n-k} - (1/\kappa_n) \cdot \dot{\alpha}_n \cdot \cos(\alpha_n - \psi_{n-j}) \tag{8.10}$$

Since we are still assuming that the distortion is minimal, α_n and ψ_{n-j} will still be approximately multiples of 2π, and as a result equation (8.10) reduces to

$$\frac{\partial \dot{\alpha}_n}{\partial c_k} \simeq (1/\kappa_n) \cdot (\dot{\psi}_{n-k} - \dot{\alpha}_n) \tag{8.11}$$

furthermore from equation (8.4)

$$\kappa_n \simeq \sum_j c_j \tag{8.12}$$

From equations (8.12), (8.11), and (8.9), the k^{th} component of the gradient is

$$\frac{\partial M_d}{\partial c_k} \simeq (2/\sum_j c_j) \cdot [(\alpha_n - s_n) \cdot (\dot{\psi}_{n-k} - \dot{\alpha}_n)] \tag{8.13}$$

$$\simeq (2/\sum_j c_j) \cdot [e_n \cdot (\dot{\psi}'_{n-k} - \dot{\alpha}'_n)] \tag{8.14}$$

where the prime signifies the removal of the bias ω_c, and e_n is the output error.

At this point we must consider the physical constraints to be placed on the structure of the equalizer. In the linear case, the center tap served effectively to control the overall gain by ensuring that the cross-correlation

between the baseband signal at the center tap, and the output error, was zero. However, for this structure varying the gain of the TF would have no effect, since the ideal discriminator is independent of input level (for the PLL discriminator, only the loop parameters are affected). From this physical argument we can expect

$$\overline{e_n \cdot \dot{\alpha}'_n} \simeq 0 \tag{8.15}$$

and furthermore

$$\overline{e_n \cdot \dot{\psi}'_n} \simeq 0 \tag{8.16}$$

Thus, equation (8.14) reduces to

$$\frac{\partial M_d}{\partial c_k} \simeq (2/\sum_j c_j) \cdot \overline{[e_n \dot{\psi}'_{n-k}]} \tag{8.17}$$

In most automatic and adaptive equalizers the center tap is set to unity and all other taps are set to zero for the initial guess (preset values), consequently, if the channel encountered is not too distorted, the taps need only move slightly. If the taps are not far removed from their preset values then

$$\sum_j c_j > 0 \tag{8.18}$$

and equation (8.17) can be written

$$\frac{\partial M_d}{\partial c_k} \propto \overline{e_n \cdot \dot{\psi}'_{n-k}} \tag{8.19}$$

Equation (8.19) is of the same form as that encountered for the linear case (equation (3.22)) and suggests that we once again employ the method of gradient estimation with

the function

$$f(e_n \cdot \psi_{n-k}) = \operatorname{sgn}(e_n \cdot \dot{\psi}'_{n-k})$$
$$= \operatorname{sgn}(e_n) \cdot \operatorname{sgn}(\dot{\psi}'_{n-k}) \qquad (8.21)$$

where once again the control algorithm is ($(p+1)^{th}$ iteration)

$$c_k^{p+1} = c_k^{p} - f(e_n \cdot \dot{\psi}'_n) \qquad (8.22)$$

If, as in the linear case, we make the assumption that because the channel is not severely distorted that the initial probability of error is in the order of 10^{-3} say, then the sequence at the output of the discriminator (after slicing) will be a good approximation of the one transmitted. Thus, as in the linear case we can write

$$e_n \simeq \hat{e}_n \qquad (8.23)$$

where $\hat{e}_n = \dot{\alpha}'_n - \operatorname{sgn}(\dot{\alpha}'_n)$.

Thus, the decision-directed algorithm is

$$c_k^{p+1} = c_k^{p} - \operatorname{sgn}[\dot{\alpha}'_n - \operatorname{sgn}(\dot{\alpha}'_n)] \cdot \operatorname{sgn}(\dot{\psi}'_{n-k}) \qquad (8.24)$$

It is apparent that the detector required to determine $\dot{\psi}'_{n-k}$ need not be particularly accurate, in that, we require only polarity information, furthermore, a degradation in this detector of a factor of ten ($P_e \sim 10^{-2}$) would have little effect since the tap gains would be still moving in the correct direction 99% of the time. The storage device required at each tap (an integrator or up-down counter) will also mitigate the short term effects created by incorrect decisions. It was for these reasons that the ZCDD was employed to determine $\dot{\psi}'_{n-k}$.

Equation (8.24) as in the linear case is quite easy

to implement. The multiplication required can be realized by an "exclusive or" circuit. An examination of figure (8.1) which illustrates a possible implementation structure for the algorithm reveals the additional advantages over the linear equalizer. Perhaps the most significant advantage is the fact that the delays can be achieved by employing a long string of flip-flops clocked at the sampling rate of the sampled bandpass limiter. Long strings of flip-flops are available in integrated circuit form and are quite inexpensive. In addition, the implementation of the variable tap gains is quite simple since all that is required is a switch which switches off and on the dc value stored in the integrator. The whole structure could be completely implemented digitally (i.e. using up-down counters instead of integrators, etc.) - however, a rough estimate indicates that this method would be in the order of two times as expensive at present.

8-4 Limitations of the Equalizer Structure

(a) Non-random data

As just indicated, the structure of the equalizer has many implementational advantages - however, the fact that the discriminator (PLL) is essentially input level insensitive presents a problem, in that, in practice the data terminal equipment may present a 1 or 0 to the modulator during idle conditions. If this pattern were not scrambled, it would have a disastrous effect on the equal-

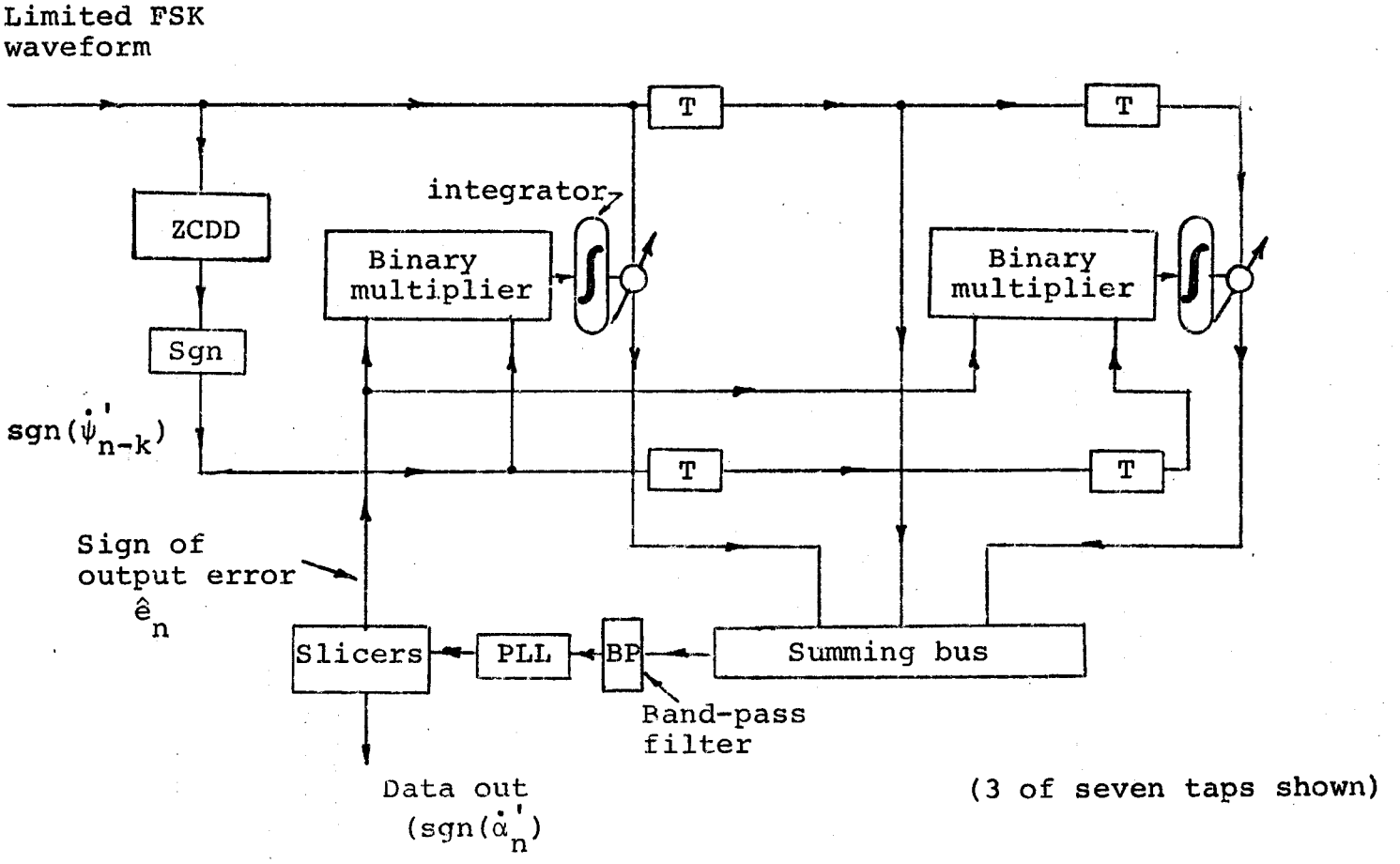


fig.8.1 Structure of Adaptive Receiver

izer because, unlike the linear equalizer this particular structure requires a constant change in the line signal which need not necessarily be random. This is apparent from figure (8.1). For example, if a steady 1200 Hz tone was received, the output slicer would interpret the baseband error to be positve or negative, consequently all the tap gains will head in one direction, because unlike the linear case there is no reflection in the baseband output from this particular variation in the tap gains. Thus, a scrambler would be required when employing this system (one was required for the DPLL in any case). The addition of a scrambler would also ensure more uniform convergence and tracking performance.

(b) Tapped Delay Line Length

As described in chapter three, the length of the equalizers structure effectively limits the degree of correction which can be achieved for a given channel impulse response.

It was not the intent of this thesis to determine the exact relationship between residual distortion and tapped delay line length, as in practice, tests would have to be performed to determine the optimum length employing representative facilities of the system to be used anyhow. However, in response to some of the earlier work of Lucky,[1] it was felt that a reasonable length of the TDL would be 7/1200 seconds, and this is the value which was implemented.

The main tap was placed in the center of the TDL in consideration of the results of Lender[2] and others, who found that this condition provided good results (in simulation and in practice) despite the fact that the impulse response observed tended to be asymetrical about the main peak in time.

CHAPTER NINE

IMPLEMENTATION - EXPERIMENTAL RESULTS

9-1 Implementation Philosophy

As can be readily appreciated, the large number of variables involved in the structure of the modem, and the non-linear nature of the system, make computer simulation a somewhat unwieldy task - furthermore, simulations of FSK systems tend to be misleading. A more direct approach was employed. The system was implemented with the philosophy of maintaining as many of the important parameters as variables such as:

(a) The signalling rate

(b) All slicing levels

(c) Tap gain time constants of the equalizer

(d) The VCO center frequency and d c gain of the PLL

(e) The tap gains of the transmitting BTF

(f) When low-pass filtering was required simple one amplifier Sallen-Key type filters were used in tandem which allowed design changes to be made relatively quickly and easily

With these parameters accessible, the following optimization procedure was employed. Initially, the modulator was designed with the post-modulation analog filter and BTF set to obtain a uniform envelope for the FSK signal employing a

2047 bit pseudo-random pattern. Then the PLL was implemented with the initial loop parameters suggested in equations (6.24) and (6.25), and the post-modulation filter as well as the loop parameters were adjusted to eliminate ISI. Next white noise was introduced into the transmission path and the above mentioned parameters were adjusted to minimize the error rate (once again employing the 2047 bit pattern). The input bandpass filter of the receiver was then designed such that it did not introduce any significant ISI but did reduce the noise bandwidth. With this initial design then completed each parameter as well as combinations of parameters were varied in order to reduce the error rate. It was found in this latter procedure that only minor variations in the parameters concerned were required and furthermore that the optimum encountered was very broad.

Having realized the basic system, the adaptive equalizer was then implemented. The tap gain time constants were adjusted to obtain optimum convergence and residual error characteristics for the worst case signal to noise ratio condition that would be expected on a typical channel (25 dB, flat weighting). Several distorting channels were employed in the latter optimization process, using the channel simulator developed. The resulting structure was then evaluated for several more channels again employing the channel simulator.

9-2 Hardware

The modem was constructed on nine cabinet mounted cards (approximately 7"x5" in dimension). A schmematic of each card is not provided since the circuits are relatively straight-forward in nature.

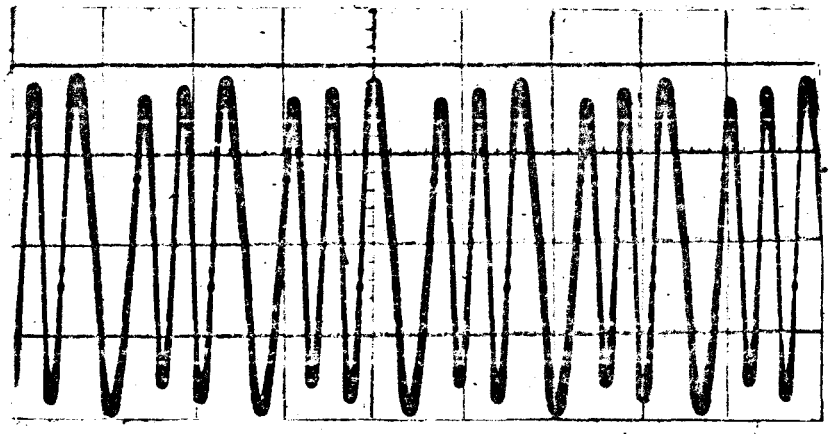
The three power supplies ($\pm$12,+5volts nominal) required are external to the cabinet.

Throughout the design of the modem integrated circuits were used wherever possible in order to take advantage of the size and performance improvements offered over discrete implementations.

It should be pointed out that the structure of the tap gain elements was in fact more complex than required. This occurred because the somewhat complicated structure employed in earlier investigations of baseband equalization was used directly in the final design.

9-3 Line Signal

Figure (9.1) shows the output of the modulator for a reversals (1010...) pattern. As can be seen from the photograph, the envelope of the signal is relatively constant, although it was observed that the envelope of the signal at the output of the pre-detection bandpass filter was even more uniform - indicating that the overall shaping was probably quite similar to that required to obtain the ideal FSK waveform.



Horizontal - 1 msec/div.

fig.9.1 Line Signal for Reversals Pattern

9-4 Amplitude Spectral Analysis of a Pulse

Figure (9.2) illustrates the amplitude spectral analysis of the modulator for a pulse of 1200 and 2400 Hz every 32T seconds. As can be observed, the two spectra are the same in the passband (excluding the discrete tones) as was expected from the earlier discussion in chapter five and furthermore they are of the form suggested by equation (5.4). The discrete tone is present at 3600 Hz for the 2400 Hz pulse because the analog low-pass filter implemented did not have a severe enough roll-off characteristic to suppress the third harmonic of the 1200 Hz clock. In passing it should be pointed out that as the pulse repetition rate decreases the closer the discrete contributions in the resulting spectrum become, consequently the more it resembles the true envelope spectrum.

9-5 Power Spectral Density

The power spectral density of the line signal is illustrated in figure (9.3) (2047 bit pseudo-random data sequence transmitted). The power in the discrete tones is approximately one half of the total power, which accounts for the 3 dB poorer performance of FSK as compared to DSB-SC-AM. Note that over 80% of the energy is concentrated between 1200 and 2400 Hz - which is desirable since the phase-frequency and amplitude-frequency responses are usually acceptable in this area.

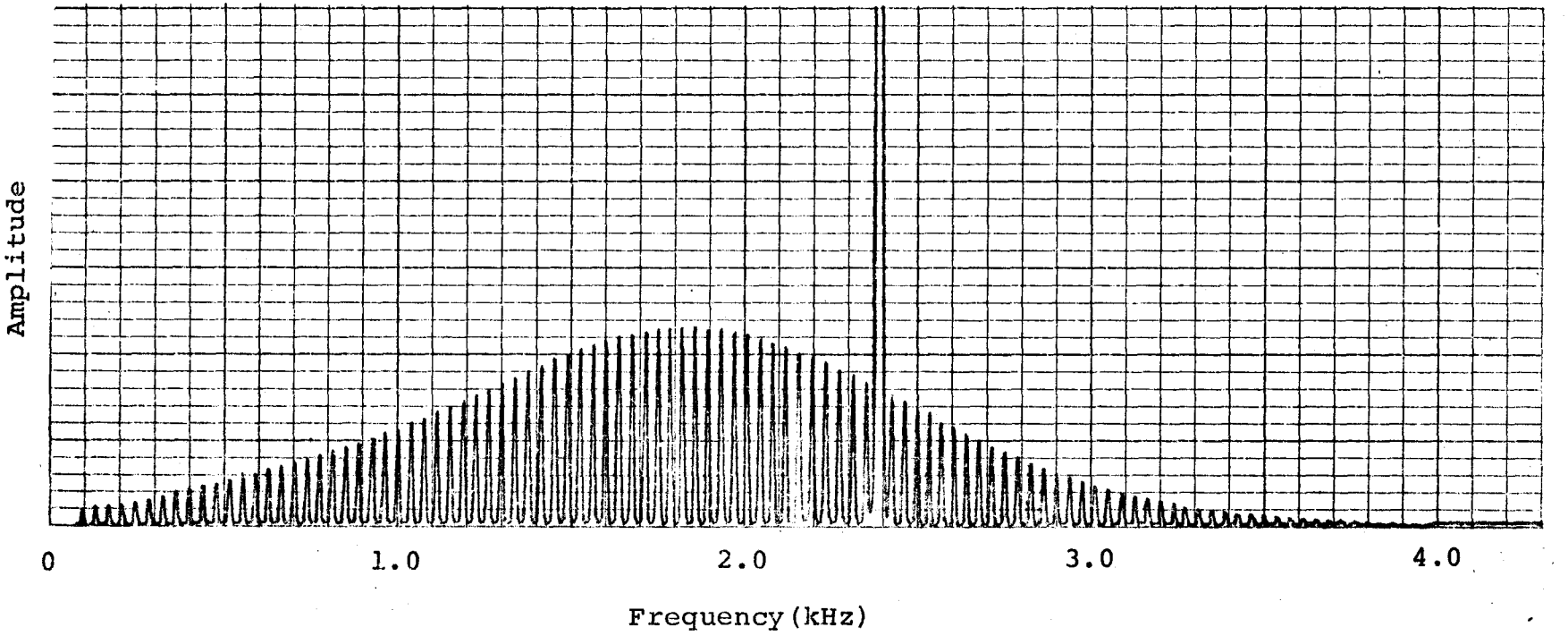


fig.9.2(a) Amplitude Spectral Analysis of Modulator Output for a Pulse of 1200 Hz Every 32T Seconds

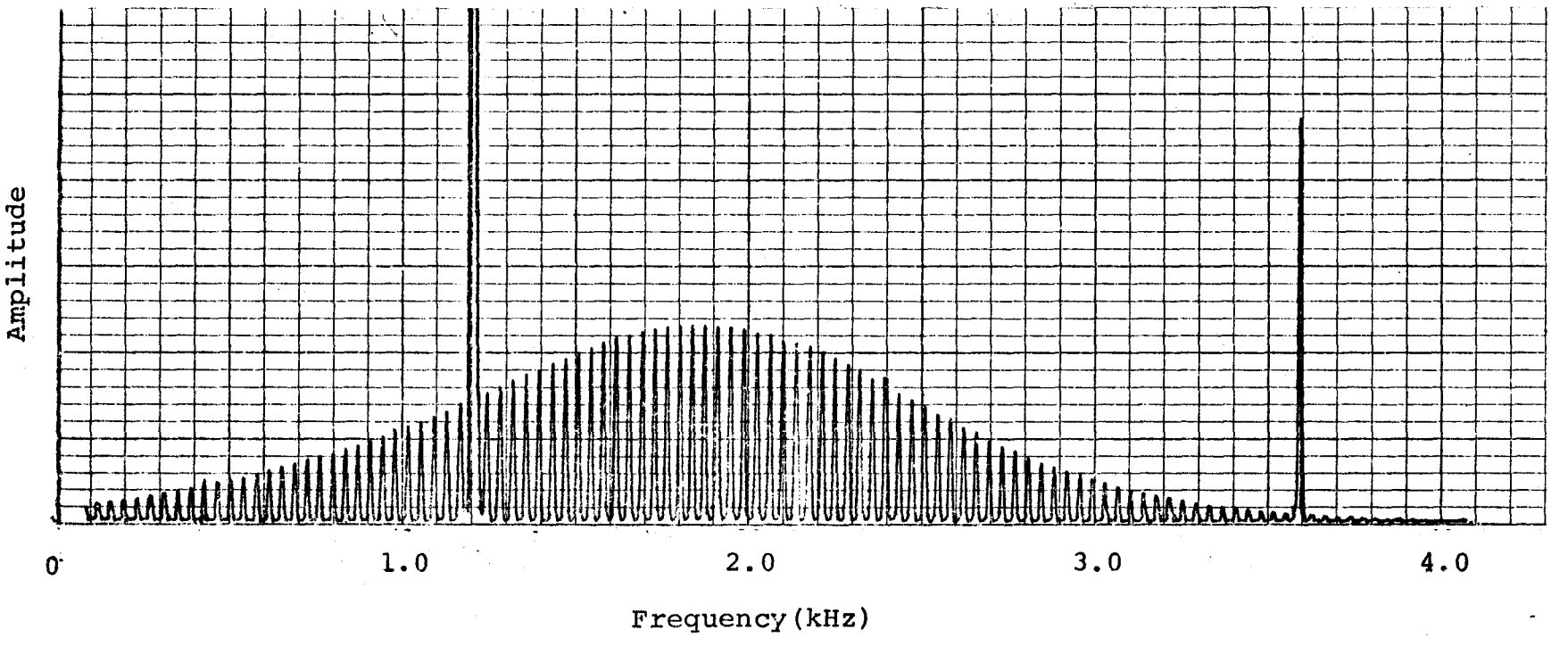


fig.9.2(b) Amplitude Spectral Analysis of Modulator Output for a Pulse of 2400 Hz Every 32T Seconds

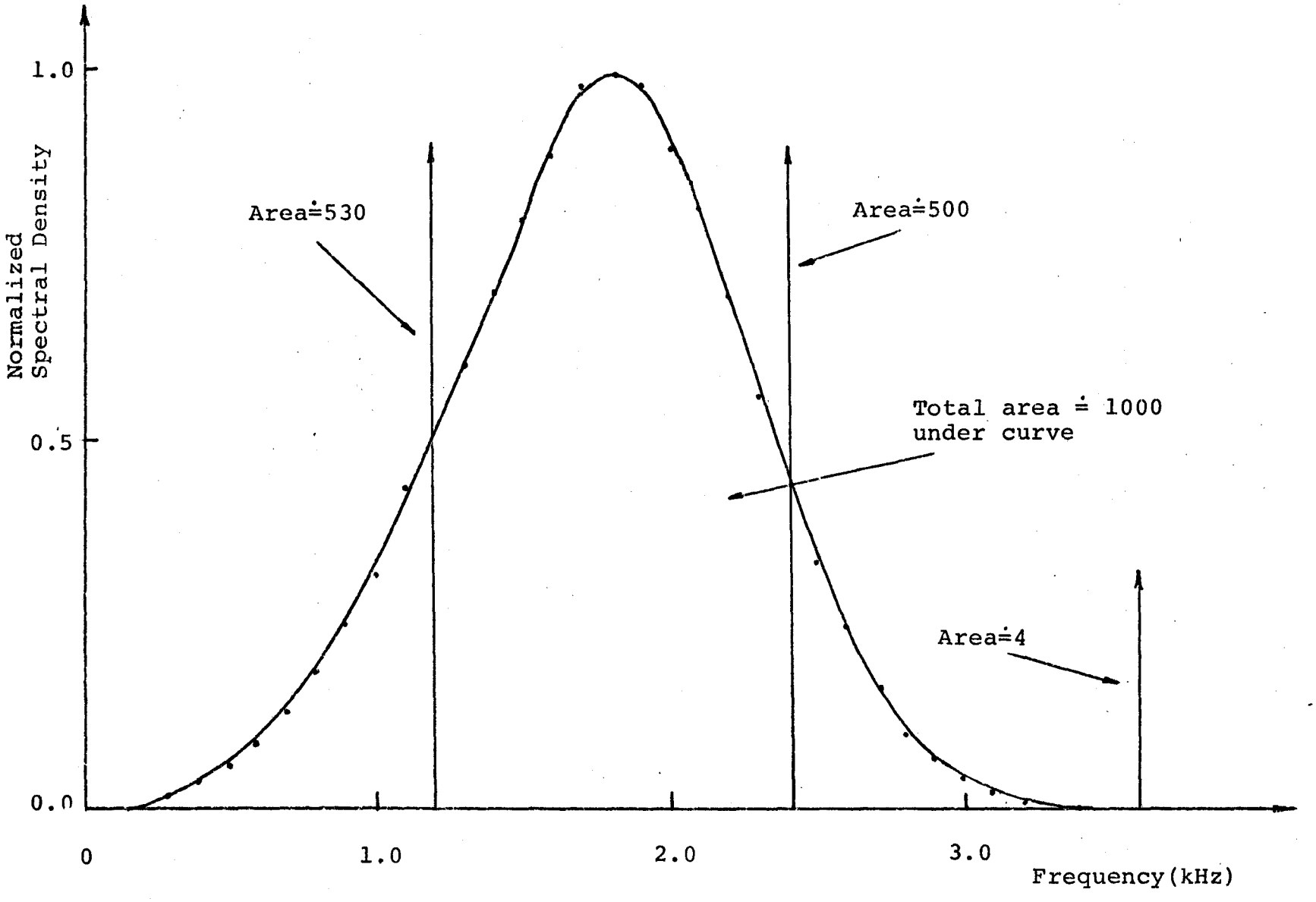


fig.9.3 Power Spectral Density of Line Signal

9-6 PLL Characteristics

Figure (9.4) illustrates the static characteristics of the loop. Examining the characteristics it is apparent that the PD has a sinusoidal transfer function and furthermore that the assumption of linearity is quite justified between 1200 and 2400 Hz. In addition, it is evident that the discriminator characteristic is quite linear. This observation implies that the VCO characteristic is also quite linear, since from equation (6.15) the linearity of the discriminator characteristic depends only on the linearity of the VCO conversion gain.

From figure (9.4(b)) and referring to appendix C the dc loop gain (in the linear region) is

$$K_{vco}K_{pd} \simeq 6 \cdot 10^{3} \tag{9.1}$$

The time constant of the loop filter in the final design is (see appendix B card 4)

$$\tau \simeq 2 \cdot 10^{-4} \text{ seconds} \tag{9.2}$$

From equations (6.12) and (6.13) (assuming linearity)

$$\omega_n \simeq 5.4 \cdot 10^{3} \tag{9.3}$$

$$\zeta \simeq .46 \tag{9.4}$$

These two values compare relatively closely to expected values indicated in equations (6.24) and (6.25). As previously discussed the difference arises because of the presence of the pre-detection filter which allows for the use of a somewhat wider loop bandwidth and smaller

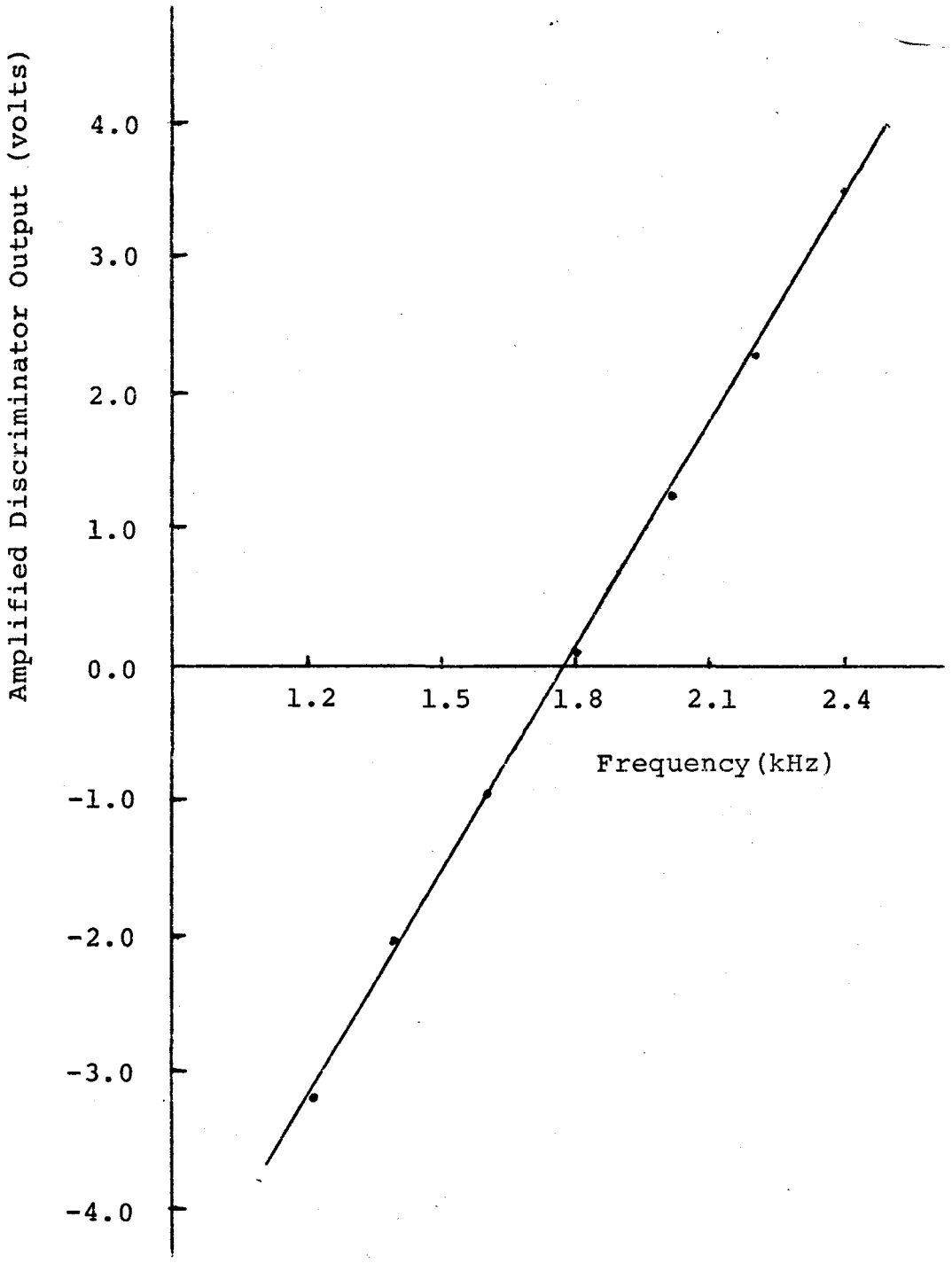


fig.9.4(a) Discriminator Characteristic

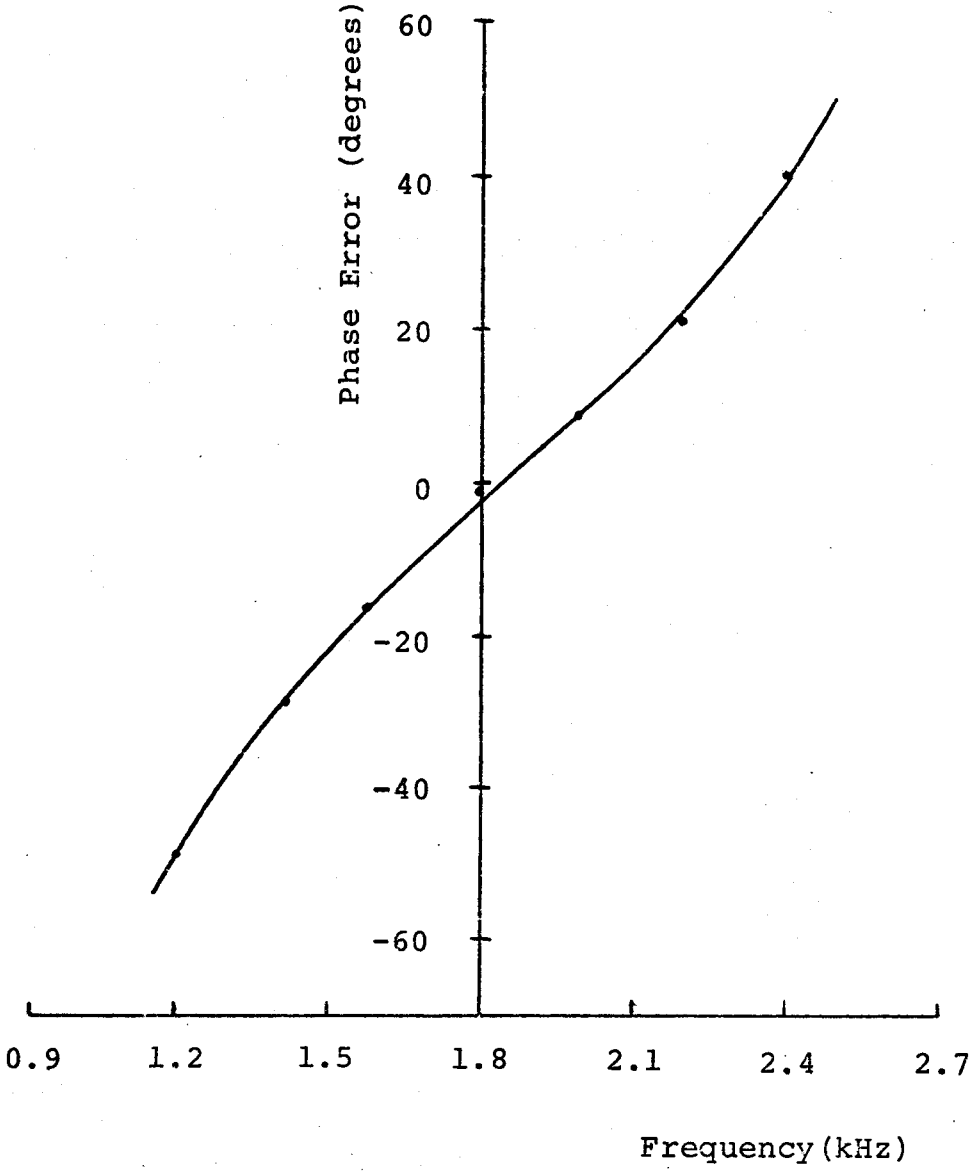


fig.9.4(b) Static Phase Error as a Function of Frequency

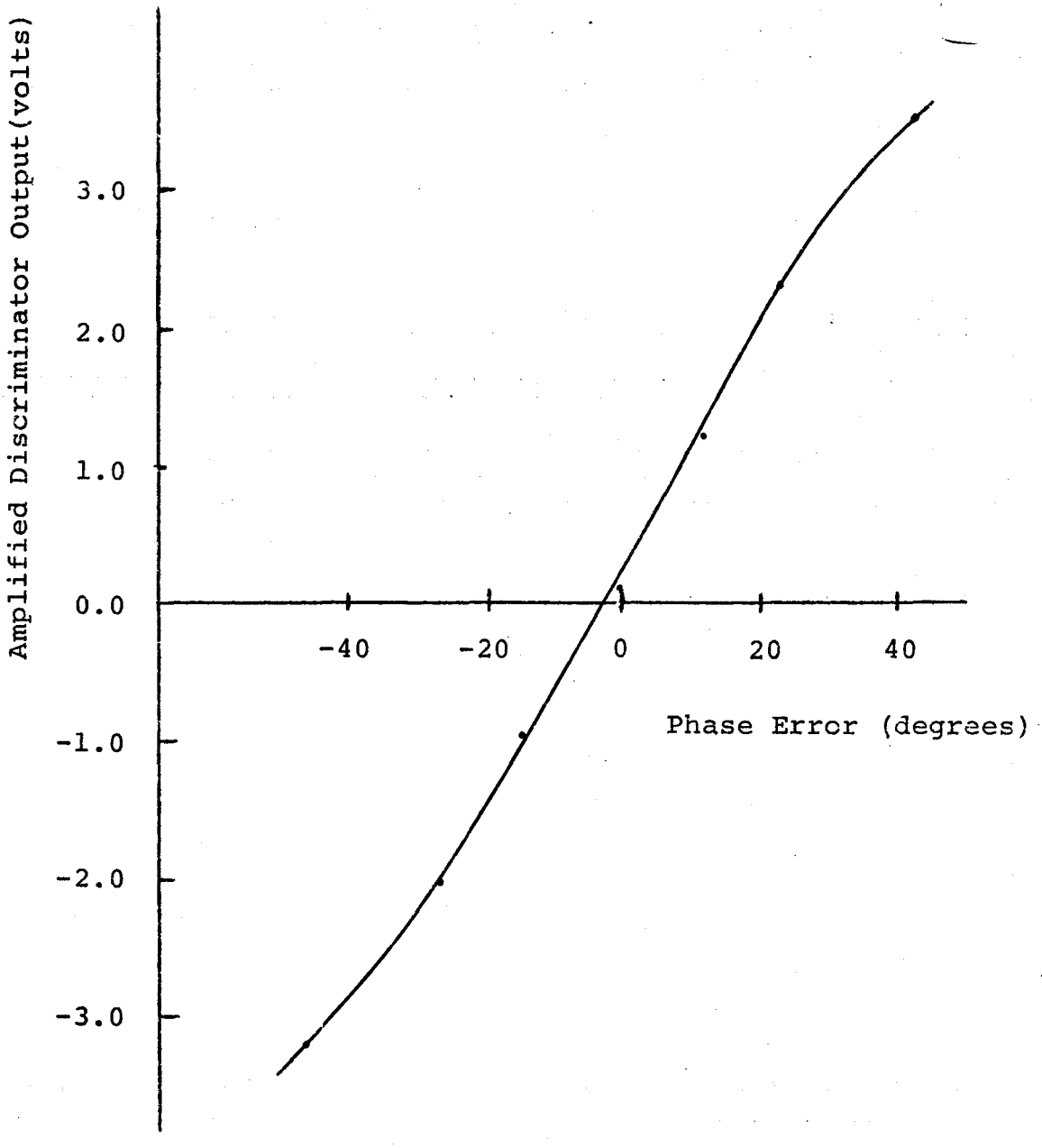


fig.9.4(c) Phase Detector Characteristic

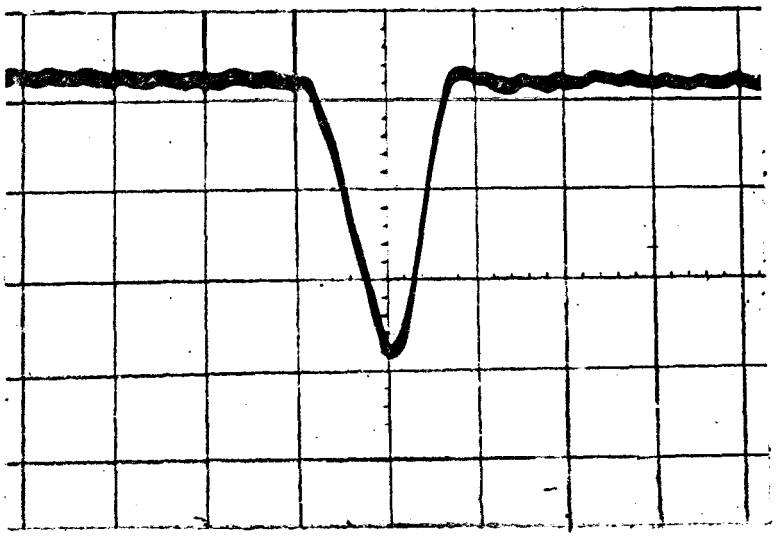
damping ratio.

9-7 Pulse Response

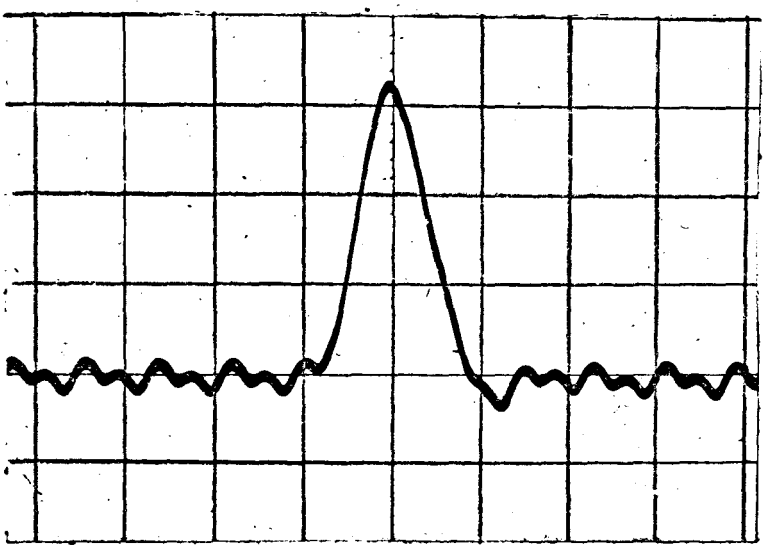
Figure (9.5) illustrates the pulse responses of the overall system (equalizer out). The shape of the pulse responses is similar to that of the raised cosine class with a 100% roll-off. The ripple in the steady state response is due to high frequency harmonics produced in the PLL but not completely removed by the post-detection filter. The effect of this ripple is small as can be seen from the eye pattern shown in figure (9.6). As can be observed from this latter photograph the residual peak distortion is small (approximately .05) and furthermore that the effects of sampling are apparently small if not negligible.

9-8 Channel Simulation

The arrangement shown in figure (9.7) was employed to determine the BPE and MSE characteristics of the modem. A schemetic of filter F1 is provided in appendix D. The signal to noise ratio was measured at the output of the second band-pass filter (F1). The correction factor for this latter filter was calculated to be 3.89 dB (i.e. the signal to noise ratio measured with the true RMS meter was increased this amount to account for the fact that we are only interested in measuring the noise power between 1200 and 2400). The auto-correlator was used in conjunction with a peak detector to obtain the maximum value of the auto-correlation

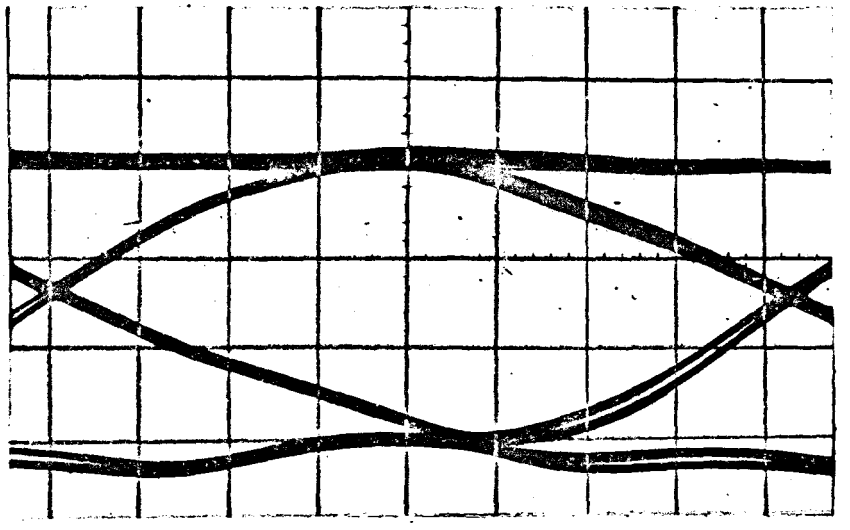


(a) Hor. - 1 msec/div.



(b) Hor. - 1 msec/div.

fig.9.5 Pulse Response. (a) Pulse of 1200 Hz Received. (b) Pulse of 2400 Hz Received



Hor. - .1 msec/div.

fig.9.6 Back to Back Eye Pattern

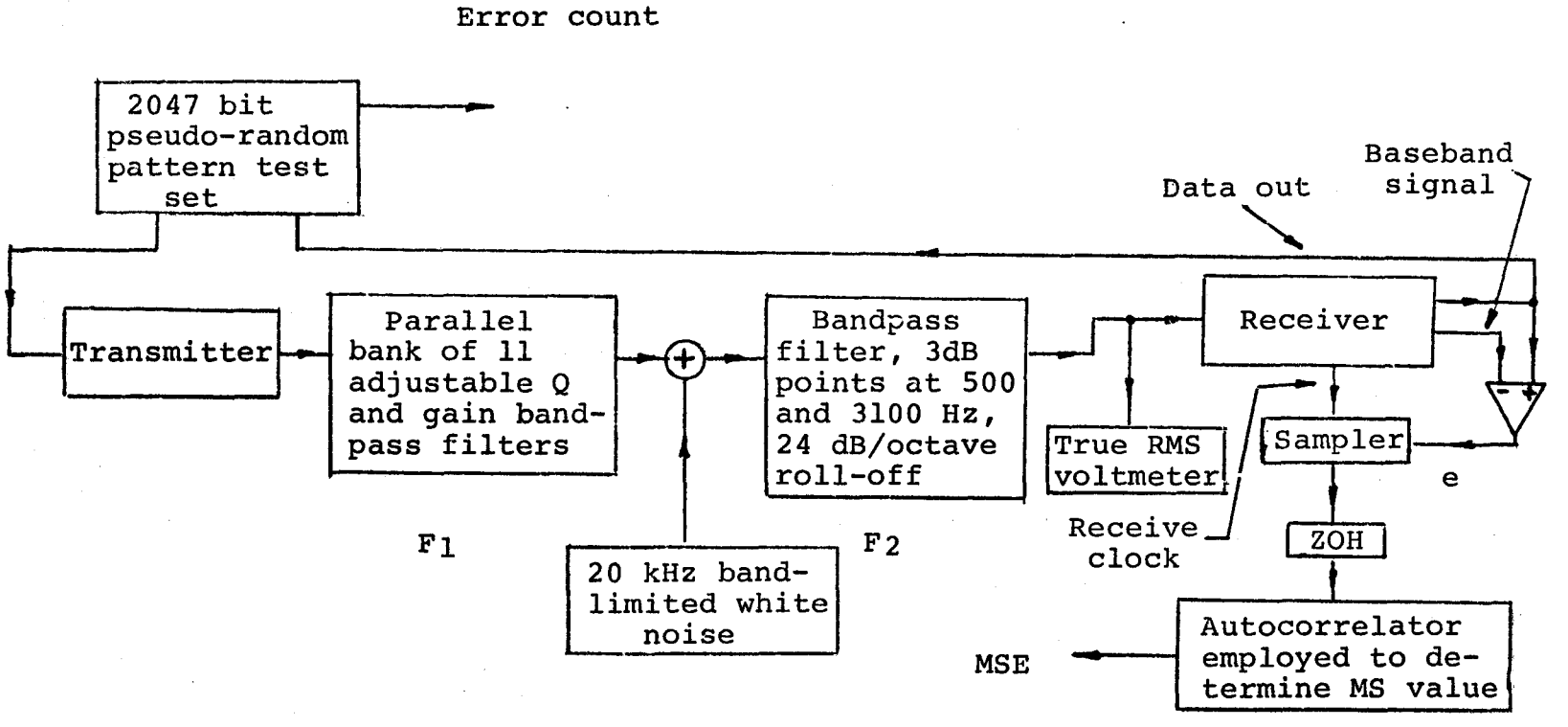


fig.9.7 Experimental Arrangement Employed to Determine MSE and BPE Characteristics

function which of course gave the mean square value of the error (approximately instantaneously).

The 2047 bit pseudo-random test set was used to estimate the BPE. Since the pattern is quite long in comparison with the effective channel memory time, the results obtained with the test set should be almost the same as those which could be obtained employing true random data. In order to obtain an accurate estimate of the BPE, the data sequence was repeated until 100 or more errors were recorded. The tests were only performed for BPE's greater than 10^{-5} or so, since for BPE's less than this, the time required to obtain enough errors for an accurate estimate becomes quite long.

9-9 BPE for the Ideal Bandpass Channel

Figure (9.8) shows the BPE for varying signal to noise ratios (F1 removed). As can be seen from the graph, the PLL type discriminator is approximately .9 dB worse than the optimum, while the ZCDD is approximately 2 dB worse than the optimum. However, one must consider the effect of the bandpass filter on the system, for it was found that when the bandpass was removed, (white noise in the system) that the PLL discriminator performed approximately .7 dB worse than previously indicated. This occurred because the noise outside the passband of F2 had some effect (approximately .5dB), and because some of the spec-

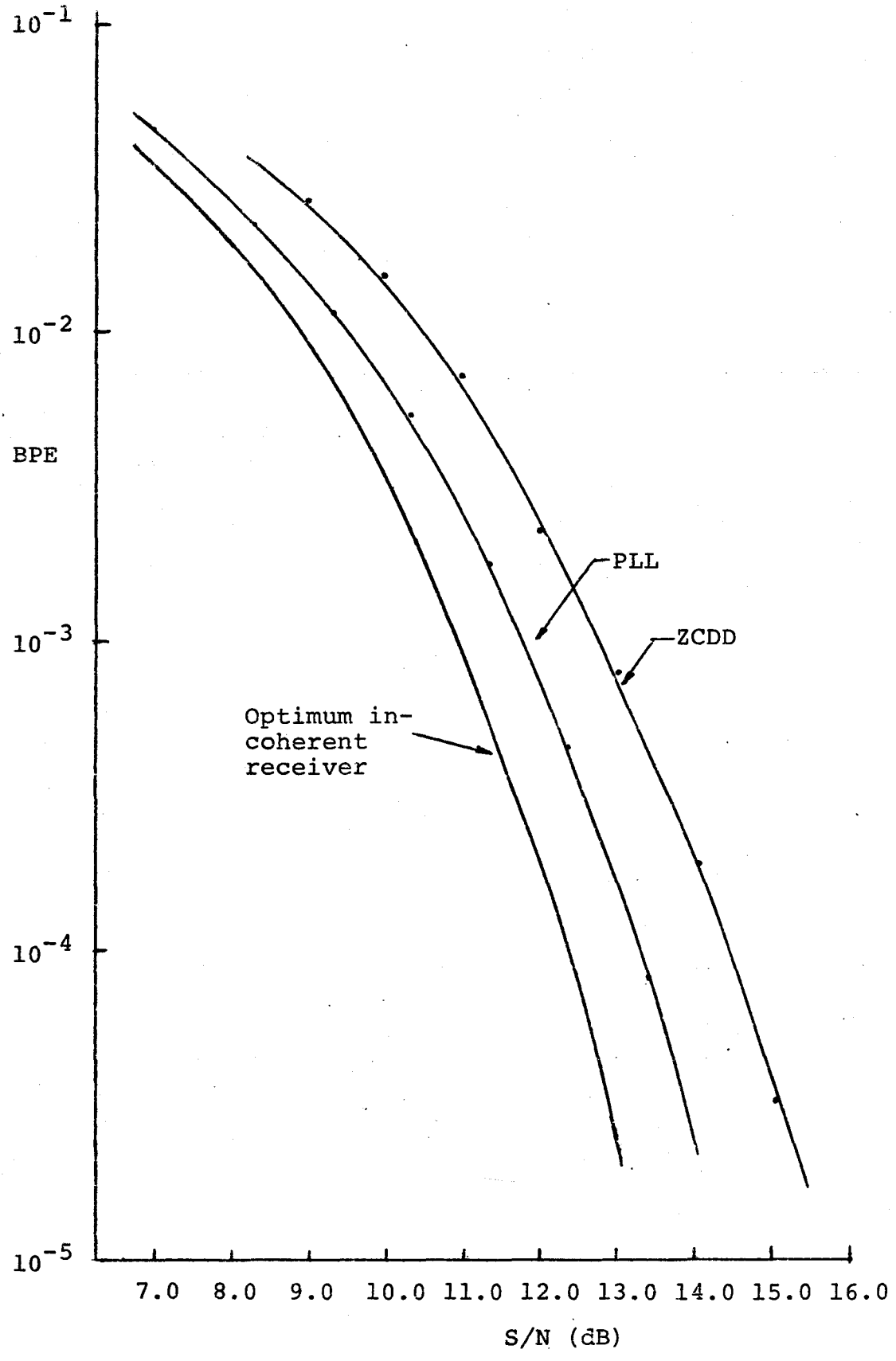


fig.9.8 BPE as a Function of SNR for PLL and ZCDD

trum of the transmitted FSK signal was outside the passband of F1 and thus represented wasted power since it was removed by the input bandpass filter of the receiver. This latter effect accounts for the other .2 dB. For the basic system then, the degradation encountered in white noise alone is in the order of 1.6 dB, which is quite acceptable considering the type of filters employed in the system, and the optimization method employed.

It should be pointed out that the additional degradation due to noise outside of the 3 kHz bandwidth would not usually be encountered in practice because of the band-limiting which occurs in most channels.

9-10 Investigation of the Non-Ideal Channel

In order to investigate the effectiveness of the adaptive system, approximately 15 channels were investigated in some detail. To summarize the performance, detailed results of three channels are presented in figures (9.9) to (9.21).

(a) Channel 1 - Moderate Distortion

The phase-frequency and amplitude-frequency characteristics of this channel (illustrated in figure (9.9) are typical of the type of channel to be expected in practice in the switched network.

Figure (9.10) shows the output MSE (approximately instantaneous) for two signal to noise ratios. As can be

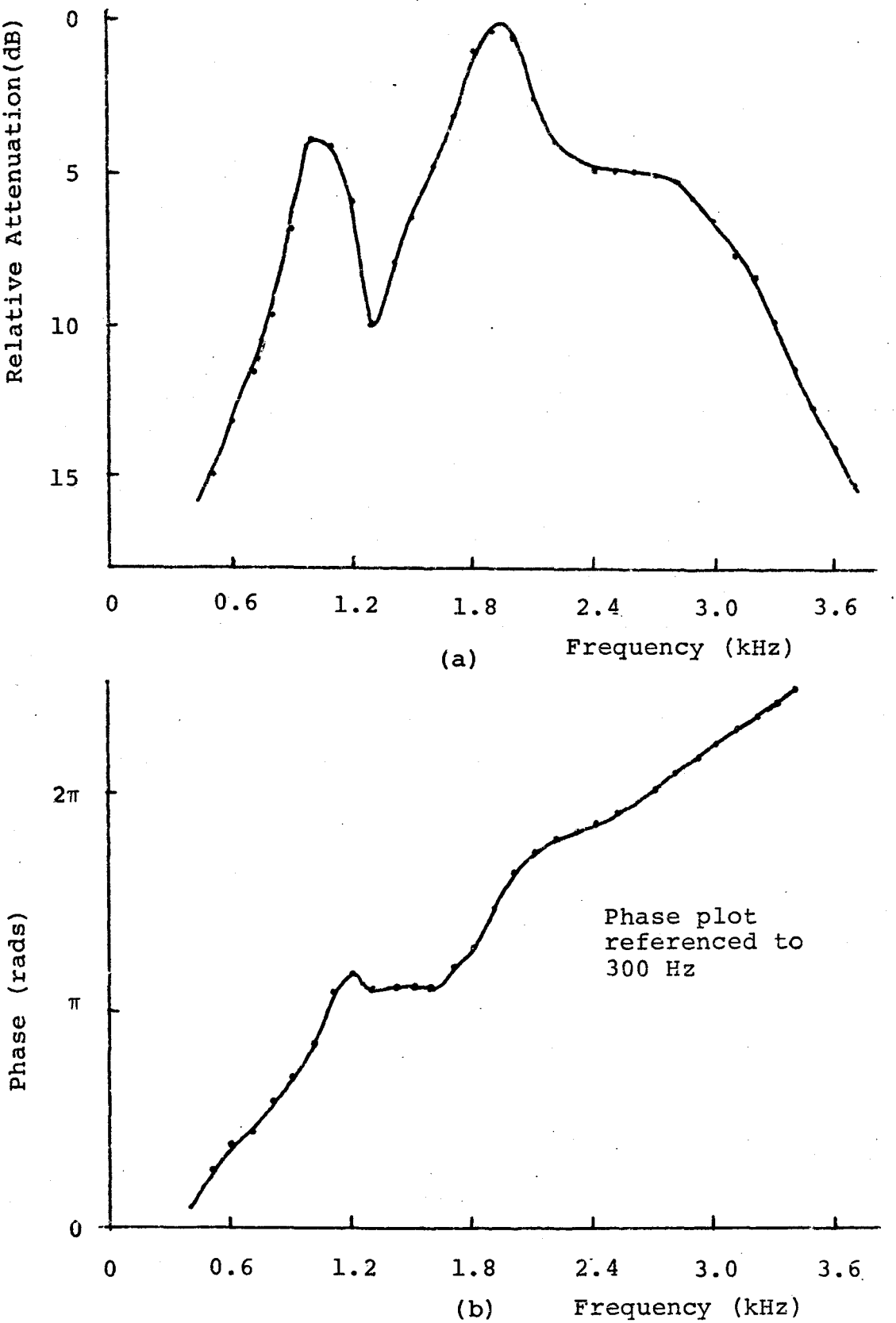


fig.9.9 Channel 1. (a) Attenuation Characteristic (b) Phase Characteristic

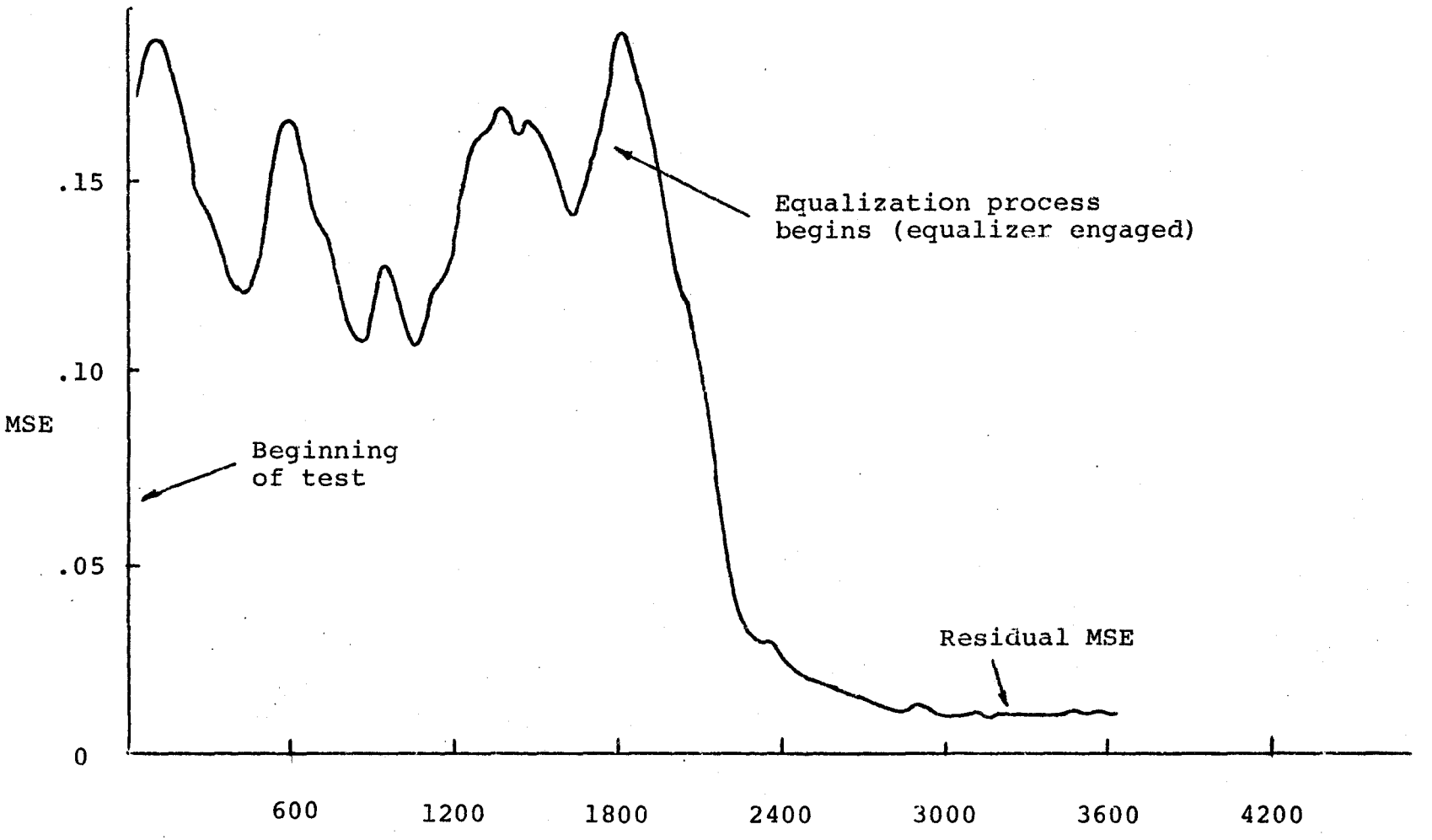


fig.9.10(a) Typical Convergence Curve at 28 dB (Channel 1)

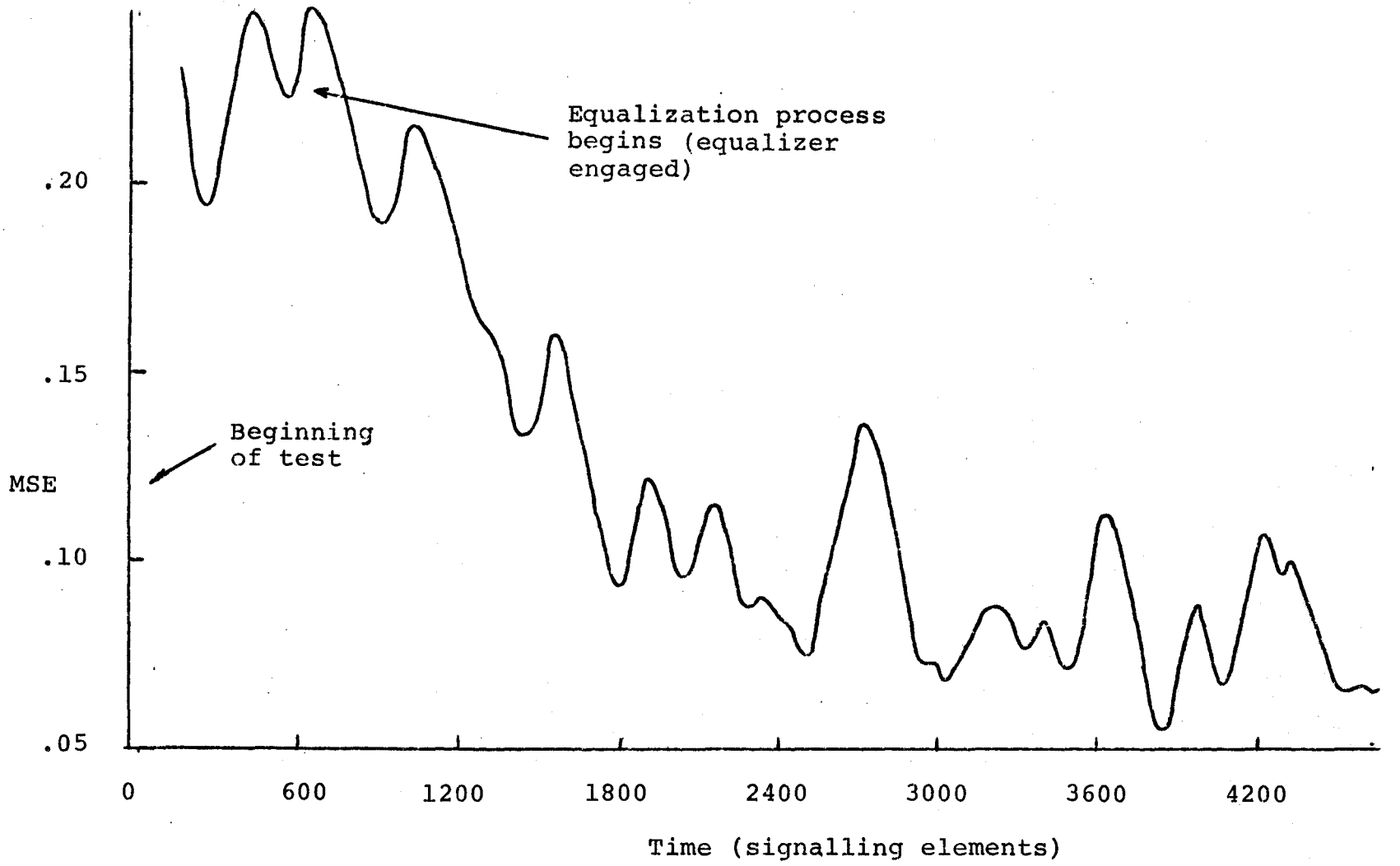


fig.9.10(b) Typical Convergence Curve at 11 dB (Channel 1)

observed, for the lower signal to noise ratio the convergence time is much longer, which is to be expected because of the increased uncertainty in the estimates of the gradient. This effect can be readily observed in figure (9.11). The convergence time and limit of convergence are directly related to the chosen tap time constant (approximately .5) seconds). For example, from figure (9.11), it is apparent that below 10 dB the equalizer does not converge (in the steady-state sense) to settings which would decrease the probability of error - however, if the time constant of the tap gains had been increased convergence could have been obtained for this SNR - at the expense of increasing the convergence time in the operating range (above 28 dB corrected, 25 dB flat).

Figure (9.12) is perhaps the most important in determining the effectiveness of the receiver, for here the effect of the adaptive equalizer can be seen in terms of error rate improvement. At 20 dB, the error rate has been decreased by approximately a factor of 15 - which is a significant reduction.

The probability distribution of the received baseband signal before and after equalization is shown in figure (9.13). As can be observed, the baseband gain has varied only slightly if at all during equalization. This result tends to verify the assumptions made earlier in equations (8.15) and (8.16), since it was observed that the

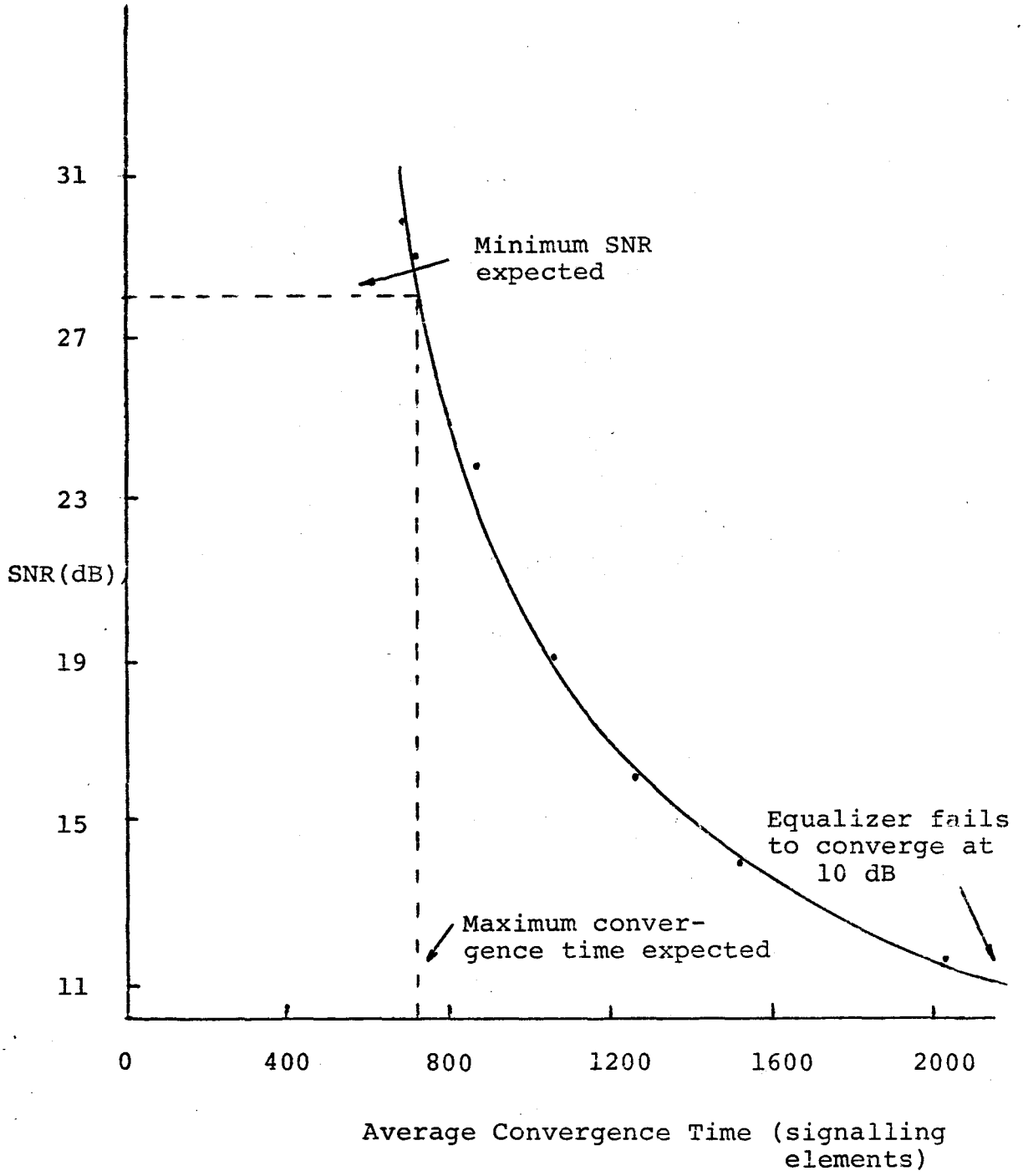


fig.9.11 Convergence Time for Varying SNR

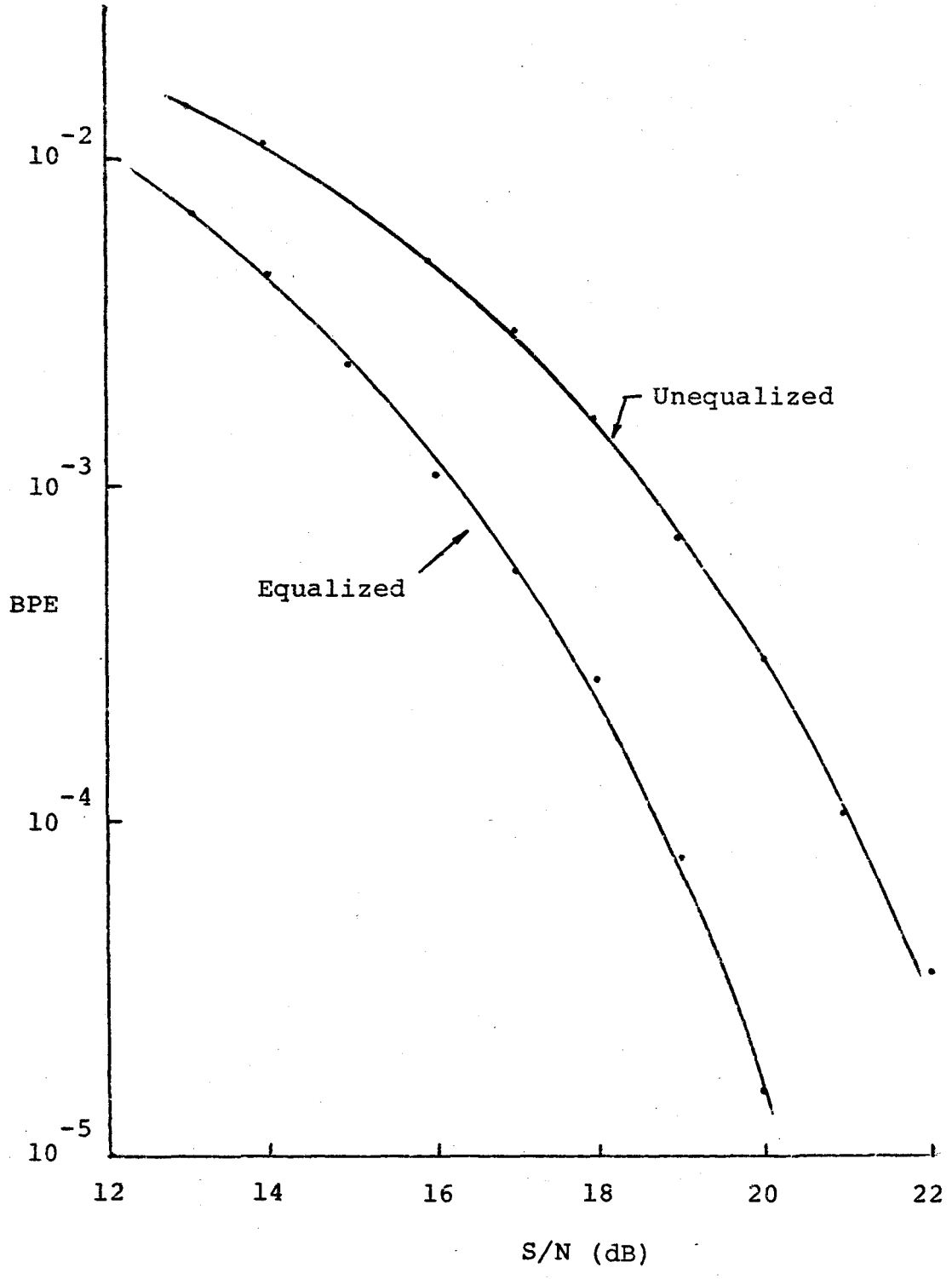


fig.9.12 BPE Characteristics (Channel 1)

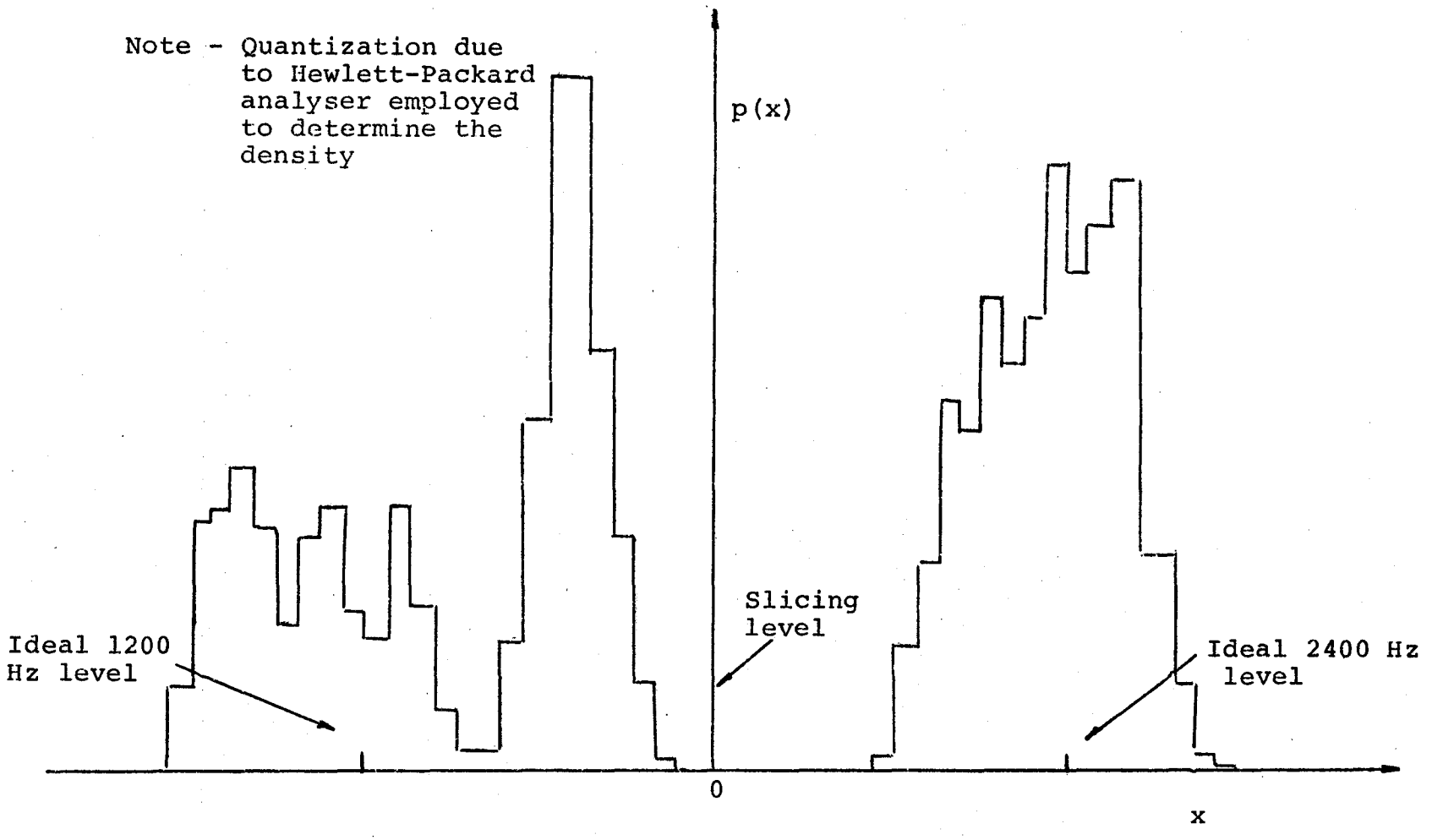


fig.9.13(a) Probabilty Density of Sampled Baseband Output before Equalization (Channel 1)

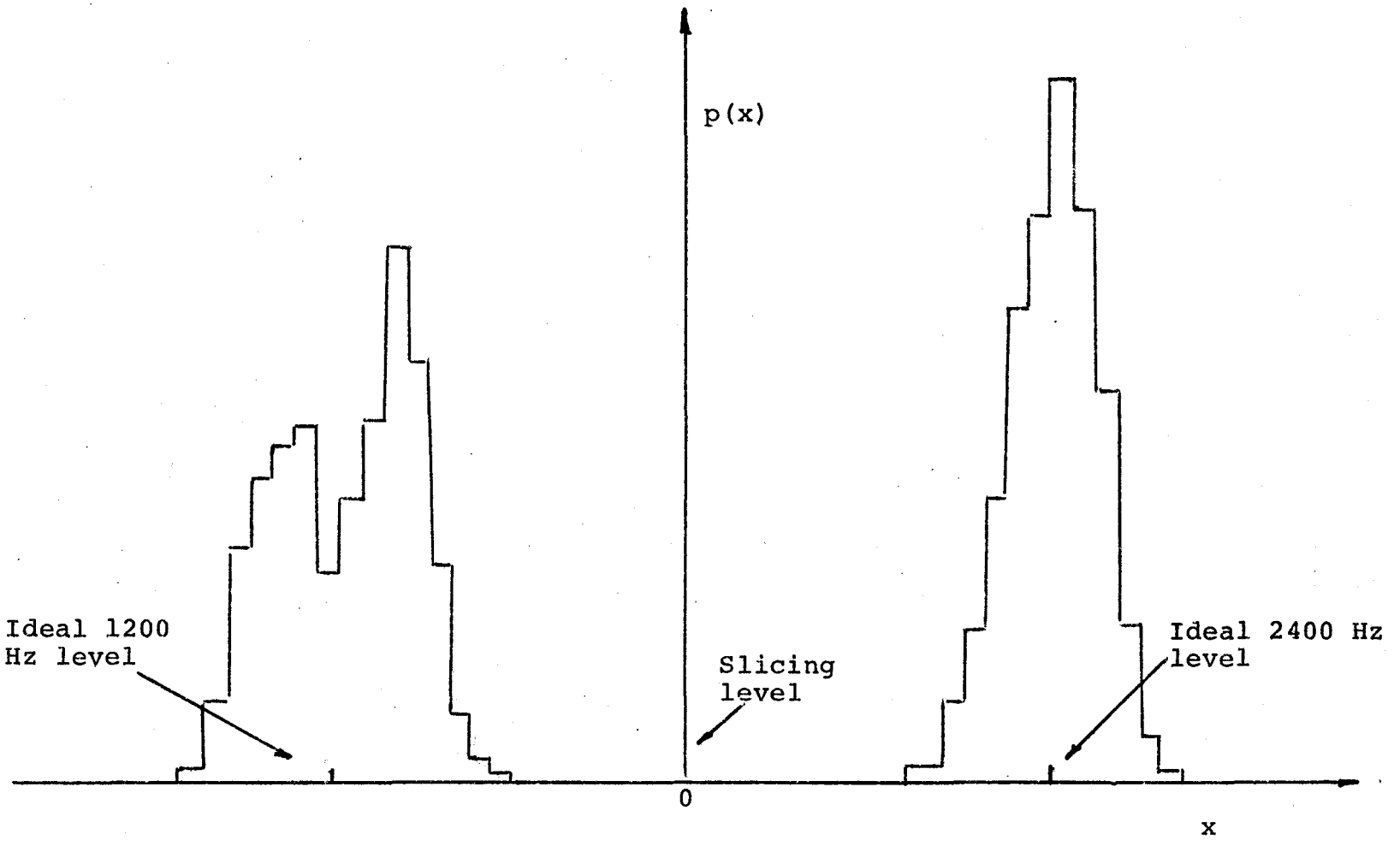


fig.9.13(b) Probablity Density of Sampled Baseband Output after Equalization (Channel 1)

gain of the equalizer at passband did change during the equalization process. Similar results were obtained for all the channels studied. Note that the decrease in the MSE is reflected in the decreased variance about the ideal received level.

(b) Channel - Slight Distortion

As can be seen from figure (9.14), the characteristics of this particular channel are quite good. There is only a 6 dB variation in the amplitude-frequency characteristic across the passband, and the phase-frequency response is quite linear. This is reflected in the unequalized BPE which from figure (9.17) is two orders of magnitude less than that of channel 1 at 16 dB. It is apparent also that the improvement that the equalizer achieves is smaller, which is to be expected since the channel was only slightly distorted to begin with. The convergence times are also somewhat decreased because the tap gains did not have to move as far in this case. Note that the residual MSE is less than the previous case indicating that the channel memory time lies essentially within the structure of the equalizer. The BPE will be much greater than 10^{-6} for the worst case SNR expected (25 dB flat, 28 dB corrected).

(c) Channel 3 - Severe Distortion

This particular channel is quite severe in nature as can be observed from the characteristic shown in figure (9.18). At 1200 Hz, the amplitude characteristic is down

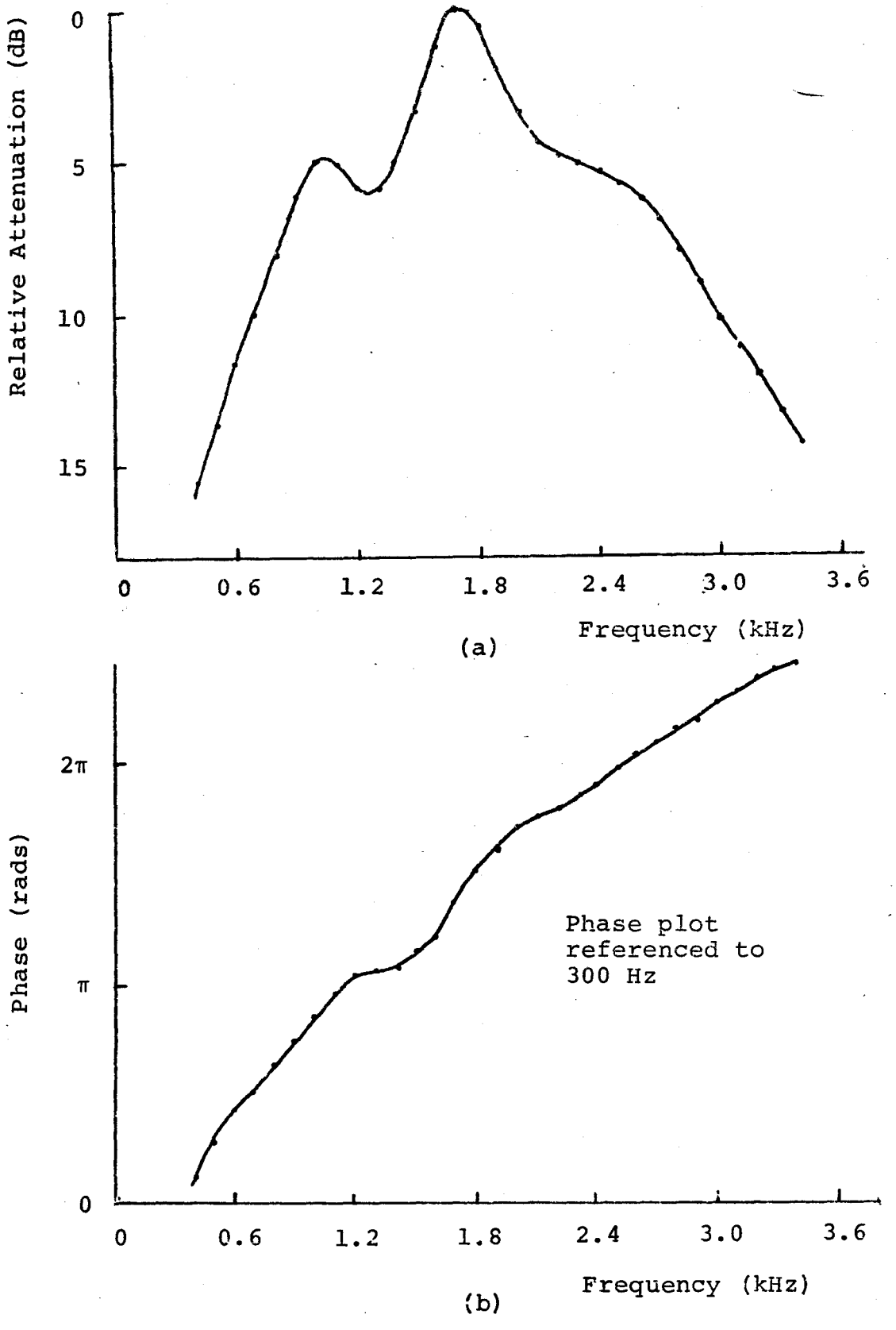


fig.9.14 Channel 2. (a) Attenuation Characteristic (b) Phase Characteristic

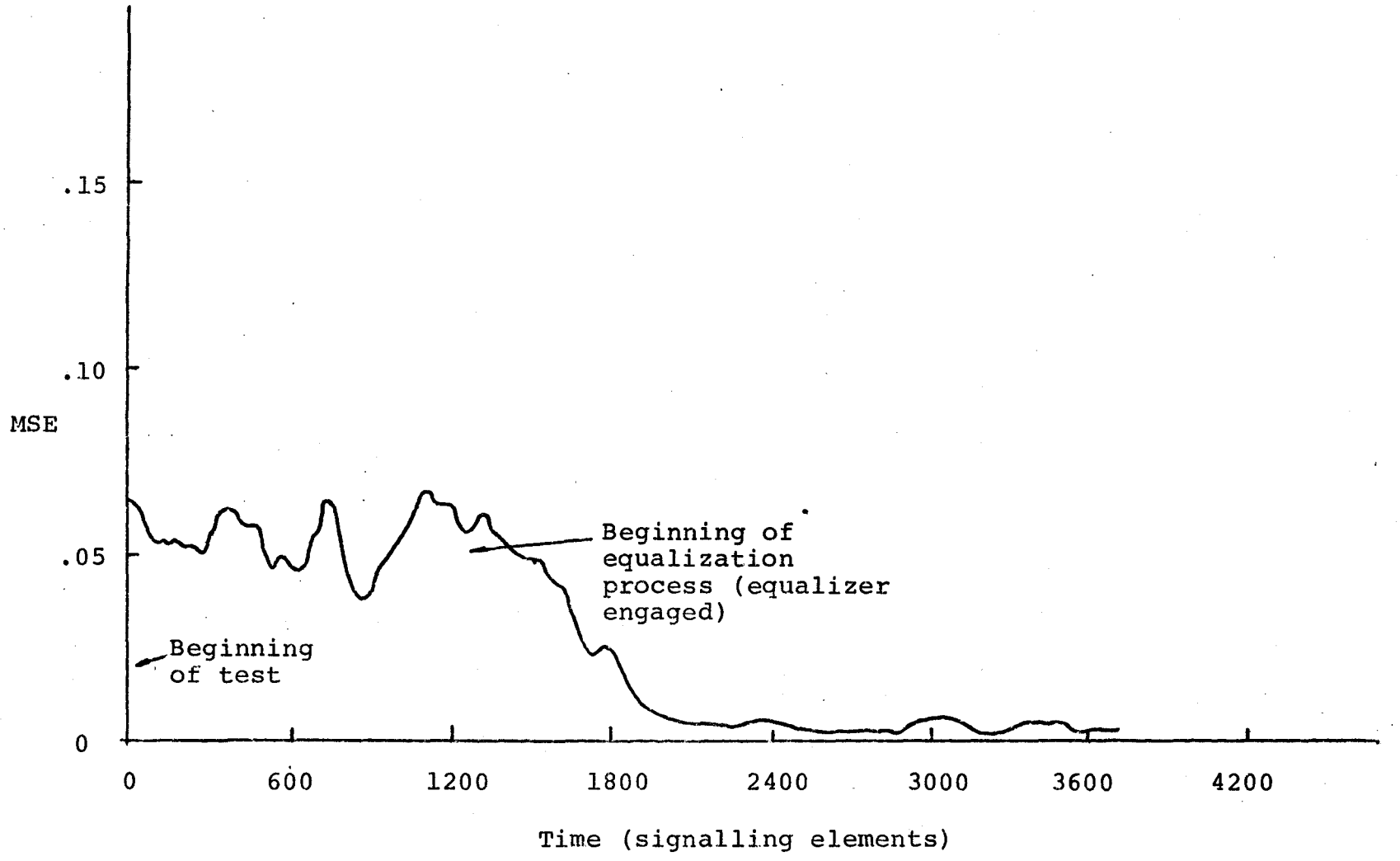


fig.9.15(a) Typical Convergence Curve at 28 dB (Channel 2)

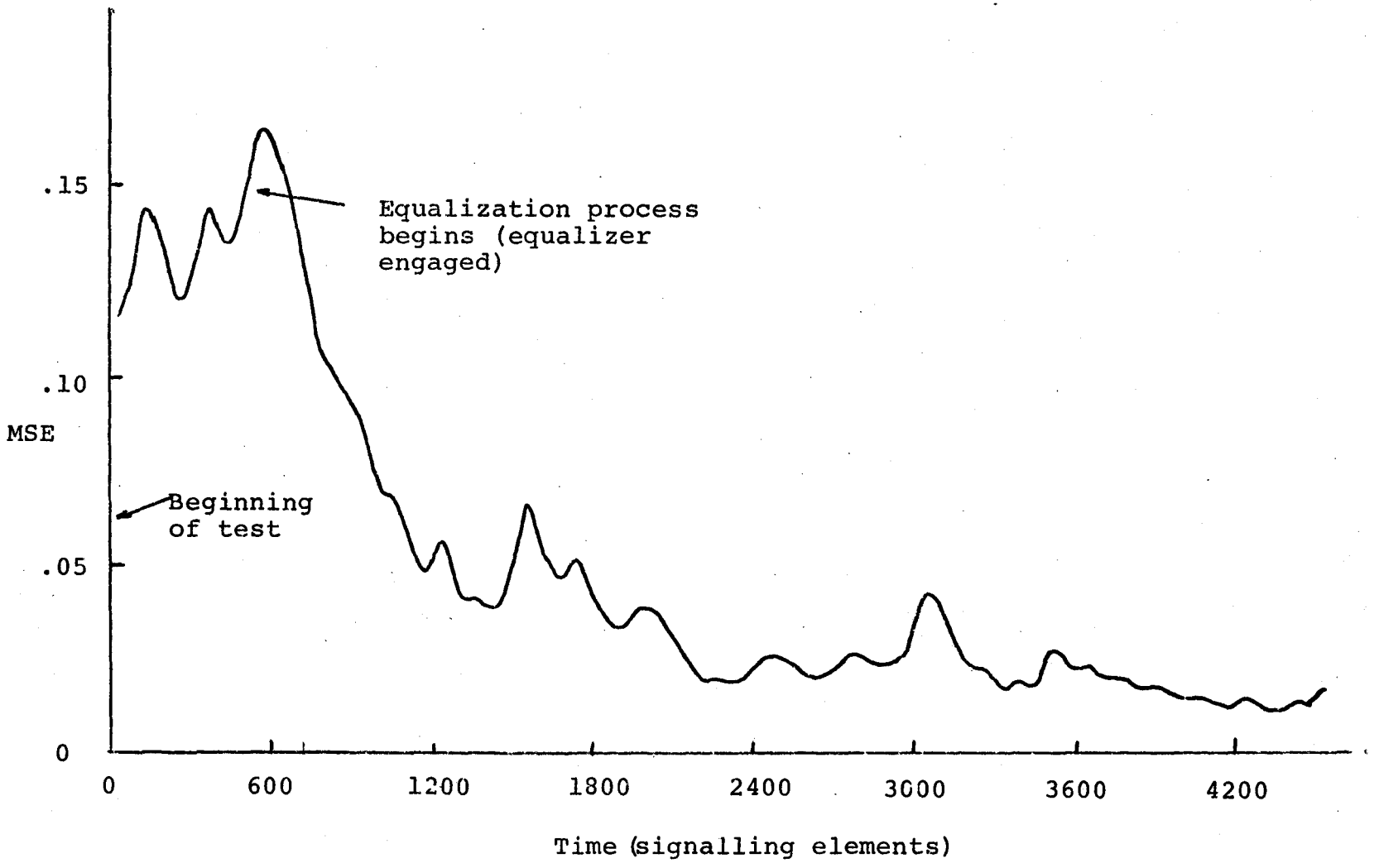


fig.9.15(b) Typical Convergence Curve at 11 dB (Channel 2)

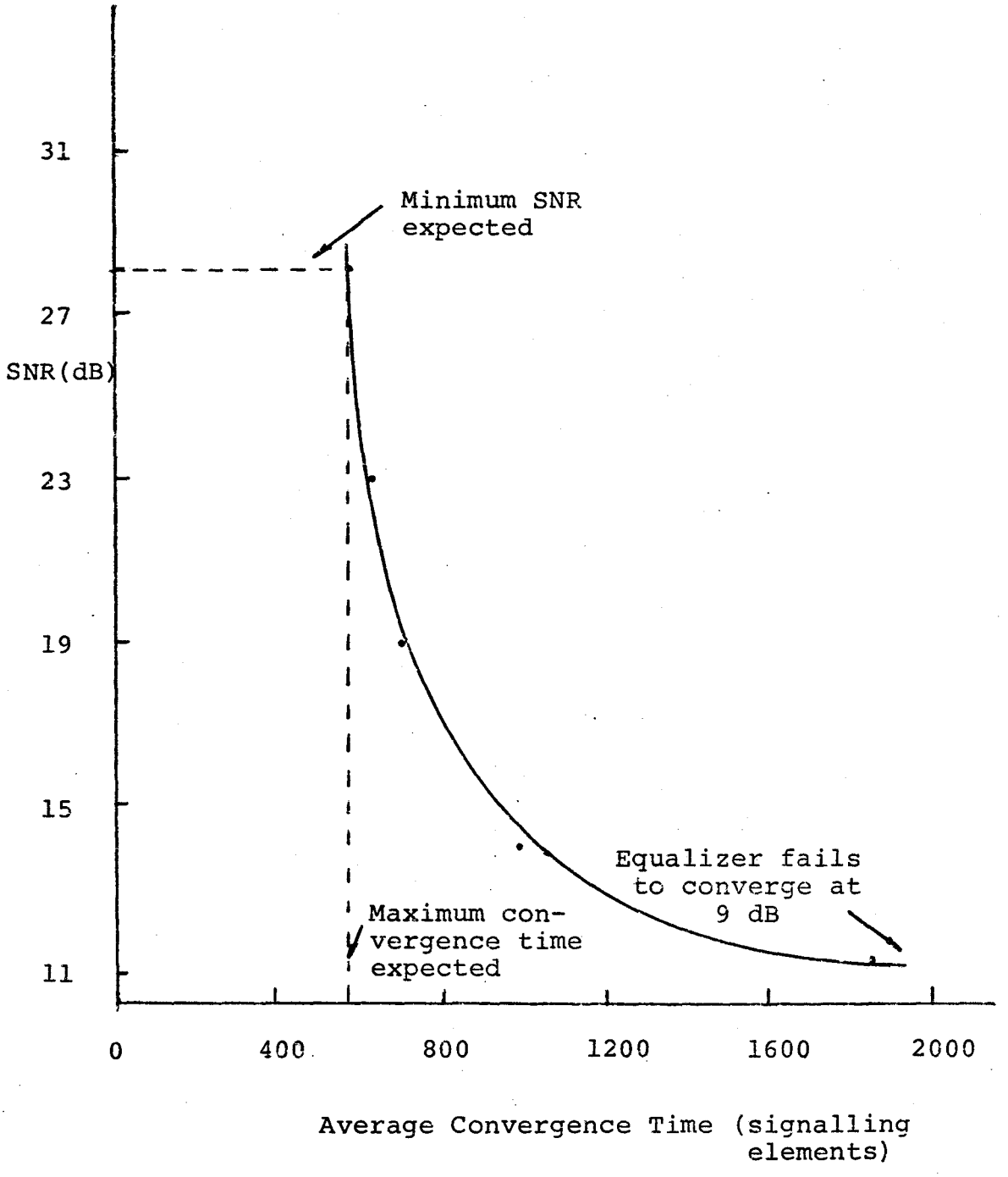


fig.9.16 Convergence Time for Varying SNR

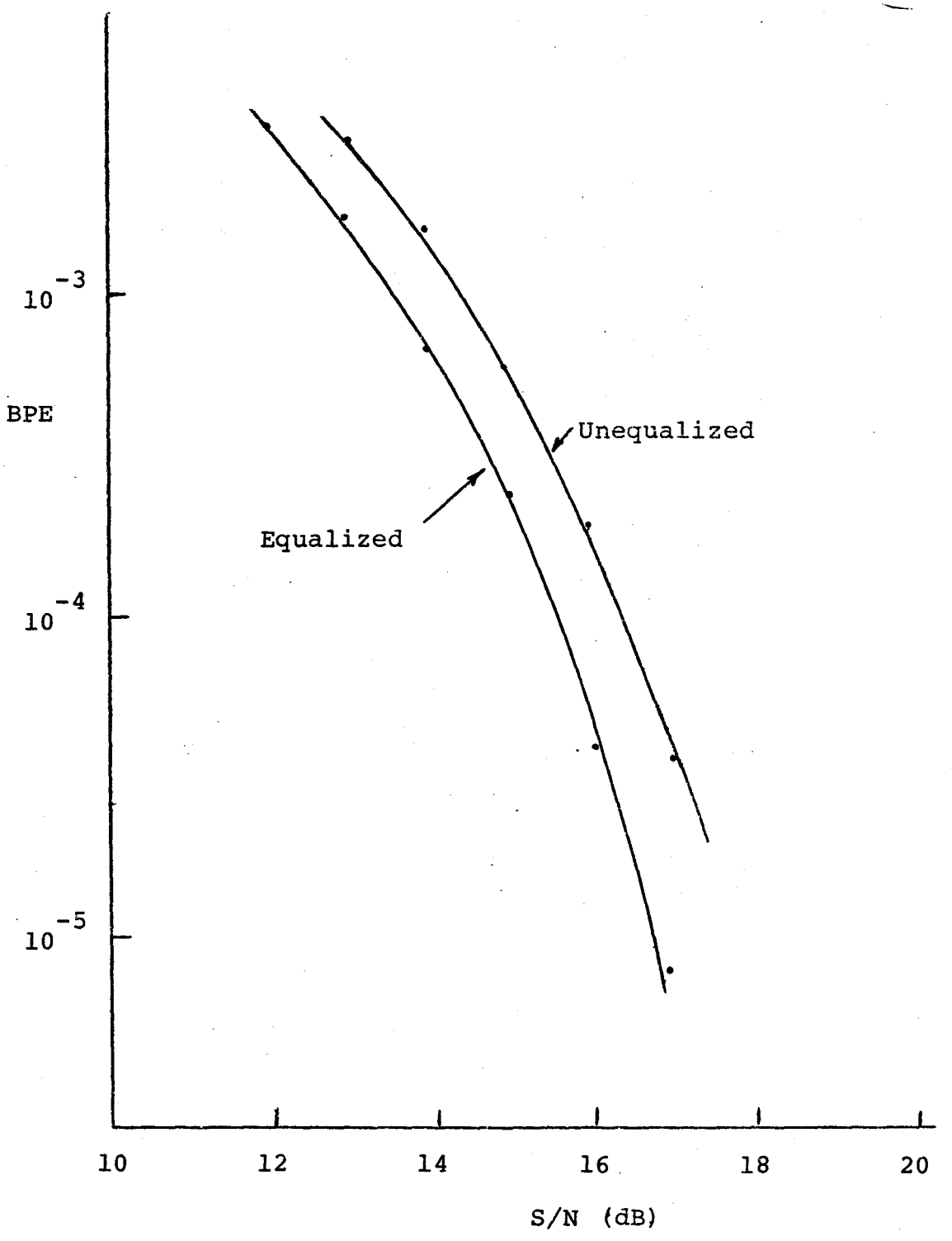


fig.9.17 BPE Characteristics (Channel 2)

approximately 12 dB from that at 1900 Hz. This response would result in poor voice transmission, and for this reason would not likely occur in practice. The peak distortion (D) observed was greater than one. This is apparent from figure (9.12), where it can be observed that the BPE is finite even in the absence of noise due to the severe ISI. It is also apparent from this latter graph that the equalizer cannot acieve significant correction below a SNR of 28 dB. This occurred because the severe band-limiting increased the channel impulse response to the point where it was outside the structure of the equalizer, which was only 7T seconds long. As a result, significant ISI remained along with a correspondingly high error rate. However, it is apparent that the adaptive structure would provide reasonably satisfactory performance (BPE $< 2 \cdot 10^{-5}$) for SNR's greater than 31 dB. At this latter SNR, the BPE of the adaptive structure is approximately 100 times better than that for the non-adaptive.

9-11 Eye Pattern

Perhaps one of the more impressive features of the equalization process is the improvement in the eye pattern. As can be seen from figure (9.22) the "eye" is much wider after equalization, (the channel employed had characteristics similar to those of channel 1). In contrast to the earlier described case where the equalizer structure was implemented at baseband, the increased eye opening here is

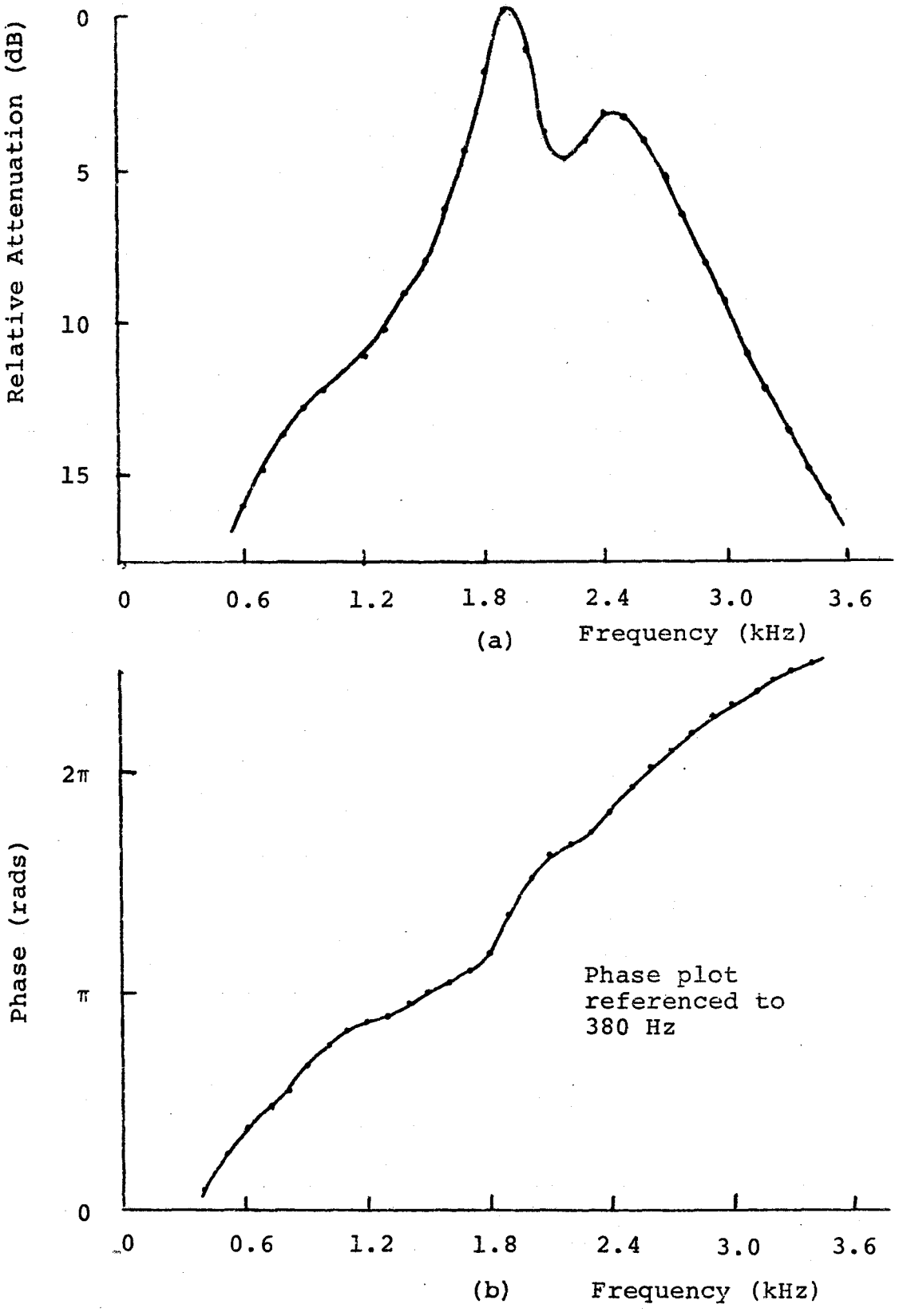


fig.9.18 Channel 3. (a) Attenuation Characteristic (b) Phase Characteristic

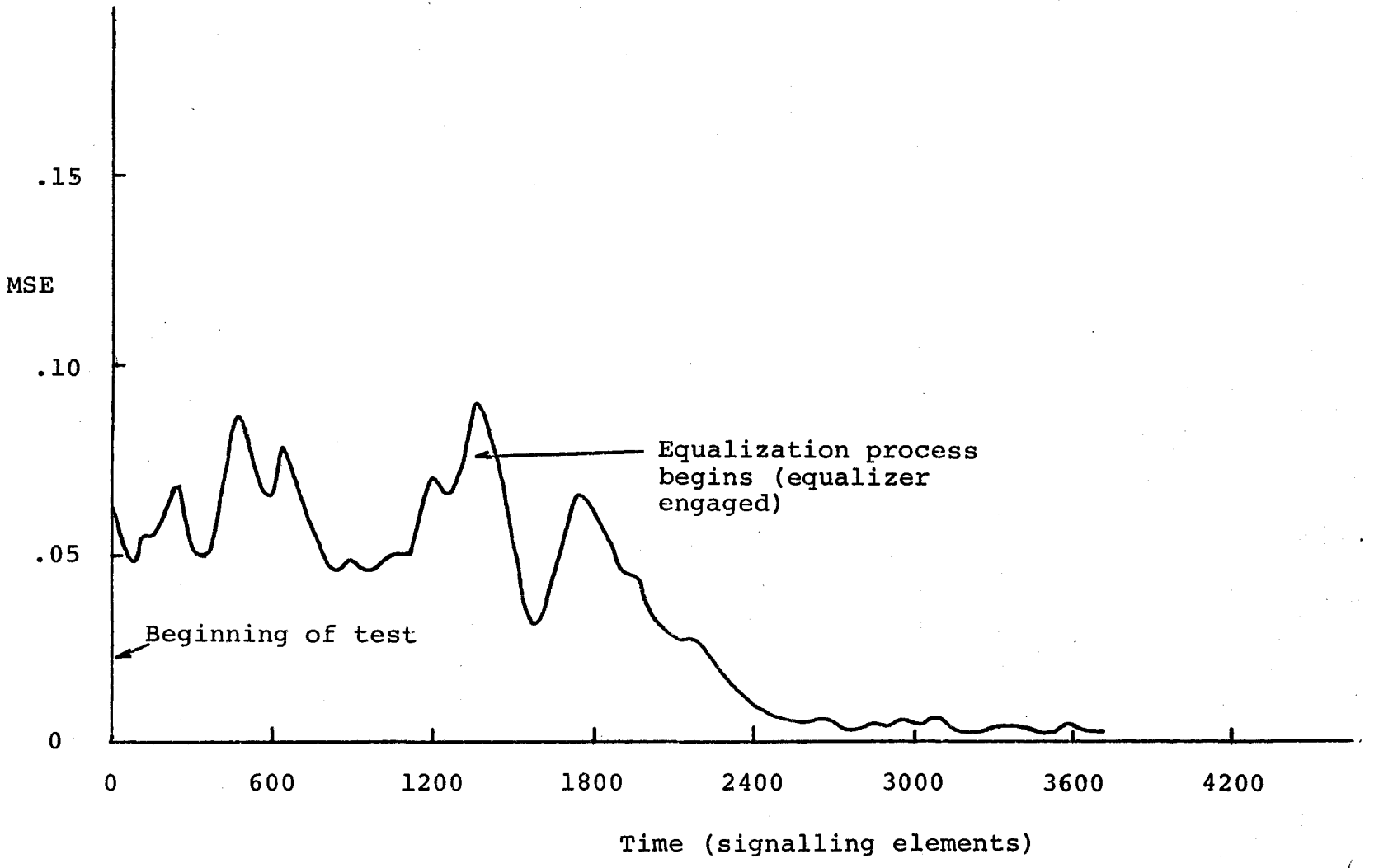


fig.9.19(a) Typical Convergence Curve at 28 dB (Channel 3)

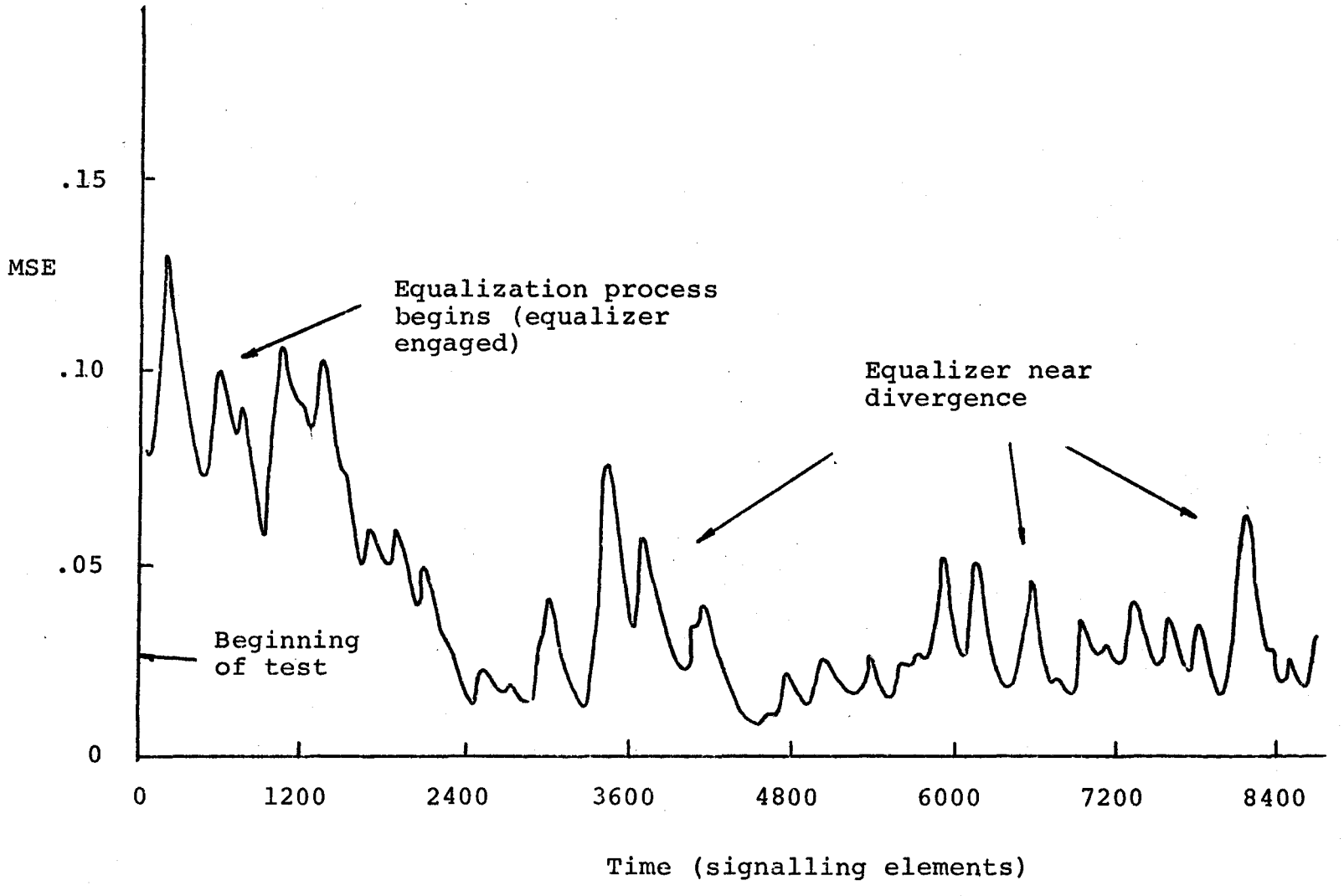


fig.9.19(b) Typical Convergence Curve at 14 dB (Channel 3)

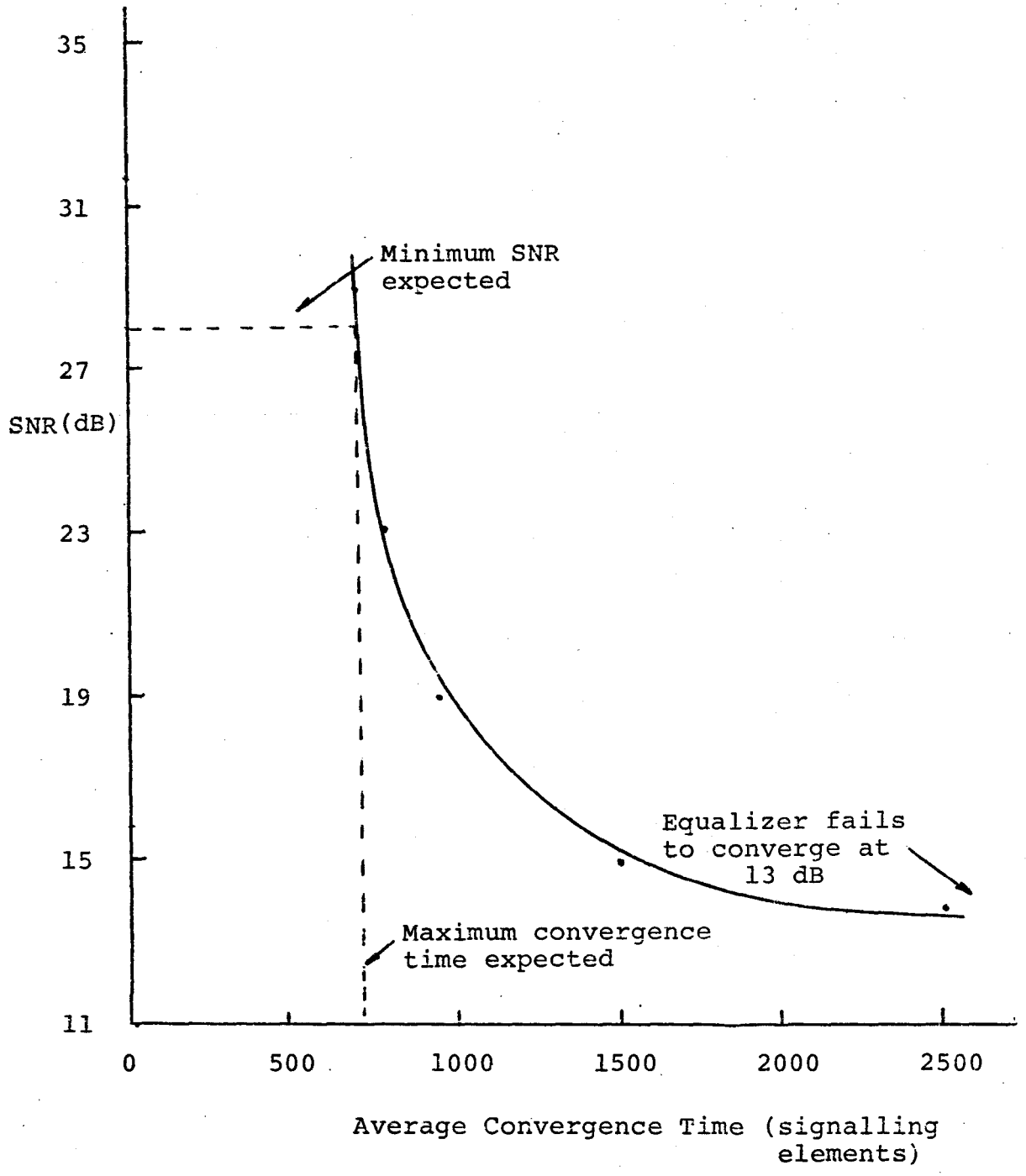


fig.9.20 Convergence Time for Varying SNR

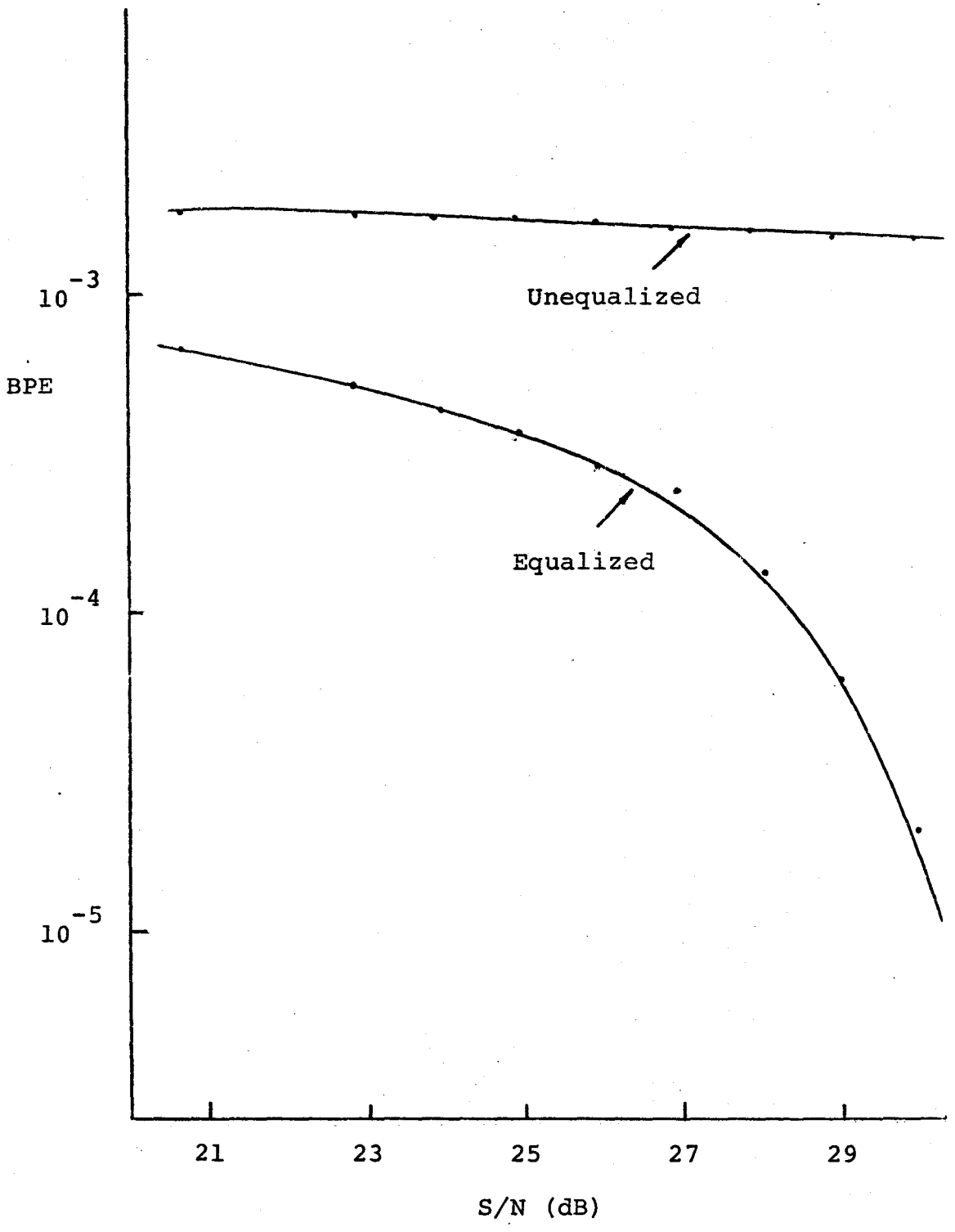
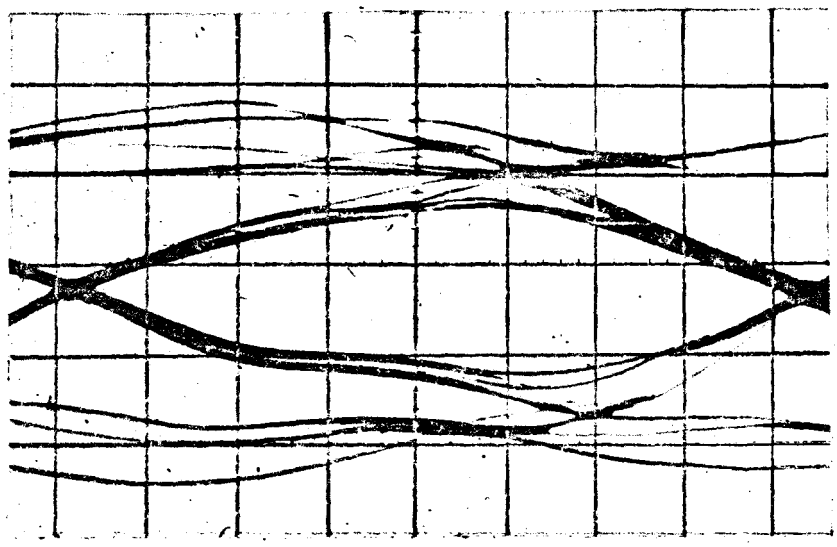
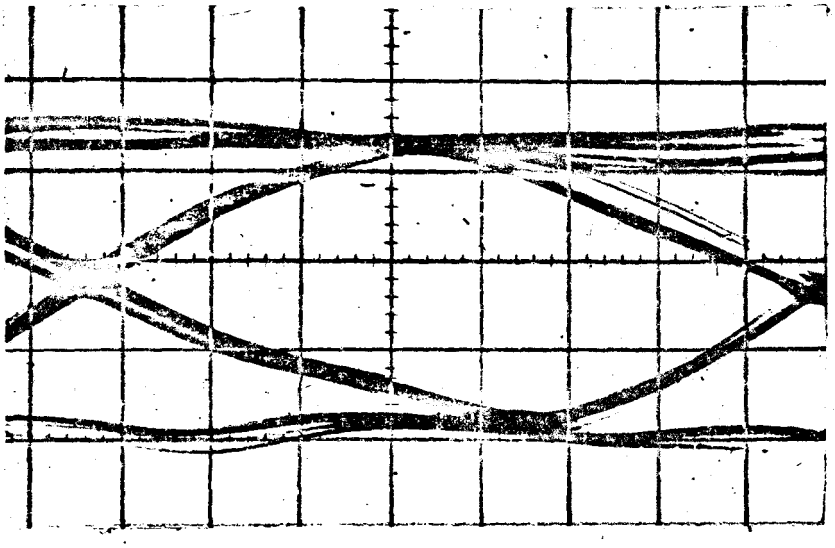


fig.9.21 BPE Characteristics (Channel 3)



Hor. - .1 msec/div.

(a)



Hor. - .1 msec/div.

(b)

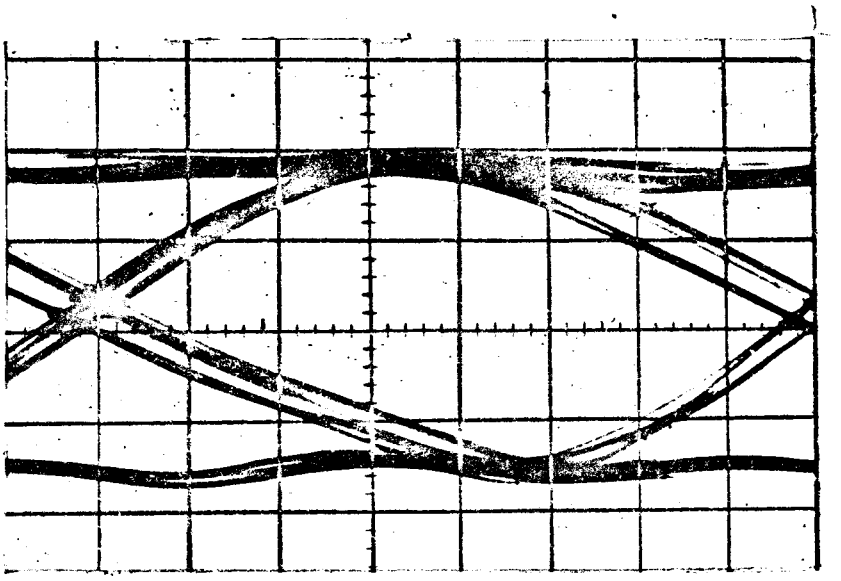
fig.9.22 Eye Pattern Characteristic.(a) Before Equalization. (b) After Equalization

directly related to the reduction in error rate.

The sequence of photographs illustrated in figure (9.23) demonstrates the holding properties of the equalizer. It is apparent (and was observed from the BPE) that the equalizer does not drift significantly during the first 15 seconds or so. The holding was achieved by effectively removing the input to the integrators of the tap gains. This relatively long time (18.000 bauds) can be used to advantage if an interruption is encountered during transmission. A simple envelope detector could be employed to hold the taps in their previous settings if loss of carrier is incurred. Thus short interruptions could be bridged.

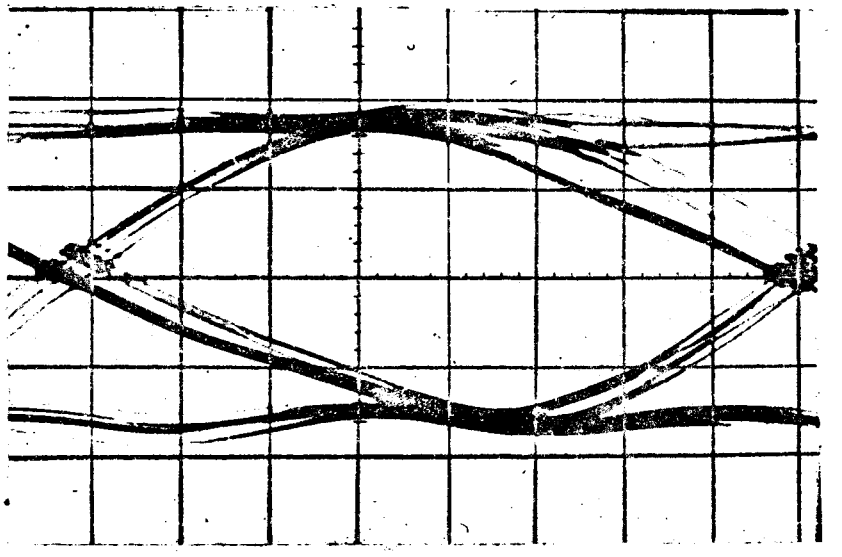
9.12 General Observations - Convergence Properties

It was found that the equalizer converged to settings which reduced the BPE in the absence of noise for all channels considered. However, it was observed that if the equalizer diverged considerably from the settings which reduced the BPE, that reconvergence could not be guaranteed. This striking event could only be brought about by decreasing the SNR to abnormally low values (in the order of 8 dB uncorrected), and then removing the noise. At this point it was observed that in most cases reconvergence did not take place. Although SNR's this poor would not likely be encountered in the voice network (since then the channel would be virtually useless for any kind of information



Hor. - .1 msec/div

(a)



Hor. - .1msec/div

(b)

fig.9.23 Eye Pattern during Holding. (a) After 10 Seconds. (b) After 20 seconds.

transfer) this characteristic would still in certain instances tend to limit the application of the structure. This divergence property also implies that the error surface encountered in the equalization process is not quadratic (in contrast to the linear MSE minimizing equalizer) but rather contains local minima which do not necessarily correspond to reducing the BPE. It is thought that this effect is directly related to the fact that the PLL is essentially amplitude insensitive and that the algorithm is decision-directed.

Another important observation was noted about the convergence properties for the case of non-random data. It was found that the equalizer did diverge as expected for a continuous 1 or 0 pattern - however, it did retain the proper settings if reversals (1010..) or other simple patterns were sent. This result implies that the scrambler, which as previously described would be required in practice, can be very simple in nature.

9-13 Clock Frequency Offset

Since the delay line per tap and the center frequency of the DPLL were both determined by crystal-controlled oscillators, tests were performed to determine the maximum allowable frequency offset in the transmitted signalling rate (1/T) which could be tolerated without significant degradation. A severe restriction on the tolerances created by too great a sensitivity on this parameter would make

implementation of the modem impractical. By adjusting the frequency of the transmit clock it was found that the system degradation was not significant as long as the offset was not more than a few cycles per second (±.15% of the nominal bit rate approximately). This does in fact represent quite a large tolerance and should not create any significant design or implementation problems.

During the course of the experiments performed, the transmit clock was maintained within ∿2 Hz of the DPLL idle frequency.

CHAPTER TEN

AREAS FOR FUTURE INVESTIGATION

10-1 General

The purpose of this thesis has been to implement and evaluate an adaptive FSK modem. The success of the project is evident from the last chapter, - however, much research remains to be performed on certain aspects of the system.

10-2 Channel Simulation

The channel model employed included only the effects of band-limited white noise and dispersion, as a result, further research is required to determine exactly the degradation created by second order effects, such as frequency offset and phase jitter. In addition, the channel model employed did not have the severe roll-off response which is characteristic of some FDM channels. This may result in an increase in the length of the tapped delay line required to compensate for the increased length of the channel impulse response.

Another area which required investigation is that concerning the effects of impulse noise on the system, since as previously indicated in chapter two, this form of noise is predominant in the voice network and not the gaussian noise employed in the channel model.

10-3 The Sampled Bandpass Limiter

The sampling rate of 100/T was chosen somewhat arbi-

trarily in the bandpass limiter. A reduction of the sampling rate would make far a more economical system, since the number of flip-flops required to provide the required delay time (T) would be reduced. However, to do so, the exact relationship between the sampling rate and the degradation encountered would have to be determined.

10-4 Filtering

The filters employed in the modem were simple in nature and although the receiver is only 1.6 db from the optimum, it may be feasible to provide more sophisticated filtering in order to achieve an even higher efficiency.

10-5 The Algorithm

In polled data collection systems, it is often desired to obtain a very short turn-around time. Thus, the adaption time may become a critical factor in this type of application, and as a result, it may become necessary to reduce it. This possibly can be achieved by implementing some of the techniques presently being employed in linear equalizers (for example, the "hybrid technique", where the error signal is not quantized, but rather is used directly in the cross-correlators).

In addition, a further investigation to determine the exact nature of the error surface encountered in the equalization process, would provide invaluable information, and could lead to the elimination of the divergence problem des-

cribed earlier

10-6 PSK Systems

Perhaps one of the most interesting and fruitful investigations which could be undertaken, is that of employing the equalizer structure developed to PSK systems. The PSK system is attractive because like FSK, PSK systems usually employ a limiter in the front end. Consequently, the technique of limiting, sampling, and using MOS flip-flops as delay elements can once again be employed. Furthermore, PSK systems offer the additional advantage of being only slightly less efficient than their linear counterparts. This increase in efficiency may allow the equalizer to be used at relatively high transmission rates.

CONCLUSIONS

The implementation of an adaptive FSK modem employing a tapped delay line equalizer has proven successful. Many adaptive systems have been suggested and implemented by other authors, but in contrast to this system, the techniques employed have generally been restricted to linear modulation methods - furthermore these systems tend to be not only complex but also costly to realize.

The system developed is somewhat unique in that it makes use of the limiter generally employed in FSK systems to its advantage in the realization of the tap delay elemnts and the variable tap gains. This results in a simple economical modem which can offer considerable improvement in performance for low speed transmission on the existing voice network.

The only disadvantage of the system of significance is the non-quadratic error surface encountered in the equalization and tracking process. However as previously mentioned this effect should not prove to be a limitation because of the nature of the channels encountered in the voice network.

THE APPENDICES

APPENDIX A

VOICE CHANNEL CHARACTERISTICS

This appendix is a collection of some of the more important parameters of the channels to be expected on the voice network.

1) Worst Case Attenuation Characteristics (reference to a 1000 Hz test tone)

Frequency (Hz)	Variation (dB)
300 - 3000	-3 to +12
500 - 2500	-2 to +8

2) Worst Case Delay Distortion (derivative of the phase frequency characteristic)

   Usually less than 1.8 msec over the band from 800 to 2600 Hz.

3) Absolute Delay - Cannot be specified as it depends on the distance between connections as well as the type of FDM equipment employed.

4) Frequency Offset - Normally within $\pm$ 5 Hz.

5) Phase Jitter - Usually within $\pm$ 10 degrees with a jitter rate in the order of 20 Hz.

6) Background Noise - Usually 25 dB or more below signal level.

7) Impulse Noise - Cannot usually be specified.

8) Harmonic Distortion - Normally 30 dB or more below sgnal level.

APPENDIX B

MODEM CIRCUITRY

This appendix contains the circuitry of the modem on a card by card basis.

Figure	Card	Contents
B-1	1	a) Digital modulator b) Pattern generator c) Random data generator d) Transmitter clock e) Receiving band-pass filter and limiter
B-2	2	a) Transmitting BTF b) R-C Low-pass filter
B-3	3	a) Receiver clock pulse generator b) EIA interfacing
B-4	4	a) PLL (discriminator) b) Zero-crossing detector c) ZCDD slicer
B-5	5	a) Digital PLL (for synchronization)

Figure	Card	Contents
B-6	6	a) FSK delay line(binary) b) Part of input symbol polarity delay line
B-7	7&8	a) Tap gain Controllers b) Summing bus
B-8	9	a) Low-pass filter for ZCDD b) Sign of output error and digit slicers c) Remaining portion of input symbol polarity delay line d) Preset voltage for all taps (except center tap)

Notes:

1) All logic elements are Signetics "Utilogic" - unless otherwise noted
2) All resistors are $\pm$ 5% unless noted otherwise
3) All capacitors are $\pm$ 5% unless noted otherwise
4) Unused circuitry is not shown
5) Decoupling capacitors are not shown unless important
6) S1 and S2 are 11 pole switches with rotors connected to +5 v supply (through 1 kΩ).

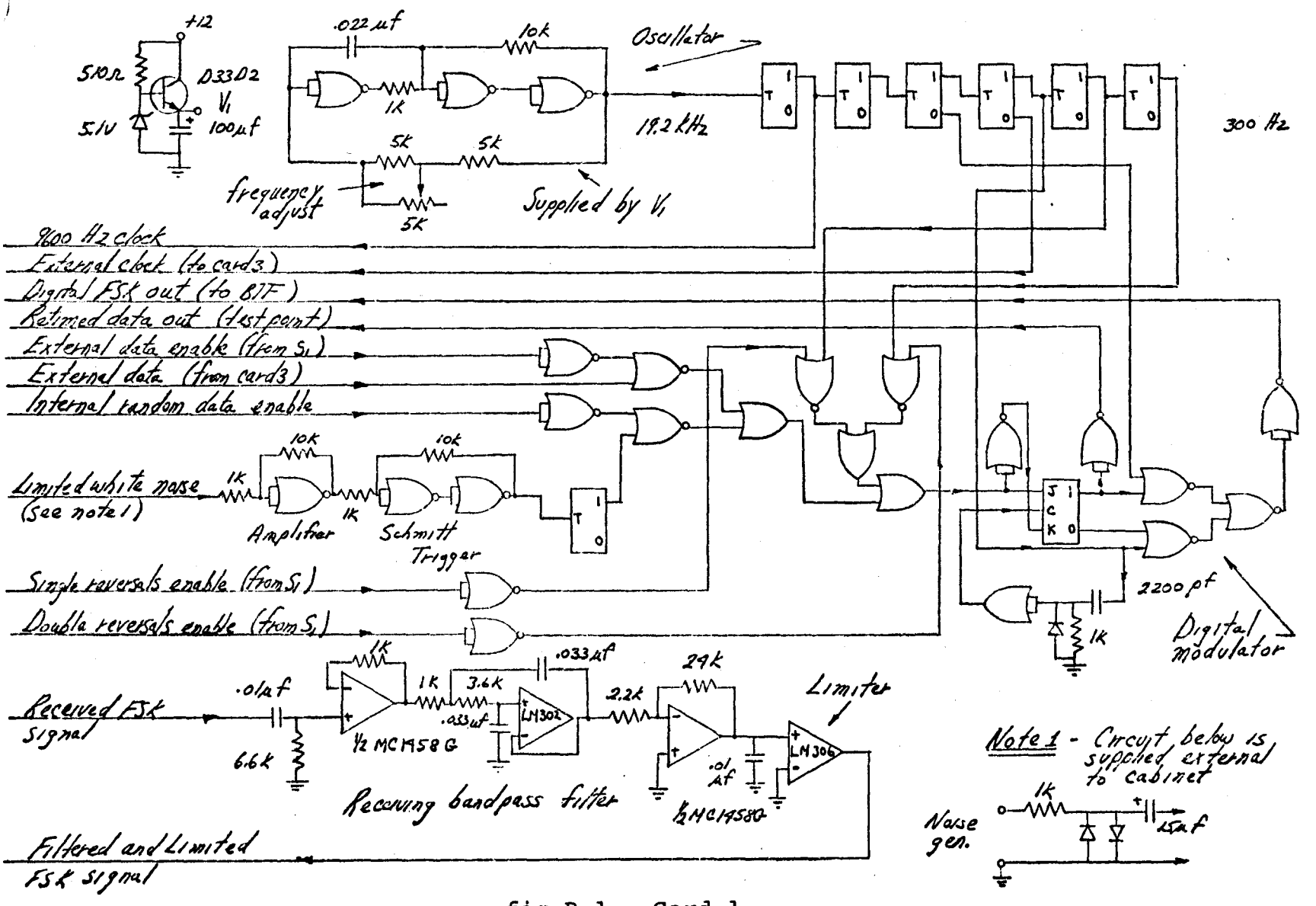

+12
510Ω
D33D2
V1
5.1V
100μf
.022μf
10k
1K
5K
5K
5K
frequency adjust
Supplied by V1
Oscillator
19.2 kHz
300 Hz
9600 Hz clock
External clock (to card 3)
Digital FSK out (to BTF)
Retimed data out (test point)
External data enable (from S1)
External data (from card 3)
Internal random data enable
Limited white noise (see note 1)
10K
1K
10K
1K
Amplifier
Schmitt Trigger
Single reversals enable (from S1)
Double reversals enable (from S1)
2200 pf
1K
Digital modulator
Received FSK signal
.01μf
6.6K
1K
1K
3.6K
.033μf
.033μf
½ MC1458G
LM302
2.2k
29k
.01 Mf
LM306
Limiter
Receiving bandpass filter
½MC1458G
Filtered and Limited FSK signal
Note 1 - Circuit below is supplied external to cabinet
1K
4.7μf
Noise gen.

fig.B-1 Card 1

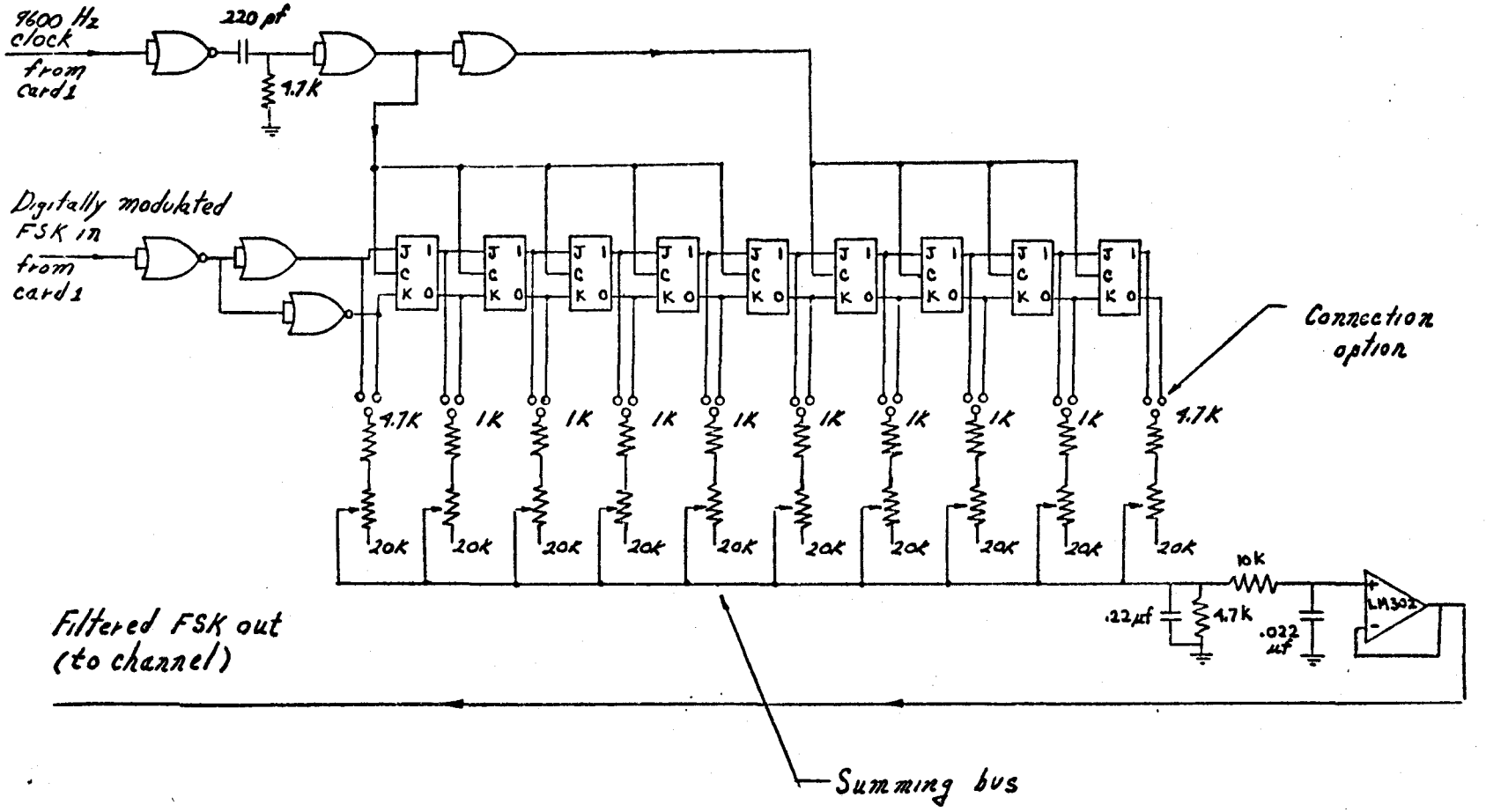

9600 Hz clock from card 1
220 pf
4.7K
Digitally modulated FSK in from card 1
J
C
K
Connection option
4.7K
1K
20K
10K
.22µf
4.7K
.022 µf
LM302
Filtered FSK out (to channel)
Summing bus

fig.B-2 Card 2

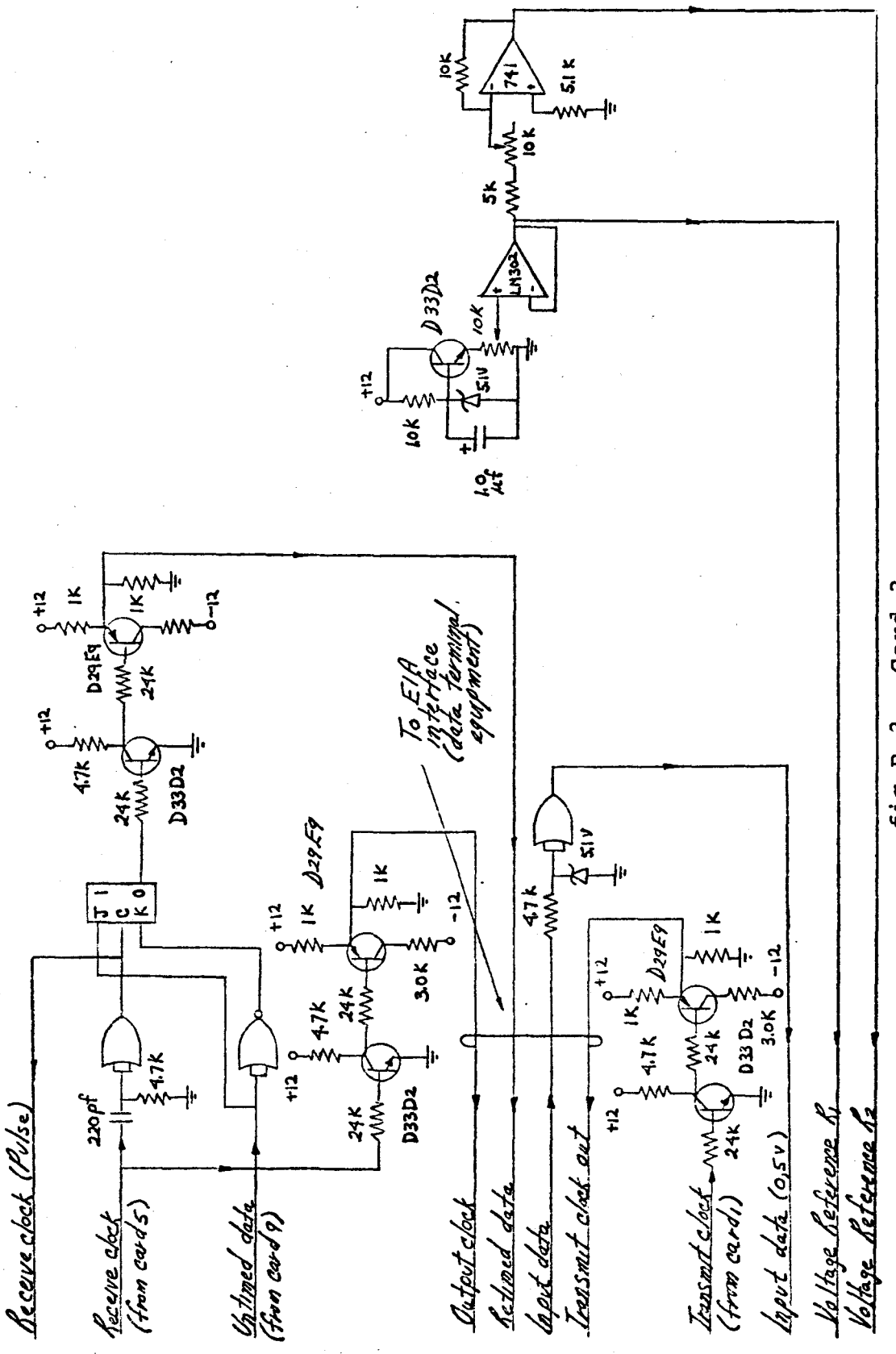

Receive clock (Pulse)
Receive clock (from card 5)
Untimed data (from card 9)
Output clock
Retimed data
Input data
Transmit clock out
Transmit clock (from card 1)
Input data (0,5V)
Voltage Reference #1
Voltage Reference #2
To EIA interface (data terminal equipment)
220pf
4.7k
J
C
K
Q
1K
24K
D33D2
D29F9
3.0K
-12
+12
5V
10K
5.1K
741
LM301
5K
10K
4.0 µf

fig.B-3 Card 3

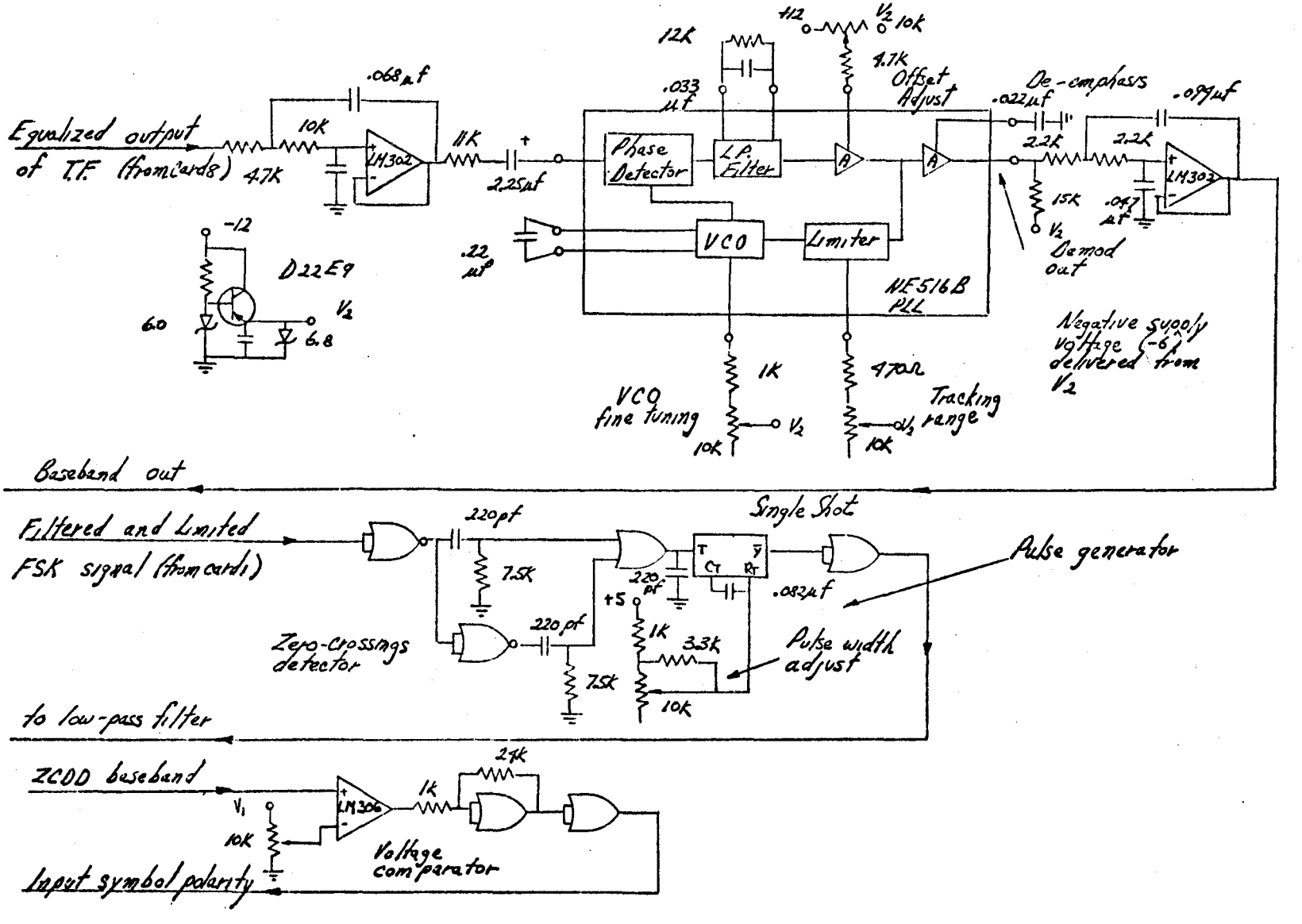

Equalized output of I.F. (from card 8)
4.7K
10K
.068 μf
LM302
11K
2.25μf
Phase Detector
L.P. Filter
12K
.033 μf
+12
V2 10K
9.1K
Offset Adjust
De-emphasis
.022μf
.082μf
2.2K
2.2K
LM303
15K
.047 μf
V2
Demod out
VCO
Limiter
.22 μf
NE516B PLL
-12
D22E9
6.0
6.8
V2
1K
470Ω
VCO fine tuning
10K
V2
Tracking range
10K
Negative supply voltage (-6) delivered from V2
Baseband out
Filtered and Limited FSK signal (from card 1)
220pf
7.5K
220 pf
Single Shot
.082μf
Pulse generator
+5
1K
3.3K
Pulse width adjust
10K
Zero-crossings detector
220 pf
7.5K
to low-pass filter
ZCDD baseband
2.7K
1K
LM306
V1
10K
Voltage comparator
Input symbol polarity

fig.B-4 Card 4

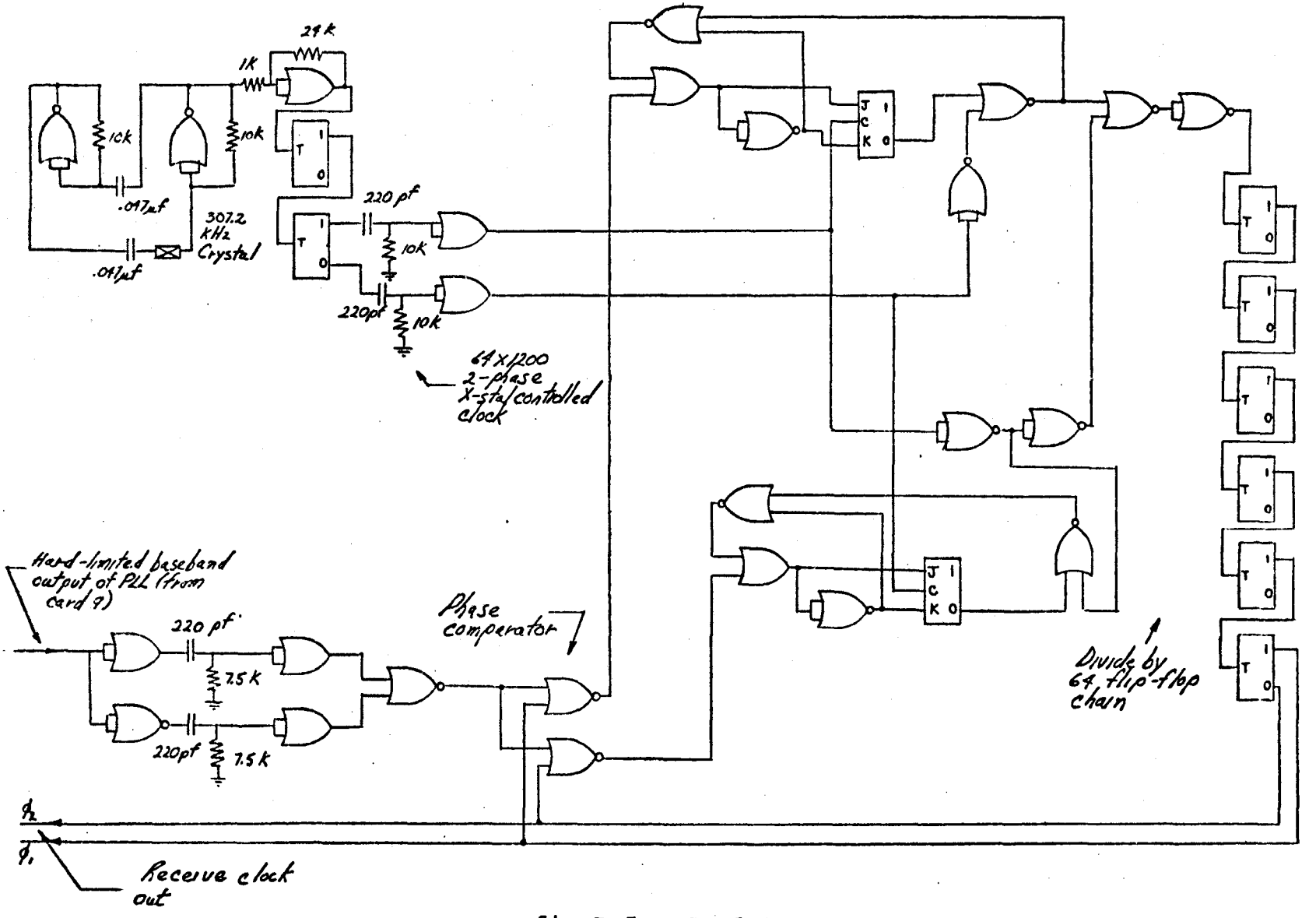

29k
1K
10k
10k
.047μf
.047μf
307.2 kHz Crystal
220 pf
10K
220pf
10k
64x1200 2-phase X-stal controlled clock
Hard-limited baseband output of PLL (from card 9)
220 pf
7.5 k
220pf
7.5 k
Phase comperator
Divide by 64 flip-flop chain
Receive clock out

fig.B-5 Card 5

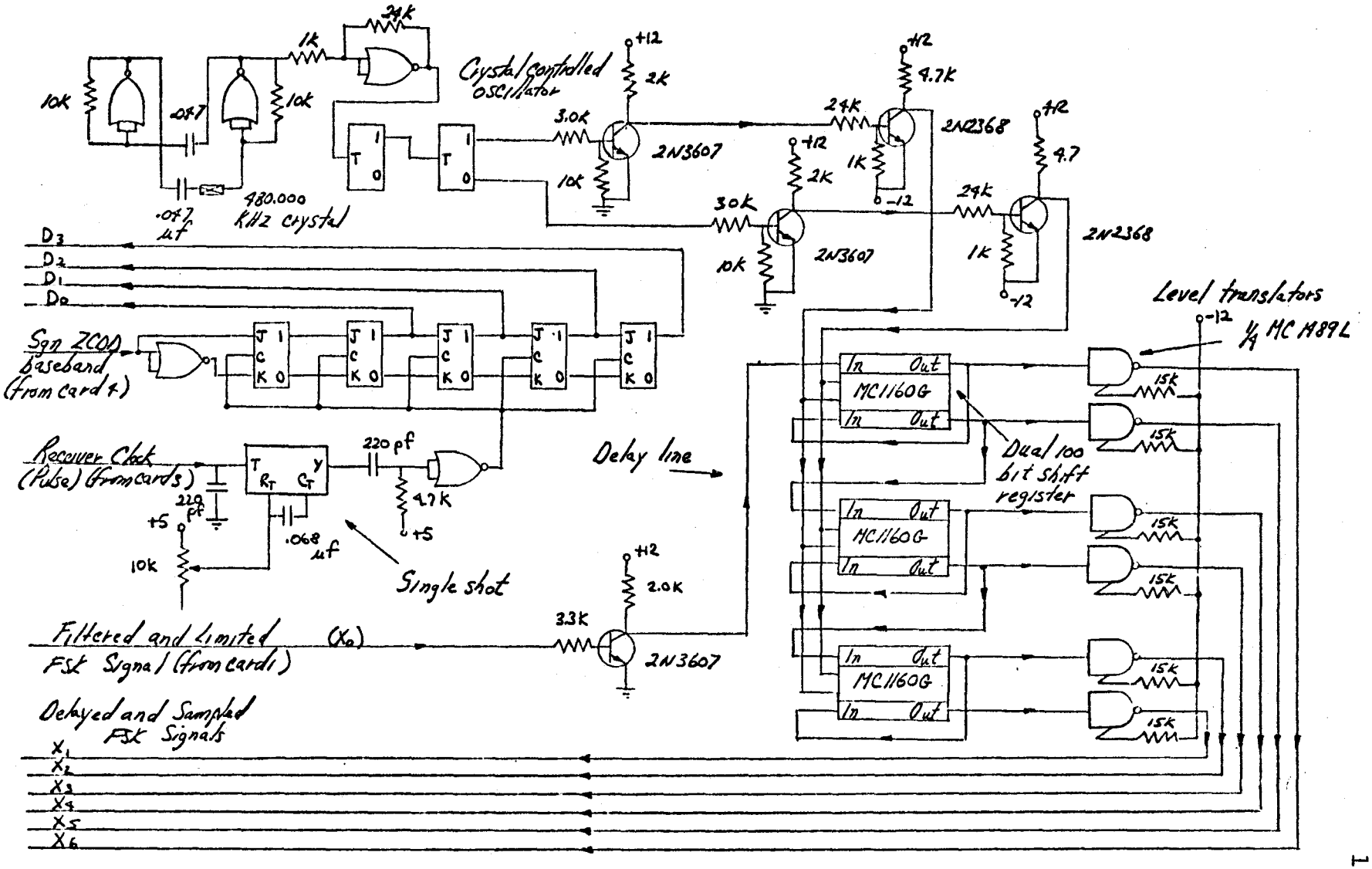

24k
1k
10k
.047
10k
Crystal controlled oscillator
+12
2k
3.0k
2N3607
10k
480.000 KHz crystal
.047 µf
+12
4.7k
24k
2N2368
1k
-12
+12
2k
30k
2N3607
10k
+12
4.7
24k
2N2368
1k
-12
D3
D2
D1
D0
Sgn ZCOR baseband (From card 4)
Level translators
-12
1/4 MC 1489L
In Out
MC1160G
In Out
15k
Dual 100 bit shift register
Receiver Clock (Pulse) (From card 3)
220 pf
220 pf
+5
10k
.068 µf
Single shot
Delay line
4.7K
+5
+12
2.0K
3.3K
2N3607
Filtered and Limited FSK Signal (From card 1)
(X0)
Delayed and Sampled FSK Signals
X1
X2
X3
X4
X5
X6

fig.B-6 Card 6

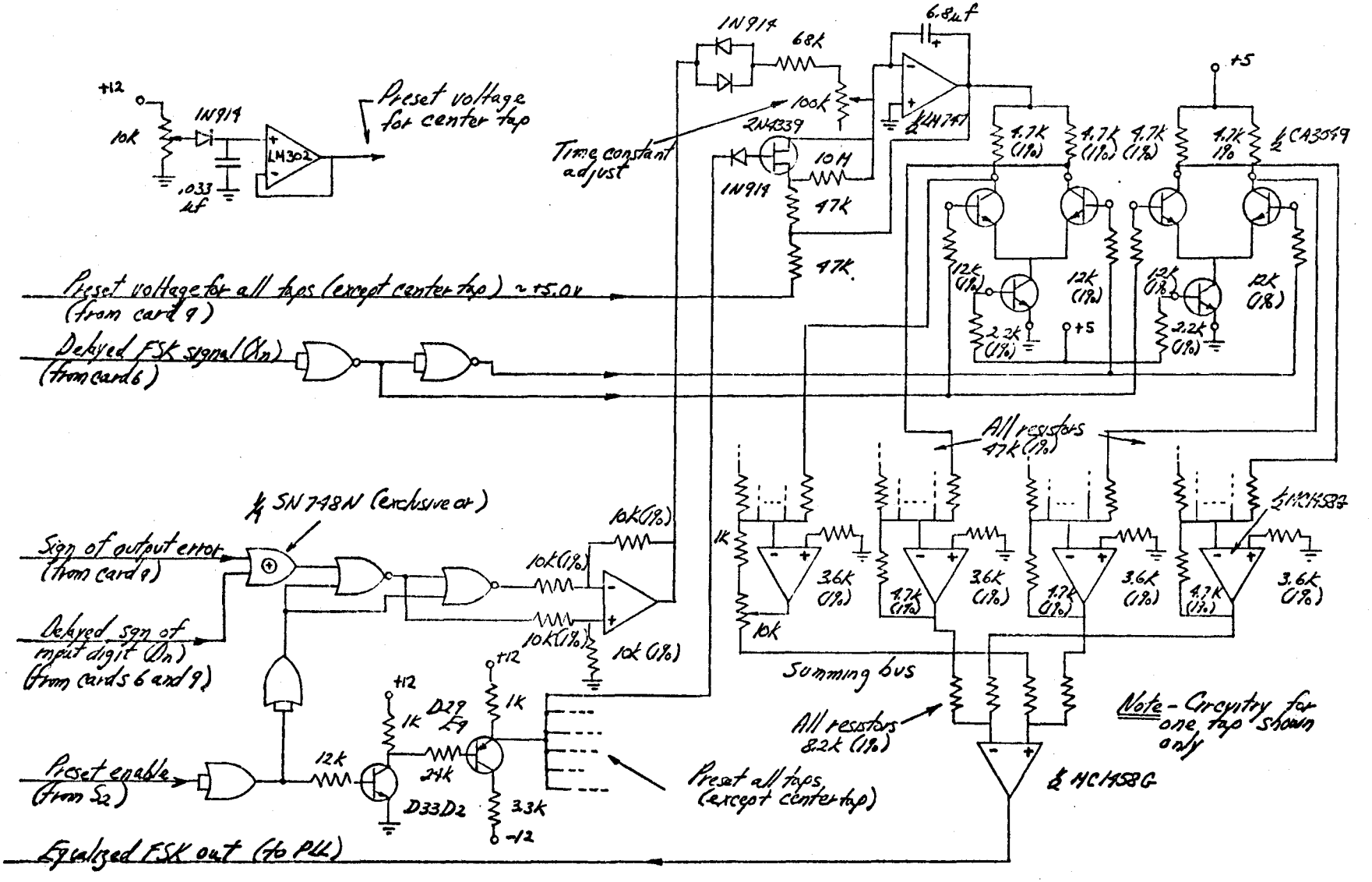


fig.B-7 Cards 7&8

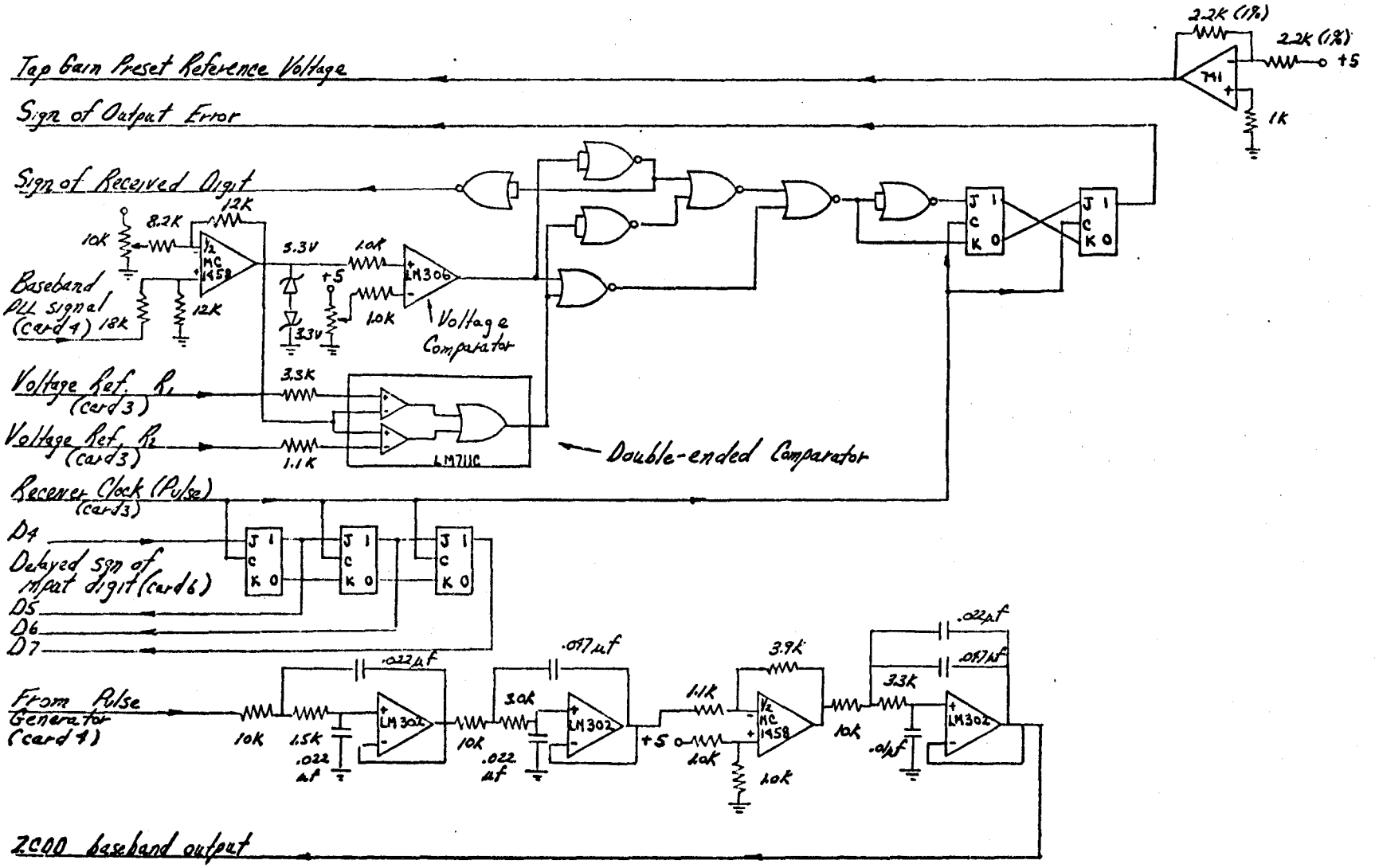

Tap Gain Preset Reference Voltage
Sign of Output Error
Sign of Received Digit
Baseband PLL signal (card 4)
Voltage Ref. R1 (card 3)
Voltage Ref. R2 (card 3)
Receiver Clock (Pulse) (card 3)
D4
Delayed sgn of input digit (card 6)
D5
D6
D7
From Pulse Generator (card 4)
ZC00 baseband output
2.2K (1%)
2.2K (1%)
+5
1K
12K
8.2K
10K
12K
18K
5.3V
3.3V
10K
LM306
Voltage Comparator
3.3K
1.1K
LM711C
Double-ended Comparator
.022μf
.047μf
10K
1.5K
.022 μf
10K
3.9K
1.1K
10K
.022μf
.047μf
3.3K
10K
.047μf
LM302
MC 1458

fig.B-8 Card 9

APPENDIX C

DETERMINATION OF THE LOOP GAIN OF A PLL

This appendix contains the background theory for determining the dc loop gain of a PLL

From chapter 6, the phase error in the frequency domain can be written

$$\theta_e(s) = \frac{s\theta_i(s)}{s + K_{pd}K_{vco}L(s)} \qquad \text{(C-1)}$$

and the steady-state response of the PLL to a step of $\Delta\omega$ radians is given by

$$\lim_{t\to\infty} [\theta_e(t)] = \lim_{s\to 0} \left[\frac{\Delta\omega}{s + K_{pd}K_{vco}L(0)}\right] \qquad \text{(C-2)}$$

since $L(0) = 1$,

$$= \frac{\Delta\omega}{K_{pd}K_{vco}} \qquad \text{(C-3)}$$

Thus the overall gain ($K_{vco}K_{pd}$) can be easily obtained by making a simple measurement of the static phase error - frequency characteristics. We have of course assumed here that the PD is linear.

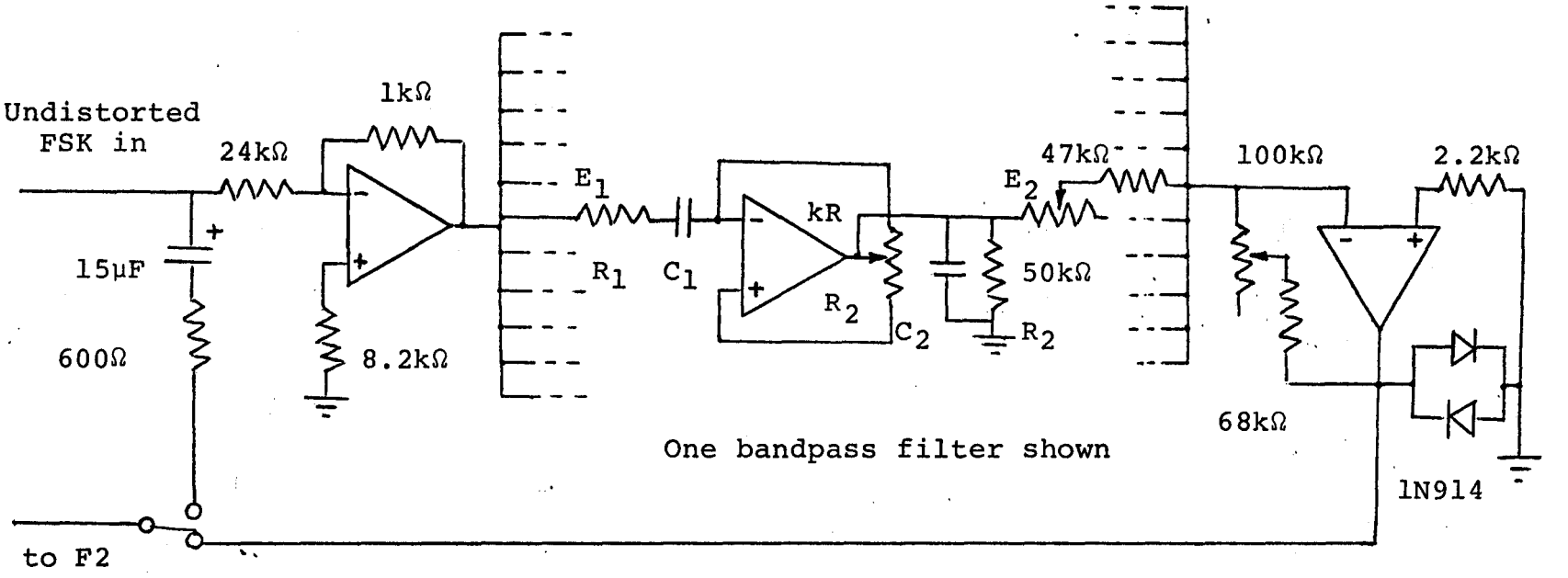
APPENDIX D

CIRCUIT OF FILTER F1 OF CHANNEL SIMULATOR

A circuit diagram of the F1 of the channel simulator is shown in figure D-1. The selected values of R_1, C_1, R_2, C_2 are listed in table D-1.

Nominal Center Frequency (Hz)	R (kΩ)	C (μF)	R (kΩ)	C (μF)
500	6.8	.047	6.8	.047
800	4.3	.047	4.3	.047
1100	2.4	.068	2.4	.047
1300	1.8	.068	1.8	.068
1500	1.5	.068	1.8	.068
1700	1.8	.047	1.8	.047
1900	1.2	.068	1.3	.068
2100	1.5	.047	1.5	.047
2400	4.5	.015	4.7	.015
2700	3.9	.015	3.9	.015
2900	1.6	.033	1.6	.033

Table D-1 Selected Component Values for Bandpass Filters



$$\frac{E_2(s)}{E_1(s)} \simeq \frac{-ksC_1/R_1}{s^2C_1C_2+s(C_2/R_1+C_1/R_2-kC_1/R_1)+1/R_1R_2}$$

fig.D-1 Circuit of Filter F1 of Channel Simulator

REFERENCES

Chapter Three:

1 M.Schartz, "Information Transmission, Modulation, and Noise", McGraw-Hill Book Company, New York, 1959.

2 R.W.Lucky, J.Salz, and E.J.Weldon, "Principles of Data Communication", McGraw-Hill Book Company, New York, 1968.

3 D.A.George, D.C.Coll, A.R.Kaye,and R.R.Brown, "Channel Equalization for Data Transmission", Communications Research Center, Ottawa.

4 R.W.Lucky, "Automatic Equalization for Digital Communication", BSTJ, Vol.44, pp.547-88, April 1965.

5 D.Hirsh and W.J.Wolf, "A Simple Adaptive Equalizer for Efficient Data Transmission", IEEE Trans. on Comm. Tech., Vol.Com-18, pp.5-12, February 1970.

6 C.W.Niessen and D.K.Willim, "Adaptive Equalizer for Pulse Transmission", IEEE Trans. On Comm. Tech., Vol.Com-18, pp.377-95, August 1970.

7 R.W.Lucky, "Techniques for Adaptive Equalization of Digital Communication", BSTJ, Vol.45, pp.255-86, February 1966

Chapter Five:

1 R.deBuda, "The Fast FSK - A New Modulation System" Report RQ71EE2, CGE Technical Information Series, February 1971.

2 P.J.van Gerwin and P.van der Wurf, "Data Modems with Integrated Digital Filters and Modulators", IEEE Trans. on Comm. Tech., Vol. Com-18, pp. 214-22, February 1970.

Chapter Six:

1 F.M.Gardner, "Phaselock Techniques", John Wiley and Sons Inc., New York 1966.

2 A.B.Grebene, "The Monolithic Phase-Locked Loop - a Versatile Building Block", IEEE Spectrum, Vol.6, pp.38-49, March 1971.

3 R.D.Shelton and A.F.Adkins, "Noise Bandwidth of Common Filters", IEEE Trans. on Comm. Tech., Vol. Com-18, pp.828-30, December 1970.

4. J.E.Mazo and J.Salz, "Theory of Error Rates for Digital F.M.", BSTJ, pp.1511-35, November 1966.

5 "Reference Data for Radio Engineers", Howard W.Sams and Company, Indianapolis, 1968.

6 R.T.Bobilin and J.C.Lindenlaub, "Distortion Analysis of Binary FSK", IEEE Trans. on Commun. Tech., Vol. Com-19, pp.478-86, August 1971.

7 W.R.Bennet and J.R.Davey, "Data Transmission", McGraw Hill Book Company, New York, 1968.

8 R.W.Lucky, J.Salz and E.J.Weldon, "Principles of Data Communications", McGraw-Hill Book Company, New York, 1968.

Chapter Seven:

1 F.M.Gardner, "Phaselock Techniques", John Wiley and Sons Inc., New York, 1965.

2 W.R.Bennet and J.R.Davey, "Data Transmission", McGraw-Hill Book Company, New York, 1965.

Chapter Eight:

1 R.W.Lucky, "Automatic Equalization for Digital Communication", BSTJ, Vol.44,pp.547-88, April 1965.

2 A.Lender, "Decision-Directed Digital Adaptive Equalization Technique for High Speed Data Transmission", IEEE Trans. on Commun. Tech., Vol. Com-18, pp.625-31, October 1970.