1990

Design and performance study of an atmospheric laser communication system

Christopher J. Hefele
Lehigh University

Design and Performance Study of an Atmospheric Laser Communication System

by

Christopher J. Hefele

A THESIS
PRESENTED TO THE GRADUATE COMMITTEE
OF LEHIGH UNIVERSITY
IN CANDIDACY FOR THE DEGREE OF
MASTER OF SCIENCE
IN THE DEPARTMENT OF
COMPUTER SCIENCE AND ELECTRICAL ENGINEERING

Lehigh University
1990

Certificate of Approval

This thesis is accepted and approved in partial fulfillment of the requirements for the degree of Master of Science.

5-2-90

Date

Professor in Charge

Chairman of Department

Table of Contents

List of Tables

List of Figures

Abstract

An atmospheric laser communication system has been constructed which uses a laser beam to provide short-haul, line-of-sight communication between two points. The developed system has a range of 3 kilometers, and can operate at data rates of up to 20Mb/s. Using data obtained from monitoring a 3 kilometer test link, the effects of weather on path losses were modeled. These effects include extinction of the beam (due to rain, snow, fog, and other particles in the atmosphere), as well as beam intensity fluctuations due to atmospheric turbulence.

1. INTRODUCTION

1.1 Purpose

Atmospheric laser communication systems provide line-of-sight communications between two points by transmitting a modulated laser beam through the atmosphere between an optical transmitter and receiver. The goal of this investigation was to develop and test such a system using the best methods and devices commercially available. The system was designed for communications over a 3 km distance at a maximum data rate of 20 megabits per second. The system was to be asynchronous and bit rate independent; that is, it had to look like a "wire" to the user. The developed system meets these requirements.

Obviously, weather plays a significant role in the performance of any atmospheric laser communication system. The effects of atmospheric propagation were studied in detail, with specific attention to weather's effects on the system's performance. Experimental data was gathered from 1 km. and 3 km. test links for comparison with the theoretical models. The results show that an atmospheric laser communication system can easily provide an effective short-haul communication link.

1.2 Applications

Atmospheric laser communication systems are becoming sensible alternatives to traditional short-haul microwave links for many reasons. Microwave links, on one hand, require large outdoor antennas, are prone to multi-path fading, and have

bandwidths regulated by the FCC. Obtaining an FCC license is difficult, especially in cities, due to the increased crowding of the microwave frequency spectrum.

Atmospheric laser communication systems, on the other hand, suffer from none of these difficulties. Like optical fiber communication systems, their entire bandwidth is available for use. They require no licensing and are physically smaller than their microwave counterparts. This makes them ideal for portable or temporary communication links. Also, the commercial availability of inexpensive semiconductor lasers and photodetectors has made such systems relatively inexpensive. Furthermore, the small beamwidths of atmospheric laser communication systems allow for an added measure of security that is not possible while using a microwave radio link. Under conditions that cause multi-path fading in microwave links, laser communication systems provide stable communications. This makes them attractive for applications traditionally reserved for microwave links : short haul communications, satellite links, and air/ground communications.

1.3 System Design Considerations

When choosing among various design schemes for a laser communications system, one needs a metric in order to determine which system is superior to another in terms of performance and/or cost. Performance can be measured by bit error rate (BER), for example, or by the signal to noise ratio (SNR) at the receiver. Costs, on the other hand, can be measured by the system price, size, complexity, and power requirements. Since each element of a system interacts with both the system's performance as well as it's costs, certain tradeoffs may have to be made when

designing to meet the desired performance specifications.

For an atmospheric laser communications system, the overall design goal is to maximize system performance and minimize the costs; that is, to maximize system link range for a fixed BER or SNR, while minimizing system price, size, complexity and power consumption of the system. Other desirable characteristics for digital communication links are bit rate transparency, bit pattern independency, and fast signal acquisition time. Since the receiver design influences system performance to a great degree, a receiver with high sensitivity and a wide dynamic range is also desirable (Van Muoi).

In designing to achieve these goals, the major factors in determining overall performance and system costs can be divided into two categories: optical considerations, and circuit considerations. In each category, choices among various design options must be made in light of the specifications for the system to be designed.

1.3.1 Optical Considerations

Laser - The simplest method to increase the optical signal level at the receiver is to use a high power laser in the transmitter. This approach, however, increases the threat of eye damage and increases transmitter power consumption. Semiconductor lasers with radiant output powers in the milliwatt(s) range are small, rugged, and can be directly modulated. Gas lasers, on the other hand, are larger, more fragile, and require more sophisticated techniques to modulate the optical beam, but are

capable of delivering a larger radiant output powers.

Optics - Narrow transmitter beam divergence and large receiving lenses minimize amount of uncollected power at the receiver. Complex optical systems that can provide narrow beam divergence and are costly to fabricate and may be difficult to adjust. Large receiving lenses gather more of the transmitted light, but large lenses that are adequately corrected are expensive.

Pointing Errors - Since laser communication systems use narrow optical beams to transmit information, alignment of the transmitting and receiving optics is critical. If the transmitted beam is not aligned properly with the receiver, or vice versa, some of the signal may be lost. A large beam divergence at the transmitter may prevent misalignment due to vibration, but at the expense of received signal strength. A large acceptance angle at the receiver may ease alignment problems, but then more background light is received, possibly saturating the detector. These problems can be avoided by means of an automatic pointing system that keeps the transmitter and receiver perfectly aligned. This, however, increases cost and complexity of the total system, particularly with high accuracy / narrow beamwidth systems.

Path Loss - Due to physical size limitations, the receiving lenses only collect a small portion of the transmitted beam. This creates large losses ranging on the order of many 10's of dB's. Furthermore, weather conditions influence the optical transmissivity of the atmosphere, and may introduce additional losses under bad weather conditions such as rain, snow or fog. Even in ideal weather, atmospheric turbulence causes the received beam to "twinkle", that is, to vary in intensity over

time. Background radiation from the sky is also added to the signal. All four of these effects are unavoidable; however, they must be understood if we are to understand system performance in the field.

1.3.2 Circuit Considerations

Photodetector - The photodetector sensitivity is obviously a very important factor in determining how a system performs, since only a small portion of the transmitted beam is collected. Avalanche photodiodes can be used to detect the transmitted beam, since they have a high sensitivity due to their internal multiplication gain processes. These detectors require a high voltage power supply to set the detector gain. How this gain is set effects not only the photodiode's sensitivity to light, but also it's noise characteristics. Large detector areas can decrease the focusing requirements of the lens, but at the expense of detector bandwidth.

Receiver Preamplifier - The first amplifier stage in the receiver is a main source of noise in the system. Proper selection of the first stage amplification scheme as well as proper matching of the amplifier to the photodetector can decrease system noise, increase the signal to noise ratio, and thus increase system performance.

Modulation Scheme - The proper choice of a modulation scheme (or even not to modulate) can influence system SNR by as much as 6 dB, which can influence link range by a factor of two. Of course, the simplest methods perform more poorly than more sophisticated methods, which are also more expensive and require more complex circuitry.

System Bandwidth - High system bandwidths allow for the transmission of data at high speed. Large bandwidths are difficult to obtain and increase circuit complexity, system price, and system power requirements. Furthermore, system performance is optimized for an anticipated maximum bit rate; thus a high bandwidth system used for low speed communications will not have optimum performance.

Error Correction Codes - In a digital communications system, error coding techniques can be used to encode the transmitted data. These codes can be used at the receiver to reconstruct the transmitted data even though the received data contains bit errors. Of course, complex digital circuitry must be present at the transmitter and receiver sides of the link in order to encode and decode the data stream. Furthermore, the system will not be asynchronous; that is it must be designed for fixed data rates. However, the equivalent reduction in the signal to noise ratio can be as large as 10 dB.

2. MODULATION

2.1 Introduction to Modulation

Several difficulties in atmospheric laser communication systems can be overcome by *modulating* the data to be sent. Instead of sending data directly, digital modulation techniques use an analog time-varying signal to represent the each of the digital states to be sent. There are several advantages to using a modulation scheme over direct (baseband) modulation. First, the receiver signal-to-noise ratio can be improved by up to 3 dB. Second, modulation allows AC coupling between amplifier stages, thus simplifying circuit design and removing the effects of background light (DC) components in the received optical signal. This allows the construction of receivers having a larger dynamic range, which is important because of the wide range of signal levels received by an optical receiver under various weather conditions.

In order to chose among the various modulation schemes, their basic characteristics must be explored. The three basic digital modulation schemes we will study here are Amplitude Shift Keying (ASK), Frequency Shift Keying (FSK), and Phase Shift Keying (PSK). The basis for understanding these digital modulation schemes, however, can be found by studying amplitude modulation in general.

2.2 Amplitude Modulation

In an amplitude modulation (AM) transmitter, the amplitude of a sinusoidal carrier

is varied according to the amplitude of a modulating signal. At the receiver, the carrier signal is removed, and the original amplitude information recovered. This is considered in detail below.

2.2.1 Transmitter Modulation

If, at the input of a transmitter, we have an analog message signal m(t) with AC and DC components

$$m(t) = A_{dc} + A_m \cos(\omega_m t) \quad \left(\text{where } m(t) > 0 \right)$$

and a carrier x(t) generated inside the transmitter such that

$$x(t) = A_c \cos(\omega_c t) \qquad \left(\text{where } \omega_c \gg \omega_m \right)$$

we can form an amplitude modulated (AM) signal y(t) by multiplying x(t) and m(t):

$$y(t) = m(t) \cdot x(t)$$

This is depicted in figure 2-1. Substituting for m(t) and x(t) yields

$$y(t) = \left(A_{dc} + A_m \cos(\omega_m t) \right) A_c \cos(\omega_c t)$$

Recalling that $\cos(A)\cos(B) = \frac{1}{2}\left(\cos(A+B) + \cos(A-B) \right)$, we obtain

Amplitude Modulation - Transmitter

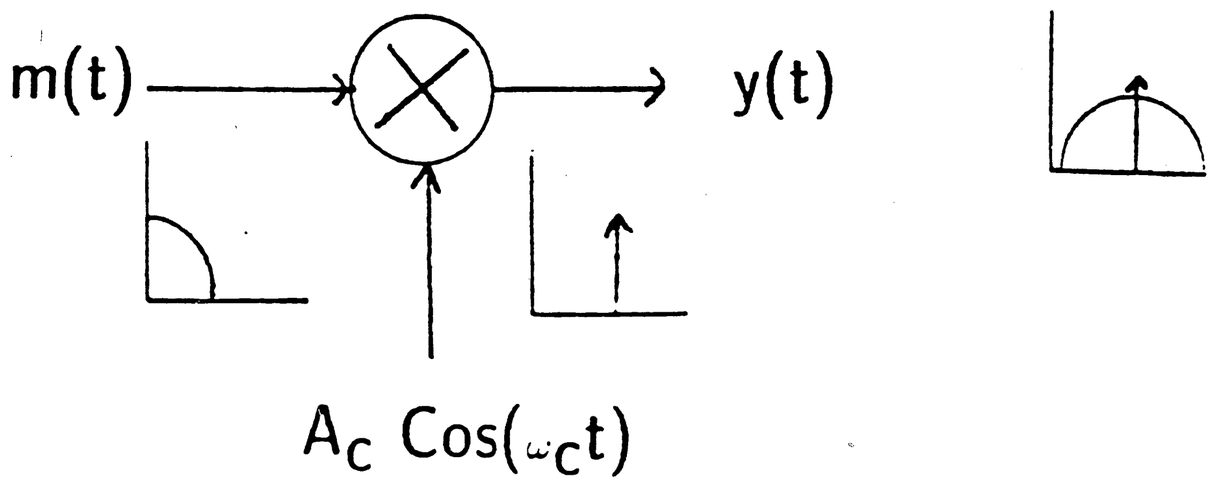


Figure 2-1

Amplitude Modulation - Receiver

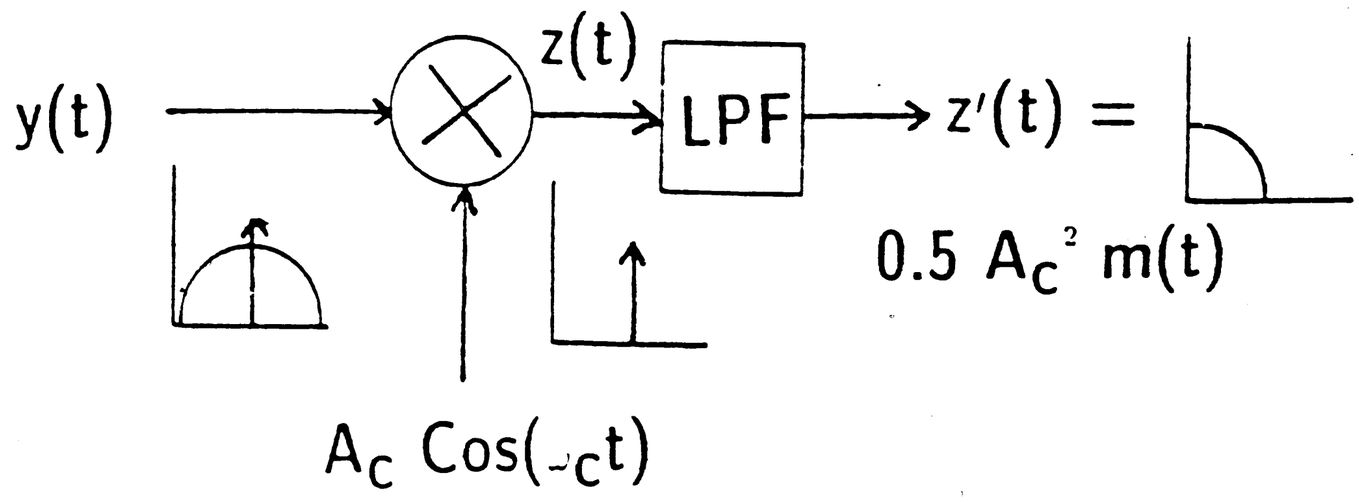


Figure 2-2

$$y(t) = A_{dc}\, A_c \cos(\omega_c t) + \frac{A_m\, A_c}{2}\left(\cos(\omega_c t+\omega_m t) + \cos(\omega_c t-\omega_m t)\right)$$

The output signal y(t) has spectral components at the original carrier frequency as well as at the sum and difference of the carrier frequency and the modulating frequency. The first term in the above expression is the *carrier* component, and the other two spectral components are the *sidebands*, specifically, the upper and lower sidebands at $\omega_c+\omega_m$ and $\omega_c-\omega_m$ respectively. If m(t) has a whole spectrum of frequency components in the range $\omega = 0$ to ω_{max}, using superposition for each component, one can see that the baseband spectrum is mirrored around the carrier frequency ω_c, the lower sideband from $\omega_c-\omega_{max}$to ω_c, and upper sideband from ω_c to $\omega_c+\omega_{max}$. This is shown in figure 2-1.

Note that while the sidebands contain information about the modulating signal's AC level and frequency, the carrier component contains *only* information about the modulating signal's DC level. Since the total signal power is shared by the sidebands as well as the carrier, the more power put in carrier, the less power is available to transmit information in the sidebands.

We can define the efficiency, E, as the ratio of energy spent to send useful information to the total energy. Thus,

$$E = \frac{<A_m{}^2 \cos^2(\omega_m t)>}{A_{dc}{}^2+<A_m{}^2 \cos^2(\omega_m t)>} \cdot (100\%)$$

For a sinusoidal-type signal amplitude modulating a carrier, it can be shown that the

maximum efficiency is E=33.33%. For square-wave type signals, the maximum efficiency is E=50%.

2.2.2 Coherent Receiver Demodulation

At the receiver, the signal y(t) is multiplied (or *mixed*) with the carrier signal x(t) to form the demodulated signal z(t):

$$z(t) = y(t) \cdot x(t)$$

This is shown in figure 2-2. The signal x(t) may be generated internally by the receiver, or it may be obtained from the signal y(t) itself by careful filtering around the carrier frequency ω_c.

Now, substituting for y(t) and x(t) yields

$$z(t) = \left[A_{dc}\, A_c \cos(\omega_c t) + \frac{A_m\, A_c}{2} \left(\cos(\omega_c t + \omega_m t) + \cos(\omega_c t - \omega_m t) \right) \right] \cdot A_c \cos(\omega_c t)$$

Which, after trigonometric substitutions, is

$$z(t) = \frac{A_{dc} A_c^{\,2}}{2} \left[1 + \cos(2\omega_c t) \right] +$$

$$\frac{A_m A_c^{\,2}}{4} \left[\cos(2\omega_c t + \omega_m t) + \cos(2\omega_c t - \omega_m t) + \cos(\omega_m t) + \cos(-\omega_m t) \right]$$

If we pass this signal through a low-pass filter with cut-off frequency between ω_m and ω_c , all terms are eliminated except the DC and ω_m terms. Thus, if $z'(t)$ is the

output of the low pass filter, we obtain

$$z'(t) = \frac{A_{dc}A_c^2}{2} + \frac{A_m A_c^2}{4}\Big(\cos(\omega_m t) + \cos(-\omega_m t)\Big)$$

Factoring out $A_c^2/2$, and since $\cos(\omega_m t)=\cos(-\omega_m t)$, the demodulated output $z'(t)$ is

$$z'(t) = \frac{A_c^2}{2}\Big[A_{dc} + A_m \cos(\omega_m t)\Big] = \frac{A_c^2}{2}\, m(t)$$

Thus, the demodulated signal $z'(t)$ is the same as the original message signal $m(t)$, only multiplied by $A_c^2/2$.

2.2.3 Incoherent Receiver Demodulation

Instead of using a multiplier (mixer), a peak detector can be used to easily decode an AM signal. Since the modulating signal changes the carrier amplitude relatively slowly if $\omega_c \gg \omega_m$, the peak of each carrier cycle represents the modulating signals amplitude. This method, which does not rely on reconstruction of the carrier frequency, is one type of *incoherent* detection, as opposed to *coherent* methods that rely on reconstruction of the carrier signal. Although incoherent AM reception can be shown to provide approximately 1 dB poorer performance than coherent detection, it has the advantage of simplicity in circuit design.

2.3 Digital Modulation Schemes

Digital modulation schemes represent each binary symbol with one of two time-

varying sinusoidal signals. These two sinusoidal signals can differ in three ways: in amplitude, in frequency, or in phase. Each of these gives rise to a form of digital modulation.

2.3.1 Amplitude Shift Keying (ASK)

Amplitude Shift Keying uses a digital source to shift a carrier between two different amplitudes. For example, if we have digital signal d(t) which takes on the values 0 and 1, we can use it to turn on and off a carrier, thus shifting it's amplitude, i.e.

$$y_{ASK}(t) = d(t) \cdot \cos(\omega t)$$

This is depicted in figure 2-3. If d(t) has a DC component, then there will be a carrier component. If 1's and 0's are equally probable, the DC level of d(t) is 1/2. It can be shown that in this case, *half* of the signal power is wasted in the carrier component, which does not convey any time-varying information.

2.3.2 Frequency Shift Keying (FSK)

Frequency shift keying uses a digital source to shift the output signal y(t) between two different frequencies. For example, if we have a digital source d(t) which takes on the values of 0 and 1, the modulated signal will be

$$y_{FSK}(t) = \begin{cases} A\cos(\omega_1 t) & \text{for } d(t) = 0 \\ A\cos(\omega_2 t) & \text{for } d(t) = 1 \end{cases}$$

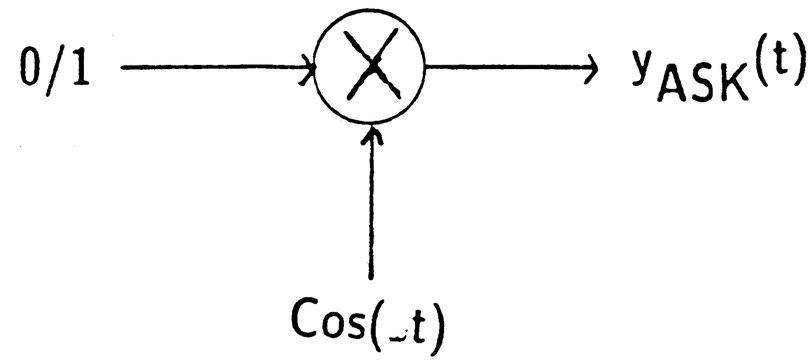


Figure 2-3

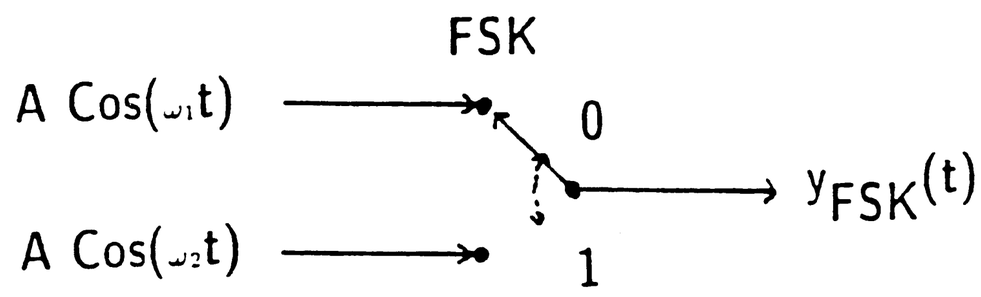


Figure 2-4

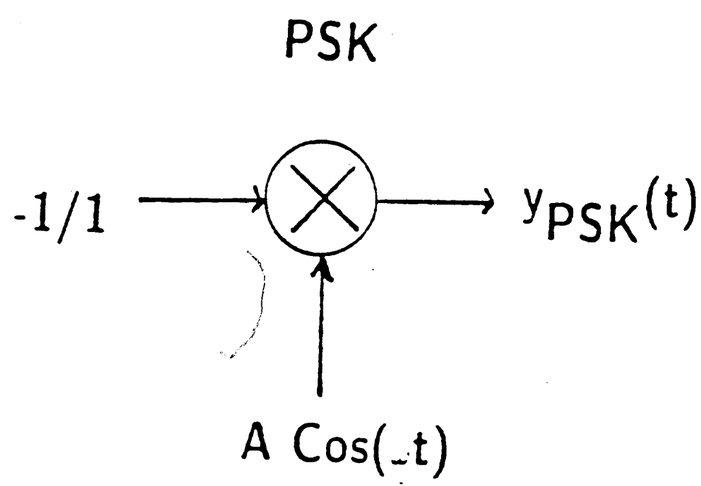


Figure 2-5

A simplified representation of an FSK modulator is shown in figure 2-4. The signal $y_{FSK}(t)$ can be thought of as the sum of two ASK signals, one with carrier frequency ω_1 and one with carrier frequency ω_2. In order to see this, we can rewrite $y_{FSK}(t)$ as

$$y_{FSK}(t) = d(t) \cdot A \cos(\omega_1 t) + \Big(1\text{-}d(t)\Big) \cdot A \cos(\omega_2 t)$$

There are further considerations when we deal with FSK. In order to prevent distortion, the two carriers must be spaced far enough apart in frequency so that the sidebands around each carrier do not overlap significantly. Also note that the signal information is sent in duplicate; i.e. d(t) modulates both carriers. Since FSK can be thought of as two ASK signals, again we find that for square-wave type signals such as d(t), half the signal power is in the carriers and is not conveying any information. Also, FSK takes up at least twice the bandwidth as ASK, since it in effect sends d(t) in duplicate.

2.3.3 Phase Shift Keying (PSK)

Phase shift keying uses a digital source, d(t) to shift the phase of a carrier wave between two different phases. The general form of PSK is

$$y_{PSK}(t) = A \cos\left[\omega t + (-1)^{d(t)} \cdot \cos^{-1}(m)\right]$$

The modulation parameter, denoted by m, determines how much the phase shifts.

Recalling that $\cos(u \pm v) = \cos(u)\cos(v) \mp \sin(u)\sin(v)$, and using the identities

$\cos(\cos^{-1}(m))=m$ and $\sin(\cos^{-1}(m))=\sqrt{1-m^2}$, we obtain

$$y_{PSK}(t)= A\cdot m\cdot\cos(\omega t) - (-1)^{d(t)} A\sqrt{1-m^2}\,\sin(\omega t)$$

The first term on the right hand side is the carrier component. The second term is an ASK modulated carrier, modulated by $(-1)^{d(t)}$. Note that this term has an average value of zero and thus no DC component, and therefore no ASK carrier component at ω due to the $\mathrm{Sin}(\omega t)$ term. Therefore, the only carrier component is due to the $A\cdot m\cdot\cos(\omega t)$ term.

By adjusting m, the modulation parameter, we change the amount of phase shift and the amount of carrier. If we choose a phase shift of $\pm\pi$ by using m=0, the resulting PSK signal is

$$y_{PSK}(t)= - (-1)^{d(t)} A\,\mathrm{Sin}(\omega t)$$

A simplified diagram of a PSK modulator is shown in figure 2-5. The signal simply switches between $\mathrm{Sin}(\omega t)$ and $-\mathrm{Sin}(\omega t)$. For this simple case, it is easy to see that all the signal power is in the sidebands conveying information, since there is no carrier component.

2.4 Effects of Noise on Bit Error Probability

If Gaussian noise is present in a communication channel, the noise will cause errors in the demodulated data. For any coherent modulation scheme sending binary data, it can be shown (Ziemer) that the probability of bit error is given by

$$P_e = \frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{E(1-\rho)}{2N_0}}\right)$$

where E is the total signal energy, ρ is the signal correlation, N_0 is the noise energy, and the error function, erfc($\cdot$) is defined as

$$\mathrm{erfc}(u) = 1 - \frac{2}{\pi}\int_0^u e^{-v^2}\,dv$$

If we define the two distinct modulator signals in a binary modulation scheme as $s_0(t)$ and $s_1(t)$, then the total signal energy, E, is found by

$$E = \frac{1}{2}\int_0^T [\, s_0^2(t) + s_1^2(t)\,]\,dt$$

The correlation coefficient, ρ, between the two signals $s_0(t)$ and $s_1(t)$ is defined as

$$\rho = \int_0^T \frac{s_0(t)\,s_1(t)}{E}\,dt$$

and is bounded by ± 1. If the two signals are very much alike, ρ approaches 1, but as they become more and more distinct (and thus more uncorrelated), ρ decreases to zero. If the signals have equal magnitudes but opposite signs, ρ=-1.

For ASK, $s_0(t) = \cos(\omega t)$ and $s_1(t) = 0$. After performing the integrations to find ρ, we find that ρ=0, and thus the probability of bit error for ASK is

$$P_{e,ASK} = \frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{E}{2N_0}}\right)$$

Recognize that E / N_0 is the signal to noise ratio (SNR), thus

$$P_{e,ASK} = \frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{SNR}{2}}\right)$$

Using similar methods for FSK and PSK, we can find the probability of bit error for FSK and PSK are, respectively,

$$P_{e,FSK} = \frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{SNR}{2}}\right)$$

$$P_{e,PSK} = \frac{1}{2}\,\mathrm{erfc}\left(\sqrt{SNR}\right)$$

For non-coherent modulation schemes, the derivation of bit-error probability is more complicated, since less is assumed known about the signal. The error probabilities for FSK and ASK have been shown to be (Ziemer) :

$$P_{e,ASK} = \frac{1}{2}\,e^{-SNR/2}$$

$$P_{e,FSK} = \frac{1}{2}\,e^{-SNR/2}$$

Non-coherent PSK is a contradiction in terms; it is simply impossible to convey any information in a carrier of totally random phase.

The graphs of each of these probability of error functions considered are shown in

figure 2-6.

2.5 Summary and Selection of Modulation Scheme

Table 2-1 summarizes the characteristics of the ASK, FSK and PSK schemes analyzed thus far. To summarize, given a fixed peak power, ASK is 6 dB worse than PSK; FSK is 3dB worse than PSK and requires twice the bandwidth; PSK is the most efficient modulation scheme of the three. Note that if we constrain *peak* power to a fixed value (instead of average power), ASK will be 3dB worse, since it's average power is not it's peak power, as is the case with FSK and PSK.

With PSK, however, are some inherent difficulties. First, a carrier signal must be regenerated at the receiver in order to demodulate the PSK signal. This carrier must have the exact frequency and phase of the original carrier at the transmitter. This is possible to achieve, but it complicates receiver design. These difficulties, however, do not change the selection of PSK as the most efficient modulation scheme of the three, and therefore the most desirable for our purposes.

Probability of Bit Error vs. SNR

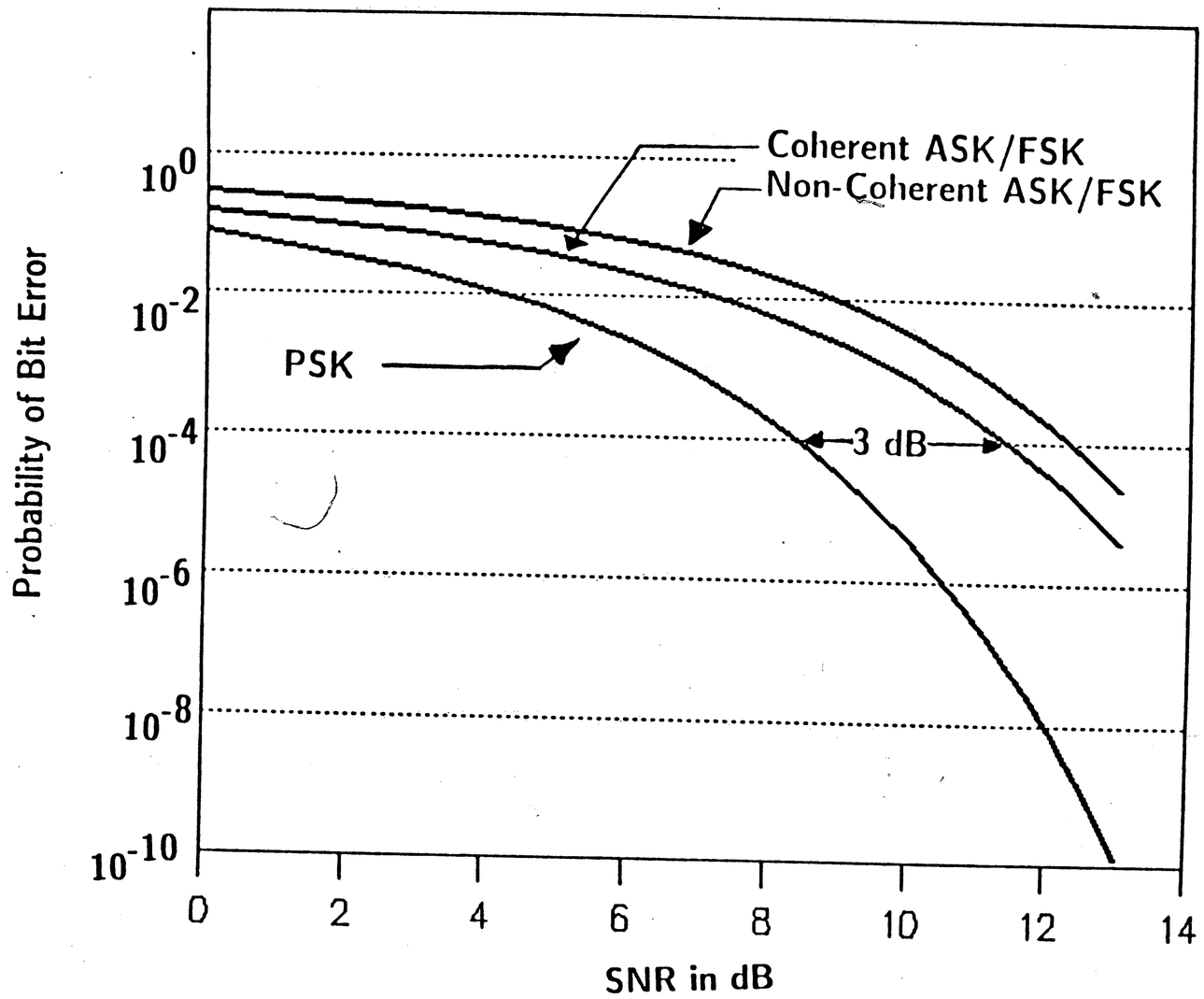


Figure 2-6

TABLE 2-1 Modulation Performance Summary

Modulation Scheme	Maximum Efficiency	Minimum Bandwidth	Avg Pwr ÷ Peak Pwr	Probability of Bit Error
Coherent ASK	50%	$2\omega_{Max}$	0.5	$\frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{SNR}{2}}\right)$
Non-Coh. ASK	50%	$2\omega_{Max}$	0.5	$\frac{1}{2}\,e^{-SNR/2}$
Coherent FSK	50%	$4\omega_{Max}$	1.0	$\frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{SNR}{2}}\right)$
Non-Coh. FSK	50%	$4\omega_{Max}$	1.0	$\frac{1}{2}\,e^{-SNR/2}$
Coherent PSK	100%	$2\omega_{Max}$	1.0	$\frac{1}{2}\,\mathrm{erfc}\left(\sqrt{SNR}\right)$

Note:

SNR=Signal to noise ratio;

ω_{Max}= highest frequency component in the modulating signal

3. MODULATOR/DEMODULATOR SYSTEM DESIGN AND ANALYSIS

3.1 Introduction

The modulator/demodulator pair that was constructed can operate at data rates of up to 20 Mb/s, using PSK modulation. The goal here is to outline the design of the modulator and demodulator from a communications theory viewpoint. Each of the elements of the modulator and demodulator are analyzed in some detail so that the performance of the entire system can be understood in some depth.

3.2 Modulator/Demodulator System Design

A block diagram of the 20 Mb/s PSK system that was constructed is shown in figure 3-1. Since the most efficient modulation scheme is PSK, it is the modulation scheme employed. The carrier frequency of the system was chosen to be 100 MHz, which gives 5 carrier cycles per 20 Mb bit period. In order to generate the 100 MHz carrier, a 50 MHz crystal oscillator generates a signal which is then sent to a frequency doubler. This approach is used due to the fact that crystals with resonant or 3rd harmonic frequencies above 50 MHz are difficult to make. As shown in the diagram, a small amount of the 50 MHz carrier is injected into the output to aid the receiver in demodulation.

3.2.1 Carrier Injection

PSK modulation requires a carrier oscillator at the receiver to be in synchronization

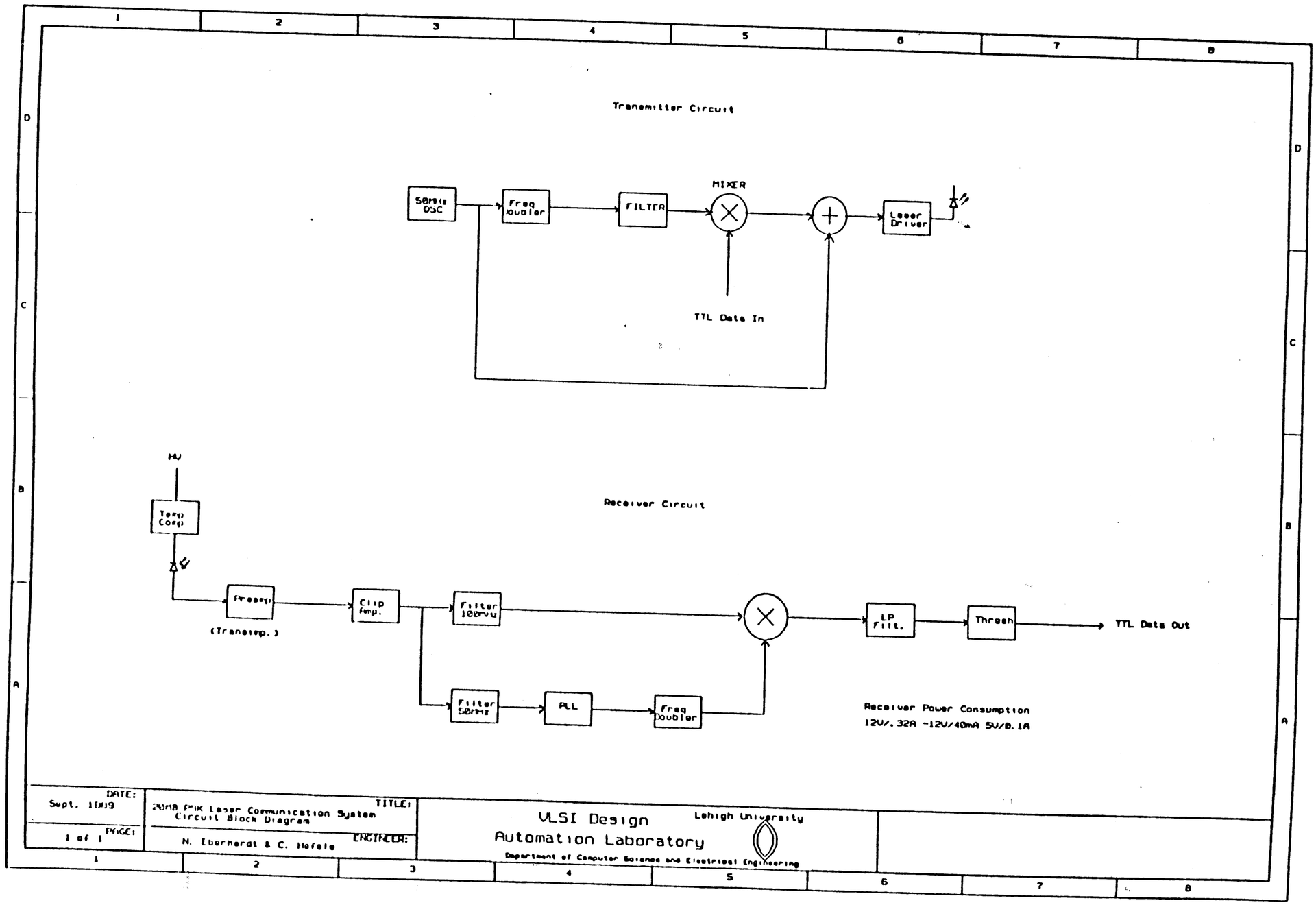

Transmitter Circuit
50MHz OSC
Freq. Doubler
FILTER
MIXER
Laser Driver
TTL Data In
HV
Temp Comp
Preamp
(Transimp.)
Clip Amp.
Filter 100KHz
Filter 50MHz
PLL
Freq Doubler
LP Filt.
Thresh
TTL Data Out
Receiver Circuit
Receiver Power Consumption
12V/.32A -12V/40mA 5V/0.1A
DATE: Sept. 1989
PAGE: 1 of 1
TITLE: 20MB FDIK Laser Communication System Circuit Block Diagram
ENGINEER: N. Eberhardt & C. Hofele
VLSI Design Automation Laboratory
Lehigh University
Department of Computer Science and Electrical Engineering

Figure 3-1

with the carrier oscillator at the transmitter. The easiest method to have the carrier present at the receiver is to add it to the transmitted PSK modulated signal, and then isolate it at the receiver. However, since there may be PSK sidebands near the carrier frequency, isolating the carrier signal would be difficult. The alternative, which is employed here, is to send the 50 MHz carrier with the PSK signal. At the receiving end, the 50 MHz signal is isolated by very narrowband filters and a phase-locked loop. To recreate the 100 MHz carrier at the receiver, the 50 MHz signal is sent to a frequency doubler.

The 50 MHz carrier level is adjustable. In order not to waste to much power in the carrier (such as in FSK or ASK), the carrier is injected at a low level, at 1/10 the amplitude of the PSK signal. This is practical because the signal is received through a very narrow bandwidth filter centered at 50 MHz; thus the carrier will not easily be masked by noise.

3.2.2 PSK Modulation and Demodulation

At the transmitter, the multiplier/mixer modulates the 100 MHz carrier with the bipolar input data at the transmitter. Then, the low level 50 MHz carrier is injected, and the signal is then sent to an amplifier to modulate the laser output. At the receiver, the photodetector detects the modulated signal, and the signal is amplified. After amplification, the signal is input to the limiter, which removes amplitude fluctuations. Then, the signal is passed to the 50 MHz carrier detector circuit (consisting of a filter and a PLL), and the 100 MHz filter that isolates the PSK sidebands. The 50 MHz component is sent to a frequency doubler in order to

reconstruct the 100 MHz carrier. The mixer then multiplies the reconstructed 100 MHz carrier with the PSK signal at 100 MHz, thus mixing the PSK modulated signal back down to baseband. Any remaining high frequency components are removed with a low pass filter. The result is passed to a Schmitt-trigger type device, with a hysteresis characteristic, that decides whether which of the two binary levels were sent.

3.2.3 Spectrum of PSK Signal

The spectrum of the transmitted waveform must be known in order to properly design filters to reject out of band noise. The system should be designed for the fastest bit rate expected, here 20 Mb/s. The power spectrum of a random binary data stream can be shown to be (Ziemer):

$$S(f) = A^2T \frac{\sin^2(\pi fT)}{(\pi fT)^2}$$

where the time between data bits is T, i.e. T=1/(Bits/Sec). If a random bit stream is used to modulate a PSK carrier of frequency f_c, then the power spectrum is centered around f_c:

$$S(f) = A^2T \frac{\sin^2(\pi(f-f_c)T)}{(\pi(f-f_c)T)^2}$$

Using this, and recalling that the system has a low level carrier at 50 MHz, the spectrum of the 20 Mb/s transmitted signal is shown in figure 3-2. The PSK sidebands about 100 MHz are apparent, as is the injected 50 MHz reference signal.

Transmitter Signal Power Spectrum using 20 Mb/s Random Bit Source

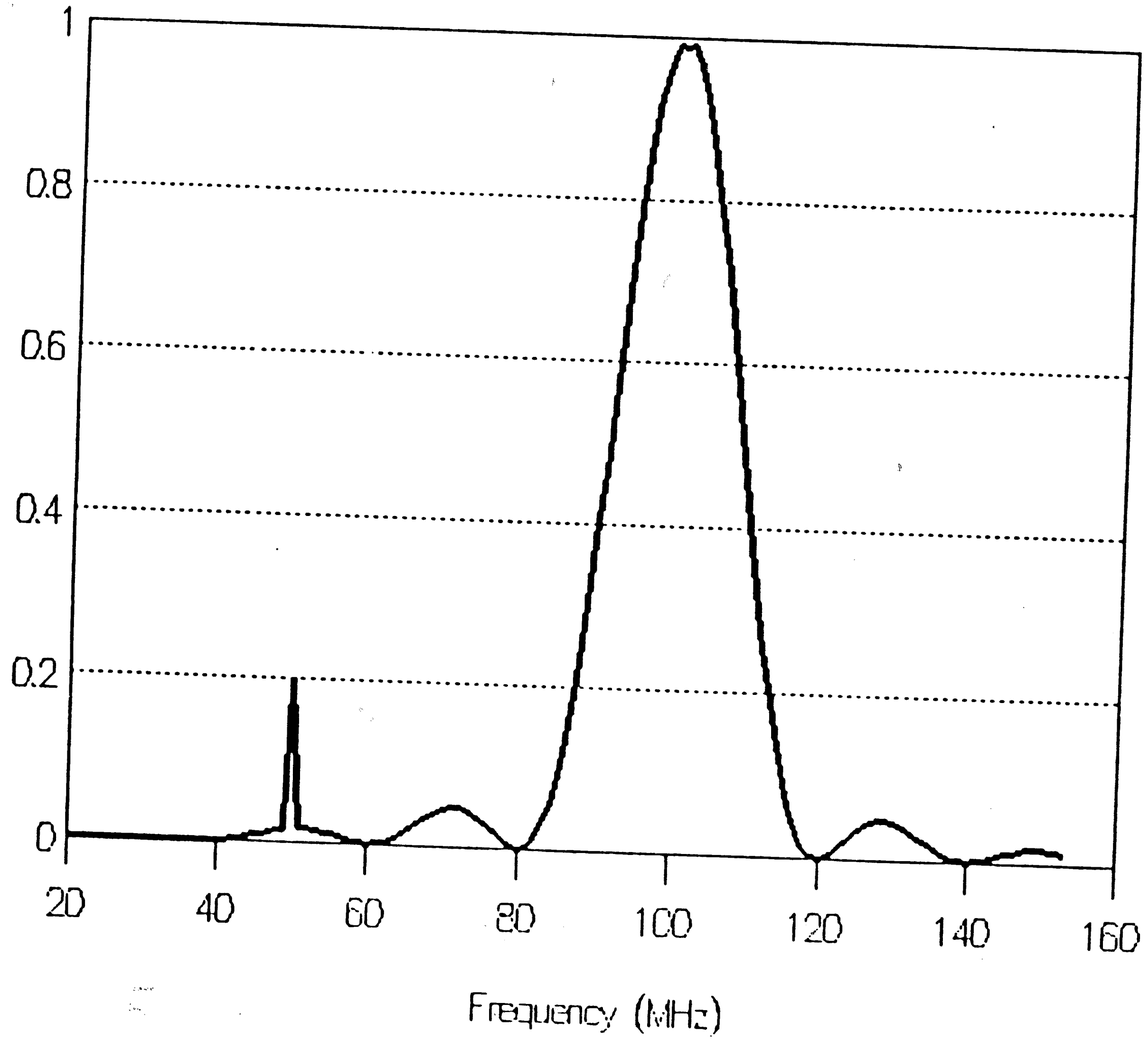


Figure 3-2

From the graph, it is easy to see that the 100 MHz filter must have a full bandwidth of 40 MHz in order to capture the main lobe of the PSK signal spectrum.

3.3 Modulation/Demodulation System Elements

The performance of the modulation/demodulation system is dependant on the performance of each of the elements of the system. From a communication theory point of view, the limiter, mixers, phase detector, frequency doublers, and phase locked loop, shown in figure 3-1, are of interest. In order to fully understand the system's behavior, one must fully understand the operation of each of these sub-systems.

3.3.1 Limiter Analysis and Simulation

Limiters are used mainly in removing amplitude fluctuations from a signal, while preserving the signal's fundamental spectral components. In certain applications, limiters can replace the use of an automatic gain control, and increase dynamic range. Limiters, however, do change both the spectrum and signal-to-noise ratio of the signals passing through them. Only by studying the effects of limiting can the advantages of limiting be fully exploited.

3.3.1.1 Effects on Spectrum

A "hard" limiter is a circuit with the characteristic

$$f(x) = \begin{Bmatrix} +A \\ -A \end{Bmatrix} \text{ if } \begin{Bmatrix} x \geq 0 \\ x < 0 \end{Bmatrix}.$$

To see how a limiter can be used to remove AM fluctuations of a signal, consider a sinusoidal signal with a modulating function M(t),

$$x(t) = M(t)\cos(\omega t) \quad (\text{ with } \infty > M(t) > 0)$$

If we apply x(t) to the input of the limiter, the output, y(t), will be approximately a square wave with a frequency ω. This approximation is valid only if m(t) changes much more slowly than $\cos(\omega t)$. The frequency spectrum of this square wave contains spectral lines at the fundamental frequency ω, as well as harmonics frequencies $n\omega$, where n is an odd, positive, nonzero integer. The Fourier series of the output square wave takes the form

$$y(t) = \frac{4A}{\pi}\left[\cos(\omega t) - \frac{1}{3}\cos(3\omega t) + \frac{1}{5}\cos(5\omega t) \ldots\right]$$

After using a low pass filter to remove all but the fundamental frequency component, we obtain

$$y'(t) = \frac{4A}{\pi}\cos(\omega t)$$

where $y'(t)$ is the output of the low-pass filter. Note that the effect of the modulation component, M(t), has been removed by the limiter. This method can be used to eliminate the need for an automatic gain control in amplifiers requiring a constant output level.

In the case of more than one input spectral component, a new analysis is required. Note that the superposition principle is *not* applicable here, since we are dealing with the non-linear limiting operation. Superposition is only valid for linear circuits. To demonstrate this, consider for a moment that the output y(t) from any memoryless nonlinearity with input signal x(t) can be represented in terms of a Taylor series expansion,

$$y(t) = f(x(t)) = \sum_{i=0}^{\infty} F_i x^i(t)$$

Now, if $x(t) = \cos(\omega_1 t) + \cos(\omega_2 t)$, then we have

$$y(t) = f(x(t)) = \sum_{i=0}^{\infty} F_i \Big(\cos(\omega_1 t) + \cos(\omega_2 t)\Big)^i$$

If we expand this polynomial, recognizing that $\cos(A)\cos(B) = \frac{1}{2}\Big(\cos(A-B) + \cos(A+B)\Big)$ and $\cos^2(A) = \frac{1}{2}\Big(1 + \cos(2A)\Big)$, we can obtain an equation of the form

$$y(t) = \sum_{n=1}^{\infty} A_n \cos(n\omega_1 t) + \sum_{m=1}^{\infty} B_m \cos(m\omega_2 t) + \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} C_{m,n} \cos(n\omega_1 t + m\omega_2 t)$$

where the coefficients A_n, B_m, $C_{m,n}$ can be determined once the F_i's are known. The first two summations consist of the *harmonic* components, which are multiples of ω_1 and ω_2. The remaining summation consists of *intermodulation* components, which are the sums and differences of the harmonics of ω_1 and ω_2. Intermodulation components are unique to nonlinear systems.

Using the above approach, the ideal "hard" limiter has been modeled by finding the expansion of $f(u)=u^{1/n}$ (n odd), and taking the limit as $n\to\infty$ in order to solve the

system (Davenport). From this analysis it can also be shown that the "hard" limiter has the highest amplitude of harmonic and intermodulation components of limiters of the form $f(u)=u^{1/n}$, given two input sinusoids.

3.3.1.2 Simulation of Limiter Performance with PSK Signal

If the limiter function or the limiter input signal becomes complicated, so does the formal analysis. Since the PSK signal of interest here is complicated, numerical methods were used to simulate the output spectrum of the limiter given the signal characteristics at the input.

The hard limiter input signal that was simulated was a 100 MHz PSK signal modulated at 20 Mb/s, with 50 MHz carrier 1/5 the size of the 100 MHz carrier. The 20 Mb/s modulating signal was a band-limited square wave, with exponential rising and falling edges having a 10%-90% rise time t_{rise} of 10 ns. If a bit change occurs at a switching time t_s, the signal is modeled as

$$x(t)= A\cos(\omega_1 t)\cdot\left[e^{-(t-t_s)\,/\,(t_{rise}/2)} - \frac{1}{2}\right] + \frac{A}{5}\cos(\omega_2 t)$$

where $\omega_1=2\pi\cdot 100$ MHz, $\omega_2=2\pi\cdot 50$ MHz, $t_{rise}=10$ ns, $t>t_s$, $A=10000$, and the limiter output levels were ± 1000.

For simulation, the above model was used to compute 2048 points of the signal x(t). The simulation duration was for 1.57 μs, which is 157 cycles of the 100 MHz carrier.

The first 200 ns of the modeled signal input to the limiter can be seen in figure 3-3. The non-instantaneous phase changes of the PSK modulation can be seen approximately every 50 ns.

The spectrum of the input signal was found by taking a FFT of the 2048 sample points of x(t). The spectrum is shown in figure 3-4. The 2 PSK sidebands are apparent, centered around 100 MHz. The injected carrier is visible at approximately 50 MHz. The other harmonics present are due to the 20 Mb/s modulation.

The output of the hard limiter is shown in figure 3-5. and the spectrum of the output is shown in figure 3-6. The spectrum remains essentially the same, however extra harmonic components and intermodulation products are present at frequencies above 200 MHz; these are not visible in the graph. These can be removed by using a bandpass filter after the limiter.

3.3.1.3 Effect on Signal-to-Noise Ratio

When Gaussian noise is added to the signal x(t), and is then passed through the limiter and output band-pass filter, the signal-to-noise ratio at the output is slightly different from the input. It has been shown (Davenport) that the SNR can be improved by 1 to 3 dB for signal-to-noise ratios over 0 dB. Thus, the output SNR can be considered essentially the same as the input SNR. Davenport's analysis, however, is only valid for a single spectral component. When more than one sinusoid plus noise is input to the limiter, formal derivations of the output signal spectrum and SNR become more complicated, due to the presence of intermodulation

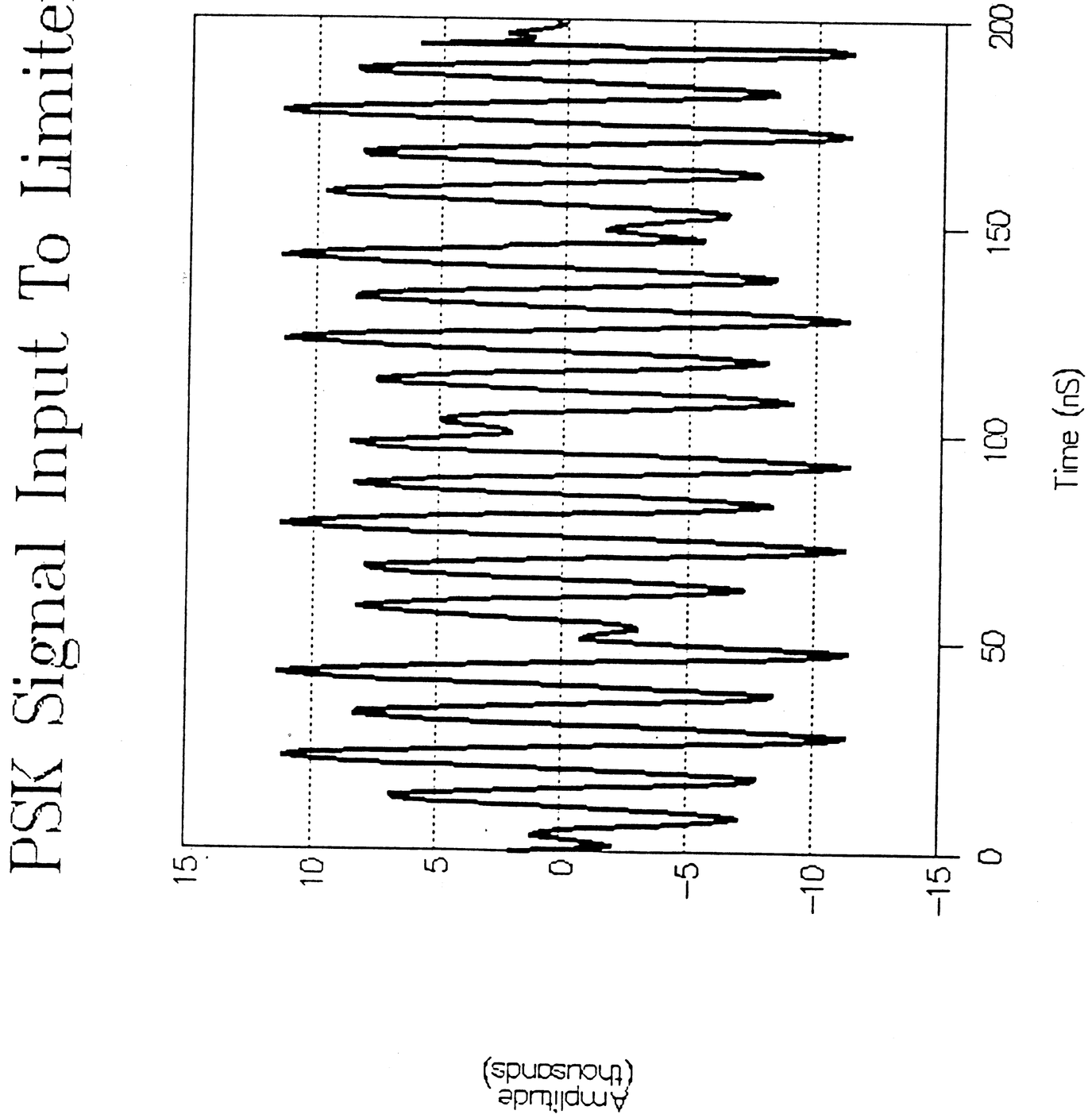

PSK Signal Input To Limiter
Amplitude (thousands)
15
10
5
0
-5
-10
-15
0
50
100
150
200
Time (nS)

Figure 3-3

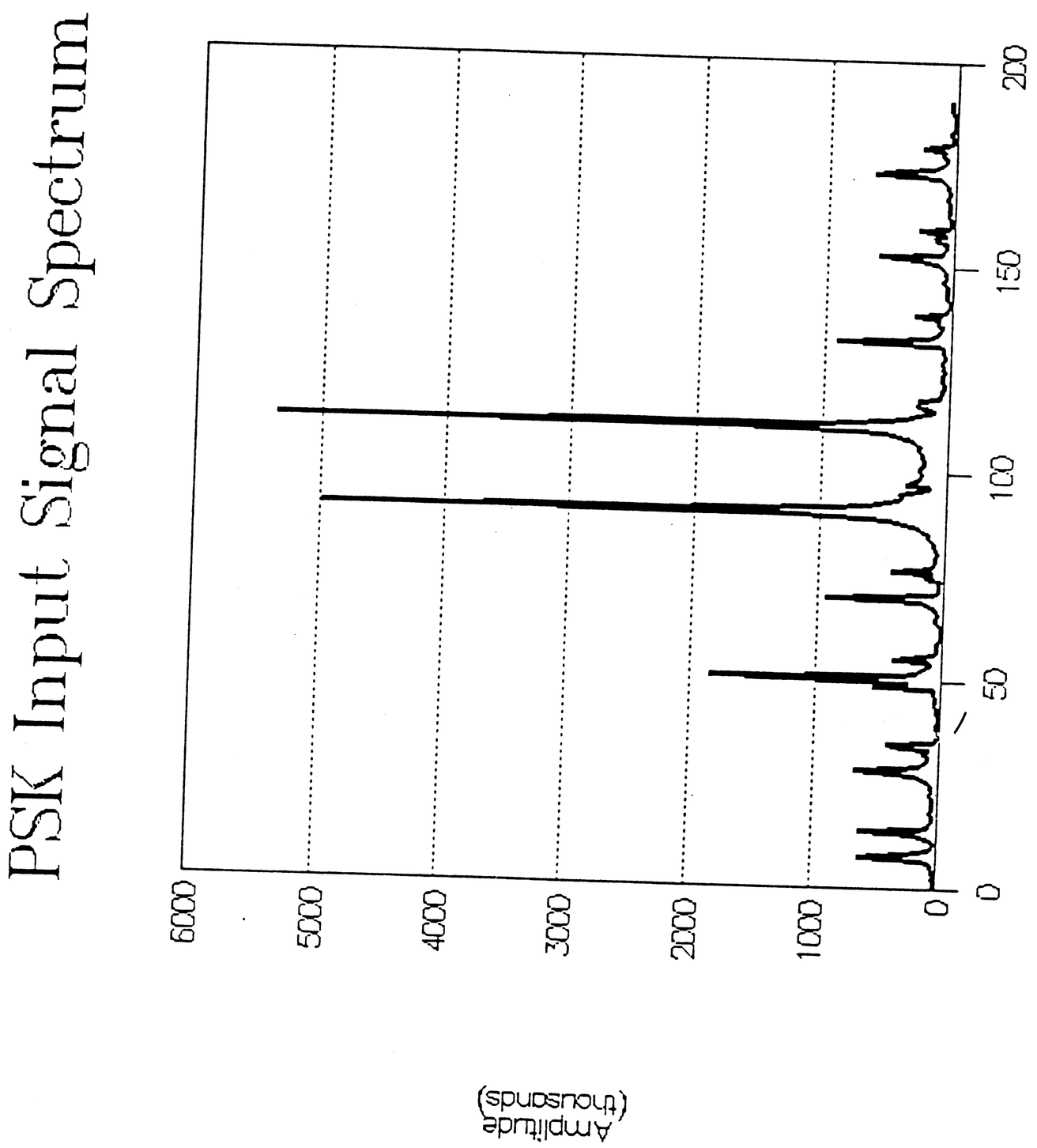

PSK Input Signal Spectrum
Amplitude (thousands)
6000
5000
4000
3000
2000
1000
0
0
50
100
150
200
Frequency (MHz)

Figure 3-4

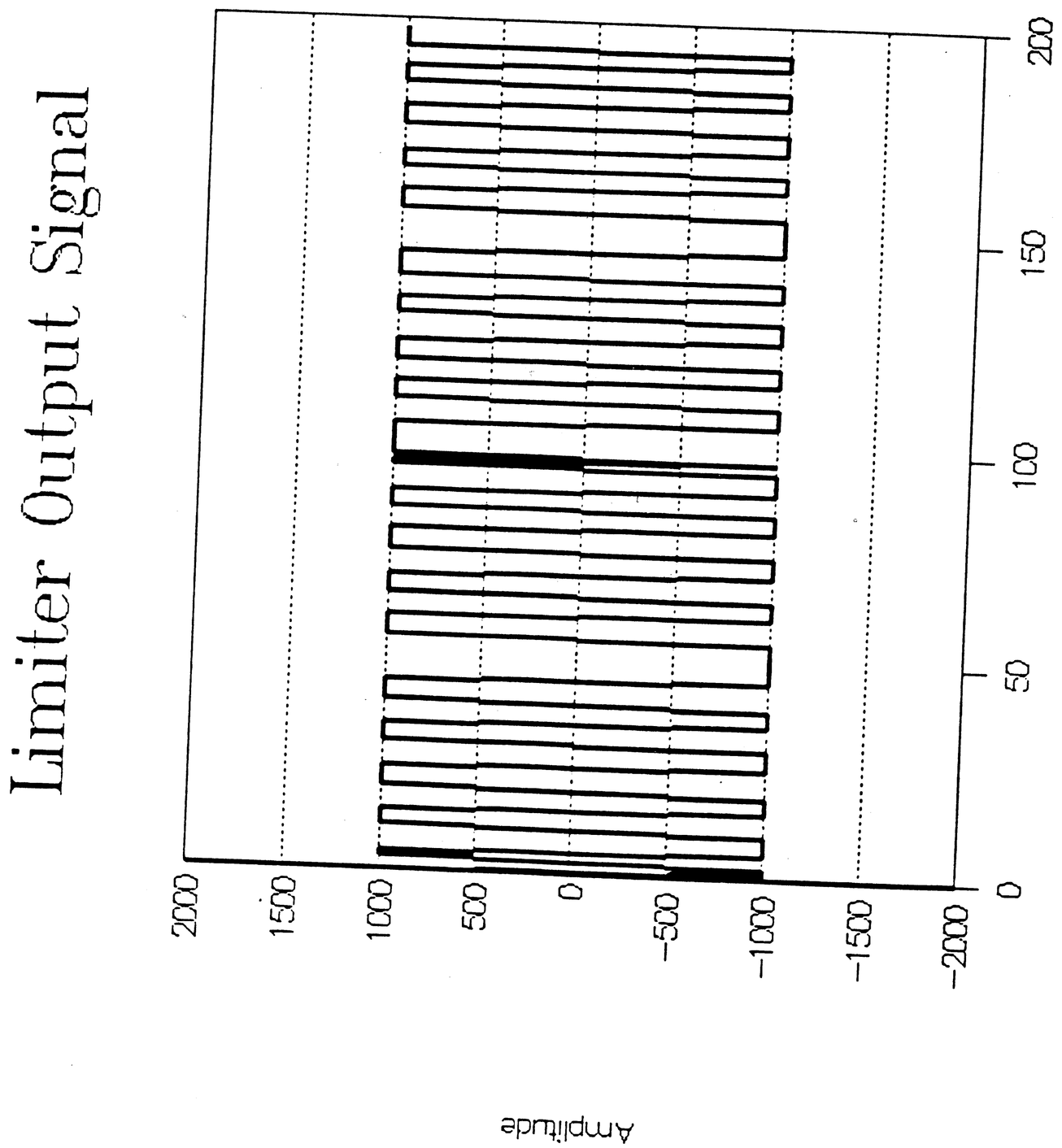

Limiter Output Signal
Amplitude
2000
1500
1000
500
0
-500
-1000
-1500
-2000
0
50
100
150
200
Time (nS)

Figure 3-5

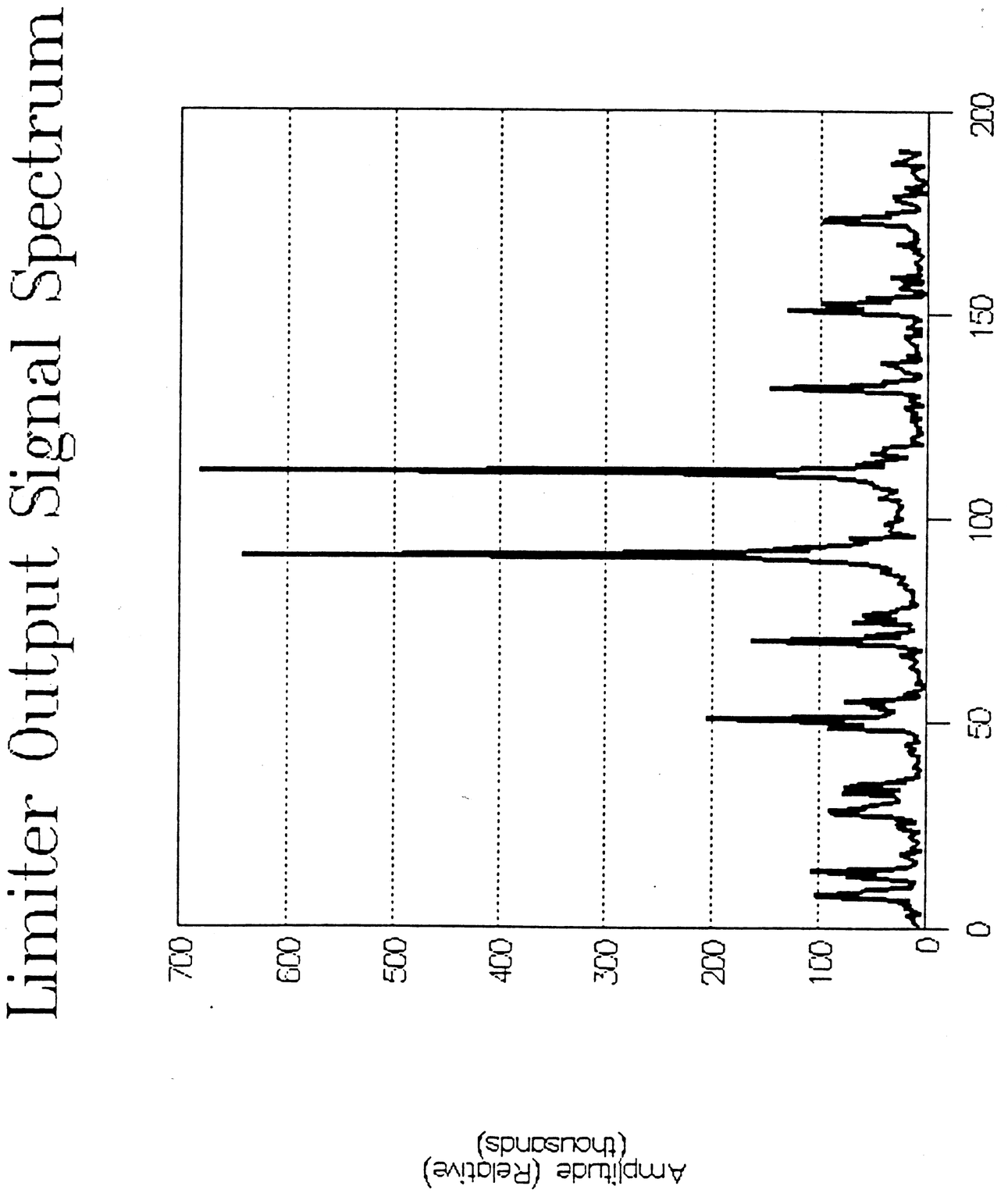

Limiter Output Signal Spectrum
Amplitude (Relative)
(thousands)
700
600
500
400
300
200
100
0
0
50
100
150
200
Frequency (MHz)

Figure 3-6

components at the output. However, the case of two sinusoids of different frequencies and amplitudes passing through a limiter has been analyzed (Jones). The results show that if the signals differ in amplitude by at least a factor of 5, the SNR's of two signals' SNR's are not affected by more than 3 dB (Jones).

3.3.1.4 Conclusions from Limiter Analysis

From this analysis, we can conclude that limiters do not appreciably distort the signal spectrum when they are followed by proper filtering around the frequency components of interest. Also, limiters (for the most part) do not appreciably affect the signal to noise ratio of the signal by more than 3 dB. Therefore, limiters are useful in eliminating AGC's while not appreciably distorting the signal.

3.3.2 Mixers

The multiplication operation is a fundamental operation for modulation and frequency translation in any radio frequency circuit. This multiplication operation is implemented by a device called a *mixer*, because it mixes two of the input frequencies.

A mixer can be implemented in several ways. One such method, termed the double-balanced mixer, is shown in figure 3-7. It consists of a ring of diodes and input and output transformers. Mixers of this type can be designed to work into the microwave frequency range, since the device design is passive and simple. Indeed, the mixers

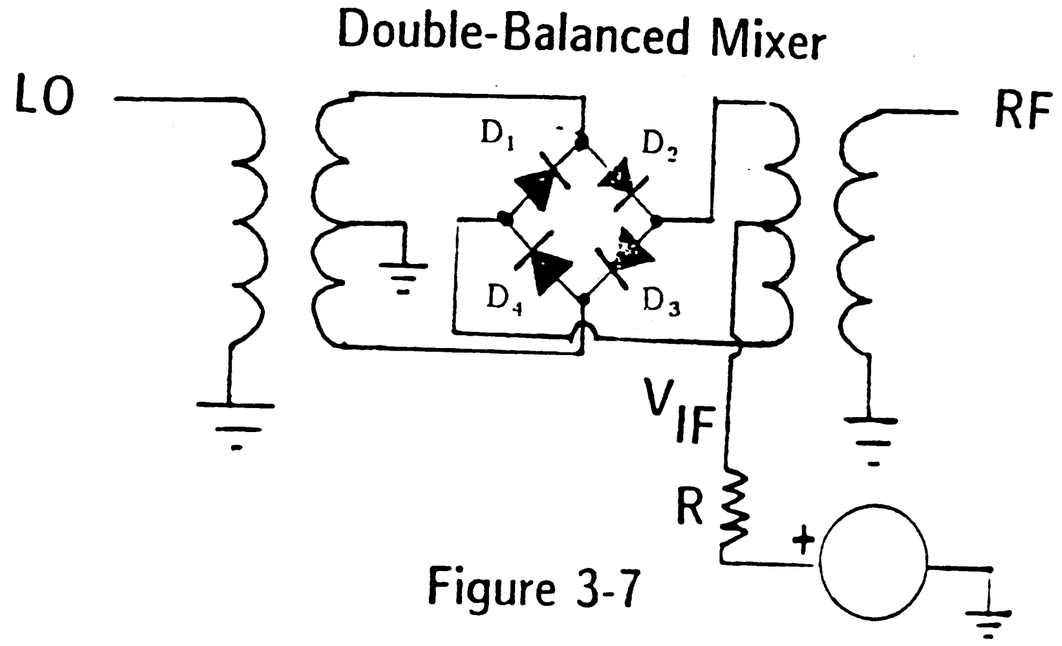


Figure 3-7

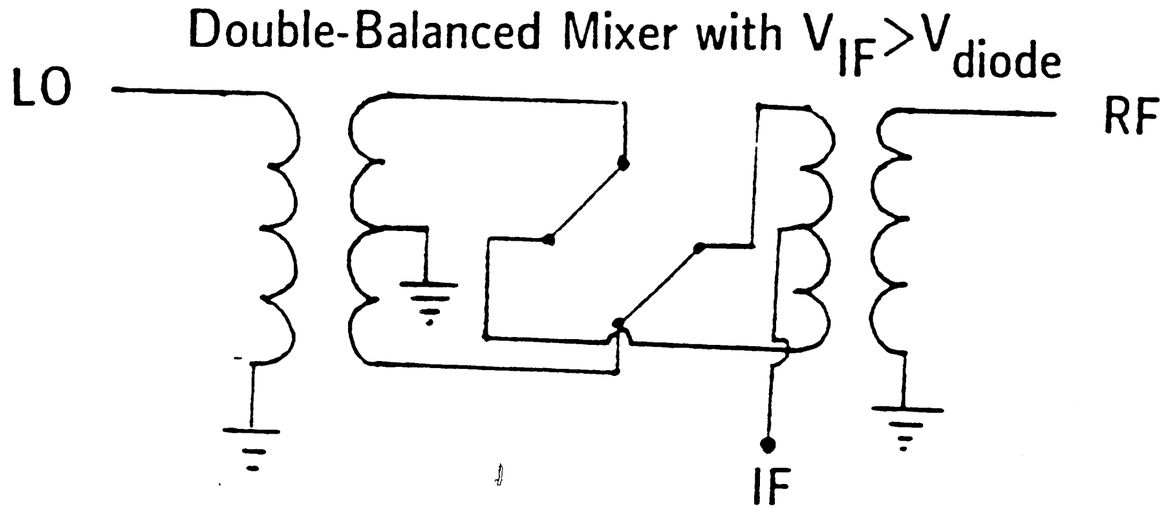


Figure 3-8

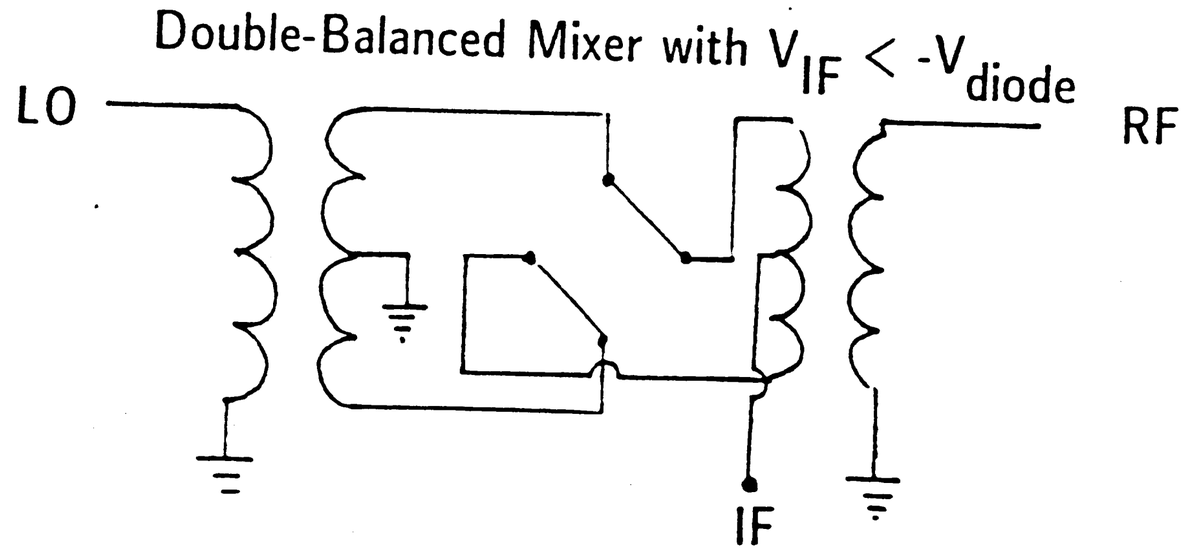


Figure 3-9

used in this project are ASK1 Mixers made by Mini-Circuits, and are designed to work from DC to 600 MHz. They are designed with 50 ohm input and output ports.

The operation of this circuit can be demonstrated in figure 3-7. The mixer has 3 ports, the LO port, the IF port, and the RF port. The LO port is the input for the local oscillator, which is generally a carrier frequency to be modulated. The IF port is the 'intermediate frequency' port, and is used to modulate the carrier. The RF port is the radio frequency output port, which outputs the modulated signal.

In order to demonstrate how the mixer operates, consider the case shown in figure 3-7, where a voltage source and a resistor are connected to the IF port. If this voltage source is higher than approximately 0.7 volts, diodes D1 and D3 will turn 'on', and the circuit is equivalent to that in figure 3-8. Similarly, if the voltage is less than approximately -0.7V, then D2 and D4 are turned 'on', and the mixer circuit is equivalent to figure 3-9. The output voltage from the mixer in these two different cases are the opposites of each other. Thus, if $V_{IF}>0$ then $V_{RF}=V_{LO}$, else if $V_{IF}<0$ then $V_{RF}= -V_{LO}$.

Now, if V_{IF} is a function of time, i.e. $V_{IF}(t)$, then the output waveform is

$$V_{RF}(t)=V_{LO}(t) \cdot \mathrm{sgn}(V_{IF}(t))$$

where sgn(x) = 1 if x>0, -1 if x<0. Note that this fact is useful for binary phase shift keying where 180° phase reversals are needed.

If $V_{IF}(t)=\cos(\omega_{IF}t)$, and $V_{LO}(t)=\cos(\omega_{LO}t)$, then the output voltage is

$$V_{RF}(t) = \cos(\omega_{IF}t) \cdot \mathrm{sgn}(\cos(\omega_{LO}t))$$

Note that $\mathrm{sgn}(\cos(\omega_{LO}t))$ is a square wave of frequency $\omega_{LO}/2\pi$ Hertz. Since a square wave can be represented by its Fourier series, i.e.

$$\mathrm{sgn}(\cos(\omega_{LO}t)) = \tfrac{4}{\pi}\cos(\omega_{LO}t) - \tfrac{4}{3\pi}\cos(3\omega_{LO}t) + \tfrac{4}{5\pi}\cos(5\omega_{LO}t)\ldots$$

Thus, the output is

$$V_{RF}(t) = \cos(\omega_{IF}t) \cdot \left(\tfrac{4}{\pi}\cos(\omega_{LO}t) - \tfrac{4}{3\pi}\cos(3\omega_{LO}t) + \tfrac{4}{5\pi}\cos(5\omega_{LO}t)\ldots \right)$$

$$= \tfrac{4}{\pi}\cos(\omega_{IF}t)\cos(\omega_{LO}t) - \tfrac{4}{3\pi}\cos(\omega_{IF}t)\cos(3\omega_{LO}t) + \tfrac{4}{5\pi}\cos(\omega_{IF}t)\cos(5\omega_{LO}t)\ldots$$

Note that the first term of the second series is the desired product; the other terms in the series can be filtered out by using a bandpass filter centered around ω_{LO} after the mixer. The result is

$$V_{RF}(t) = \tfrac{4}{\pi}\cos(\omega_{IF}t)\cos(\omega_{LO}t)$$

The above analysis used the assumption that the diodes are either 'on' or 'off'. In practice, diodes have a nonlinear exponential IV characteristic; thus the diodes must be driven with signals large enough to sufficiently turn 'on' the diodes.

Also note that the diodes of the above circuit are assumed to be exactly the same. If

the diodes are perfectly matched in a doubly balanced mixer, each port is totally isolated from the other ports (Ziemer). However, if the diodes are mismatched, some isolation between the RF, IF, and LO ports is lost . For typical mixers, such as the Mini-Circuits ASK1 mixer, the diodes are matched so as to achieve more than 30 dB isolation between the ports.

3.3.3 Mixers as Phase Detectors

We can use a mixer to detect the relative phase between two signals. Phase detectors are useful in phase locked loops, where an oscillator is adjusted in frequency/phase until the relative phase between two signals is zero.

To see how phase detection is done, consider the following. If we have two sinusoidal signals of identical frequency that differ only in phase by ϕ radians, then we can form a "phase error signal" e(t) by multiplying the two signals together with a mixer:

$$e(t) = \cos(\omega t + \phi) \cdot \cos(\omega t) = \tfrac{1}{2}\cos(2\omega t + \phi) + \tfrac{1}{2}\cos(\phi)$$

The $2\omega t$ frequency component can be filtered out with a low pass filter, leaving only the last term, $\cos(\phi)/2$. Using this fact, we can construct a phase detector circuit shown in figure 3-10. The two sinusoidal signals are input to a mixer/multiplier, the output of the mixer is passed through a low pass filter to remove the 2ω term, and the resultant output is proportional to $\cos(\phi)$. The output vs. phase error graph is shown in figure 3-11. Note that the output of the phase detector is 0 when the two

Mixer as Phase Detector

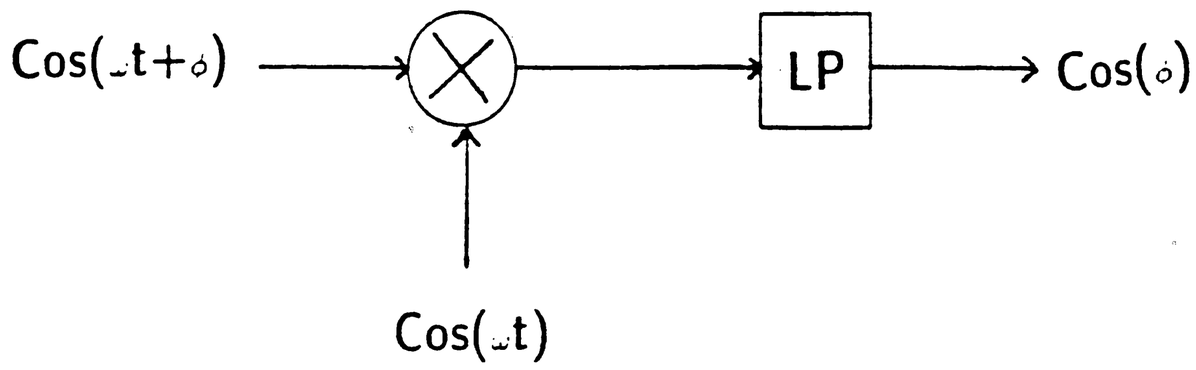


Figure 3-10

Phase Detector Output vs. Phase Error

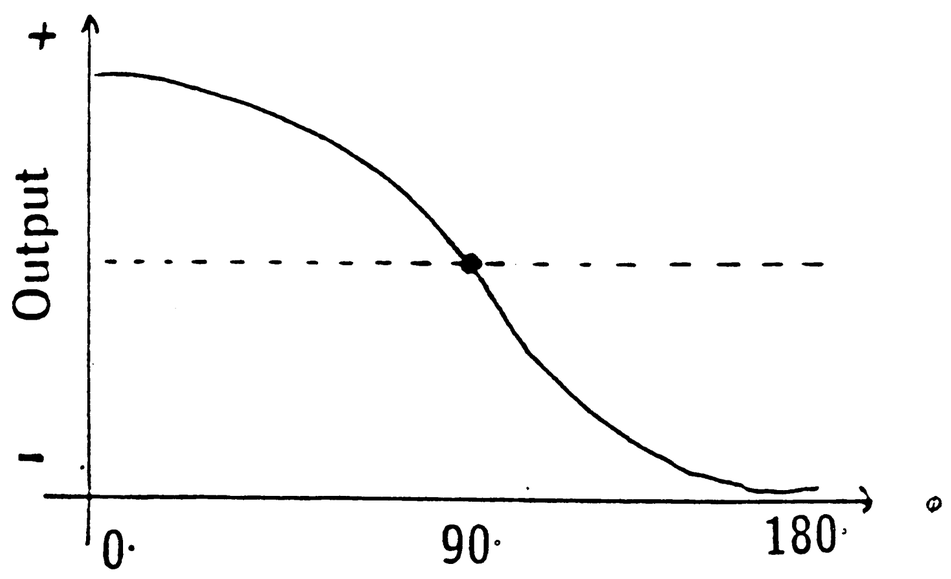


Figure 3-11

signals are 90° out of phase.

3.3.4 Frequency Doublers

A full wave bridge rectifier can serve as a frequency doubler, as shown in figure 3-12. The doubler has a similar structure to a mixer, with the IF port removed and the diodes D2 and D4 reversed. If the input signal is above 0.7 V, diodes D1 and D3 conduct the positive half-cycle so that current flows through D1, through the output transformer, and back through D3. On the negative half-cycle, D2 and D4 conduct, and the current flows in the same direction through the output transformer. A rectified sinusoid is the result, and it is AC coupled to the output through the output transformer.

The Fourier series for a rectified sine wave r(t) is

$$r(t) = \frac{4}{1 \cdot 3\pi} \cos(2\omega t) - \frac{4}{3 \cdot 5\pi} \cos(4\omega t) + \frac{4}{5 \cdot 7\pi} \cos(6\omega t)...$$

where ω is the frequency of the original sine wave. To isolate the doubled frequency component, we use a bandpass or lowpass filter after the doubler to eliminate all but the 2ω term, as shown in figure 3-13. The output of the filter is a sinusoid at twice the input frequency to the doubler.

3.3.5 Phase Locked Loop Fundamentals

Phase locked loops force a local oscillator to track an input waveform in phase and

Frequency Doubler

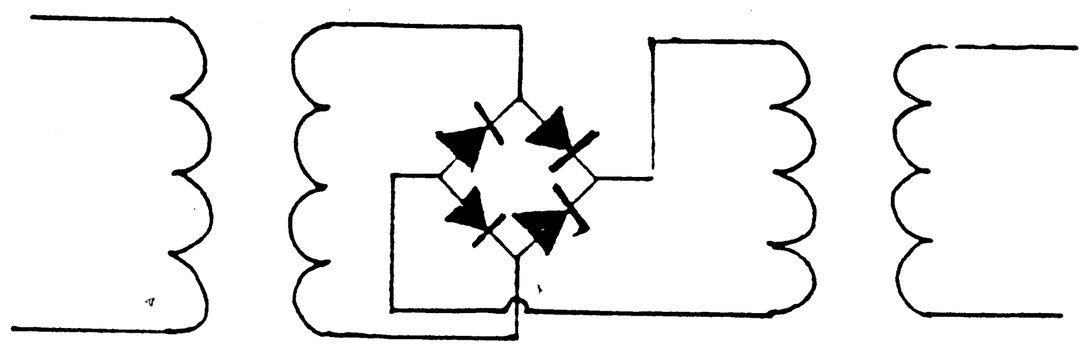


Figure 3-12

Use of Frequency Doubler

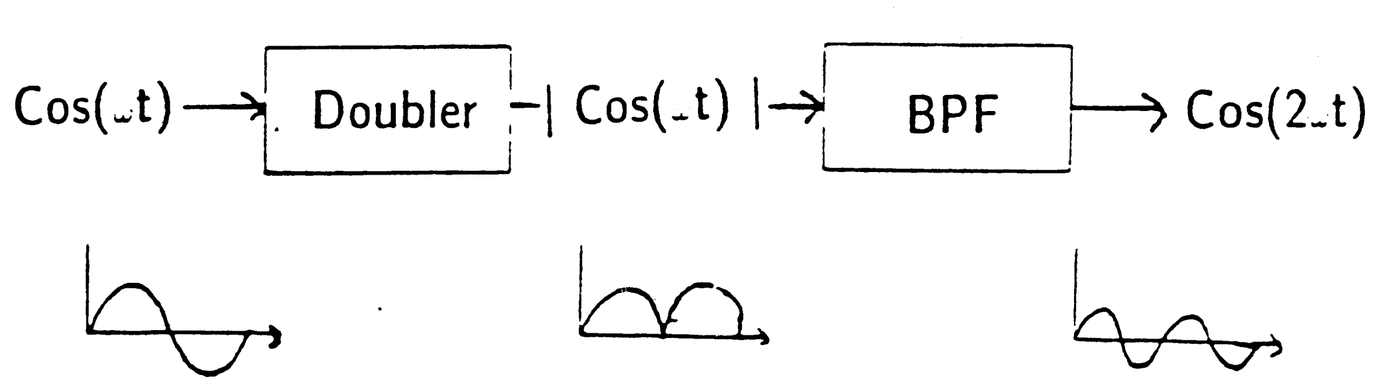


Figure 3-13

frequency. A phase-locked loop (PLL) following the 50 MHz filter in this system is used to lock on to the 50 MHz reference signal. The PLL acts as a very narrow bandwidth (200 Hz) filter, thus improving the 50 MHz reference signal SNR by rejecting out-of-band noise.

A block diagram of a typical PLL is shown in figure 3-14. Phase locked loops operate by forcing a local oscillator's frequency to match the input signal frequency. In the locked condition, any frequency or phase change in the input signal first appears as a change in phase between the input signal and the local oscillator. This phase shift then acts as an error signal and is used to change the frequency of the local oscillator in order to match the input signal. The locking onto a phase relationship between the input signal and the local oscillator accounts for the name phase-locked loop.

The phase comparator is perhaps the most important part of the PLL system since it is here that the input and local oscillator frequencies are compared. Here, a multiplier is used to mix the input and local oscillator signals. If there is a phase/frequency error between the two signals with phase and frequency ω_{in}, ϕ_{in}, ω_{VCO}, ϕ_{VCO} respectively, the resulting error waveform is

$$
\begin{aligned}
V_{Err}(t) = \; & 2\cos\left(\omega_{in}\, t + \phi_{in}\right) \cdot \cos\left(\omega_{VCO}\, t + \phi_{VCO}\right) \\
= \; & \cos\left((\omega_{in}-\omega_{VCO})\, t + \phi_{in} - \phi_{VCO}\right) + \\
& \cos\left((\omega_{in}+\omega_{VCO})\, t + \phi_{in} + \phi_{VCO}\right)
\end{aligned}
$$

In a PLL, the low pass filter is used to eliminate the frequency component at

Phase Locked Loop

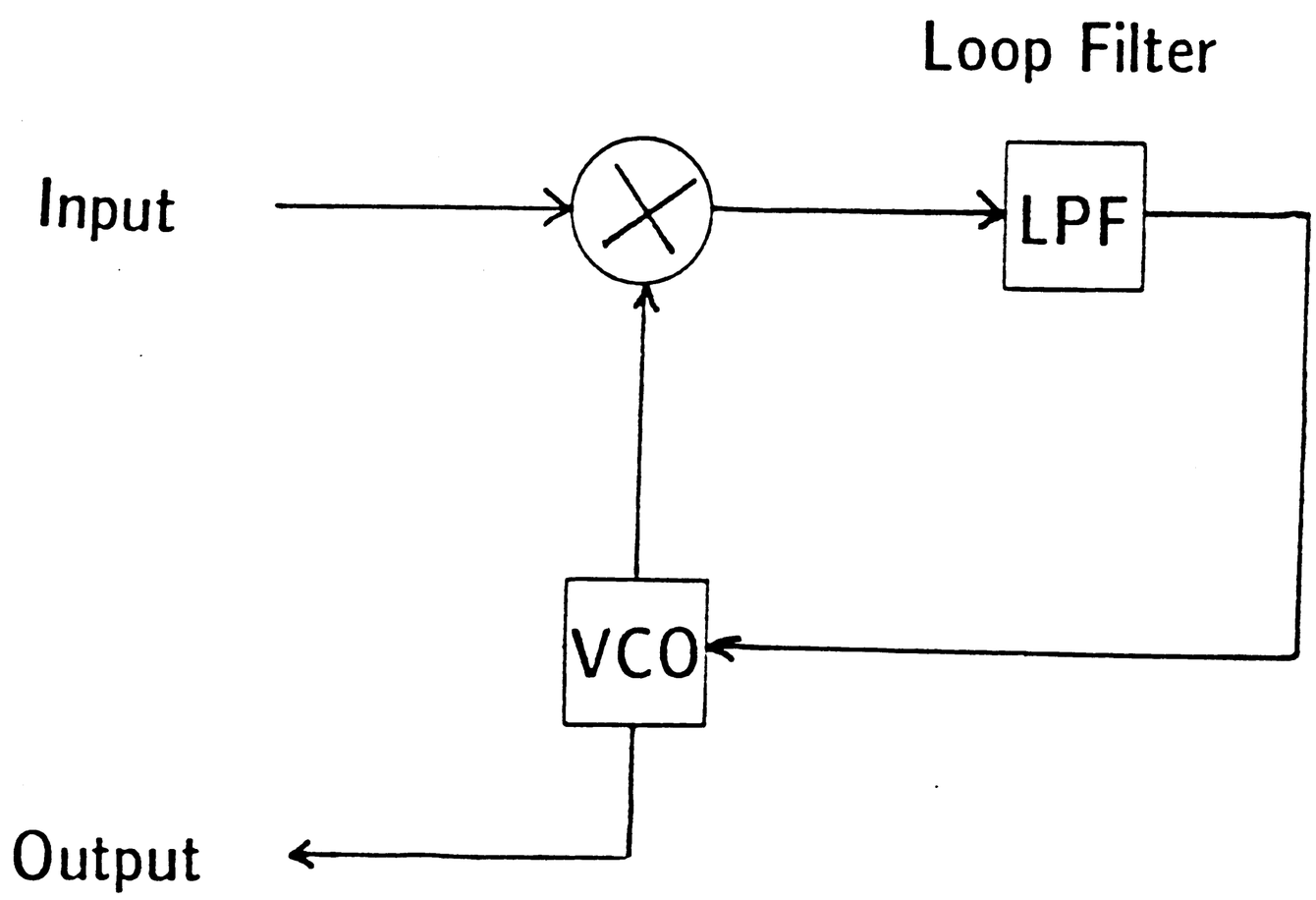


Figure 3-14

$\omega_1+\omega_2$. The resulting waveform is

$$V_{VCO}(t) = \cos\Big((\omega_{in}-\omega_{VCO})\,t + \phi_{in} - \phi_{VCO}\Big)$$

When the loop is in lock, the local oscillator duplicates the input frequency so that the difference frequency component is zero. Hence, the output of the low pass filter contains only a DC component reflecting the phase error between the input and local oscillator signals. This error signal then controls the the local oscillator's frequency. The local oscillator is usually a voltage controlled oscillator, with frequency determined by the relation :

$$\omega_{VCO} = K \cdot V_{VCO}(t) + \omega_{FR}$$

where ω_{FR} is the free running frequency of the oscillator, and K is a constant (in radians/sec per volt) determined by the PLL circuit construction.

The lock range of a PLL is the range of frequencies over which the loop will remain in lock while the input frequency is changed. Although the loop will remain in lock in this range, it many not be able to initially *acquire* lock over the entire lock range. The capture range is the range of frequencies in which the PLL can acquire lock. Capture range is always smaller than the lock range, and is determined by the bandwidth of the low pass filter. Furthermore, the low pass filter bandwidth is proportional to the maximum rate at which the input frequency can change while remaining in lock. The transient time required for a free running loop to lock is called the lock up time, and is inversely proportional to the low pass filter bandwidth. The locking transient is highly complex and does not lend itself to

simple mathematical analysis. A complete modeling of phase locked loops from a control system standpoint can be found in (Best).

4. CIRCUIT CONSTRUCTION AND ANALYSIS

4.1 Introduction

Using the modulator/demodulator system design detailed in Chapter 3, the complete circuitry for an atmospheric laser communications system was constructed. Circuitry for the modulator and laser driver was developed for the transmitter, and circuits for the photodetector, preamplifier, and demodulator were developed for the receiver. Furthermore, the signal to noise ratio of the receiver was analyzed in order to optimize the receiver's performance.

As with any radio frequency circuit, careful electrical construction techniques must be followed to ensure proper operation and to prevent unwanted radiation. The circuitry here was constructed using printed circuit boards, with each subcircuit in a shielded box, and with shielded cables carrying signals between the subcircuits. Furthermore, the liberal use of bypass capacitors on the power supply rails in each subcircuit prevents stray RF from entering any subcircuit undesirably.

4.2 Transmitter

The detailed schematic diagrams of the PSK transmitter circuitry are shown in figures 4-1 and 4-2. Portions of these diagrams correspond to the 50 MHz oscillator, frequency doubler, filter, mixer, combiner, and laser driver blocks in the system diagram of figure 3-1.

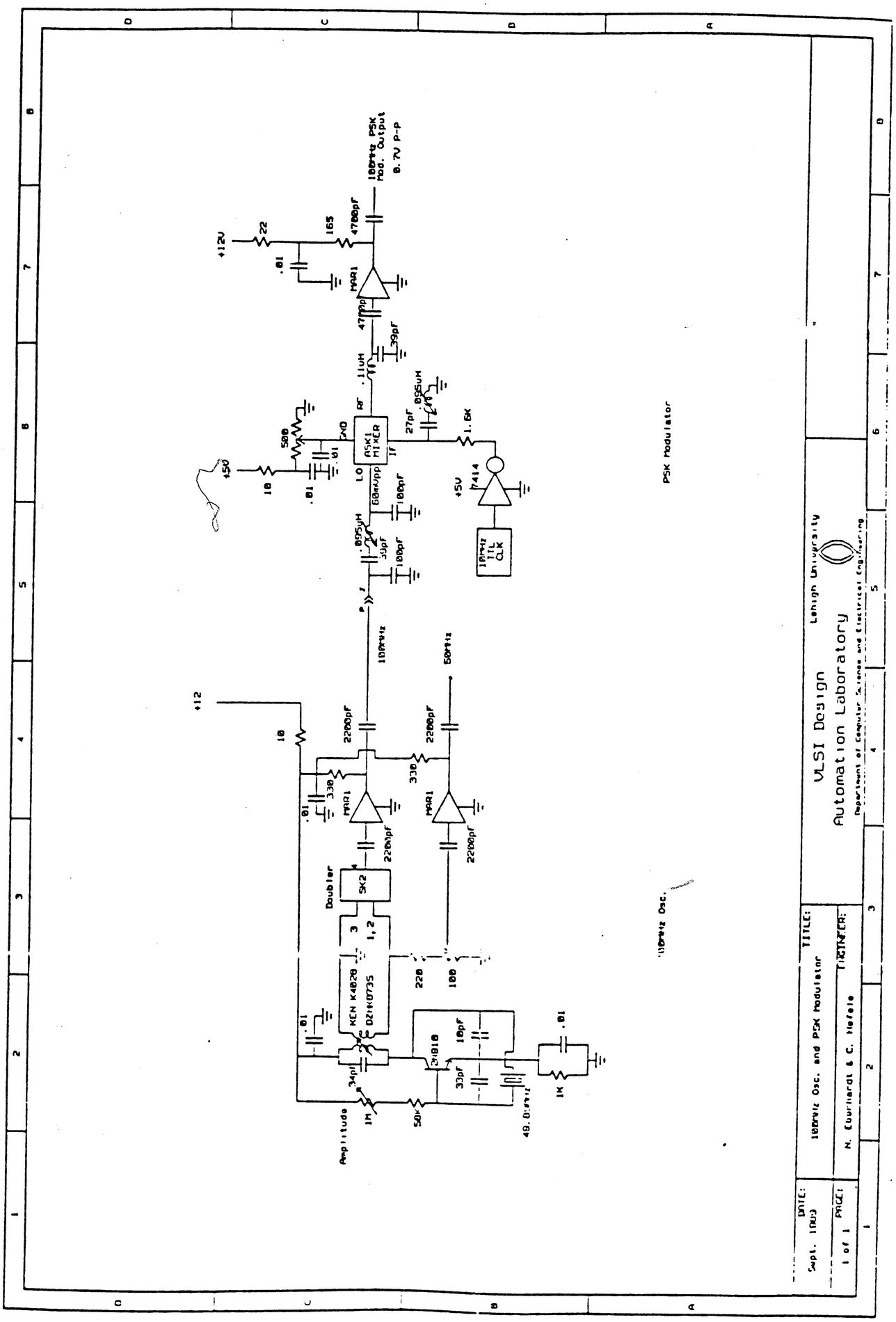

PSK Modulator
Lehigh University
VLSI Design
Automation Laboratory
Department of Computer Science and Electrical Engineering
TITLE:
100MHz Osc. and PSK Modulator
DRAWN BY:
N. Eberhardt & C. Hofele
DATE:
Sept. 1993
PAGE:
1 of 1
+12V
22
165
.01
4700pF
MAR1
180MHz PSK
Mod. Output
8.7V p-p
47pF
39pF
.11uH
RF
MIXER
ASK1
LO
IF
500
.01
+5V
18
60mVpp
100pF
.095uH
3pF
27pF
.095uH
1.6K
74F14
10MHz
TTL
CLK
180MHz
50MHz
2200pF
330
MAR1
Doubler
SK2
KEN K4828
DZ1K0735
2N918
10pF
33pF
.01
340pF
1M
50K
1K
49.0MHz
Amplitude
100MHz Osc.

Figure 4-1

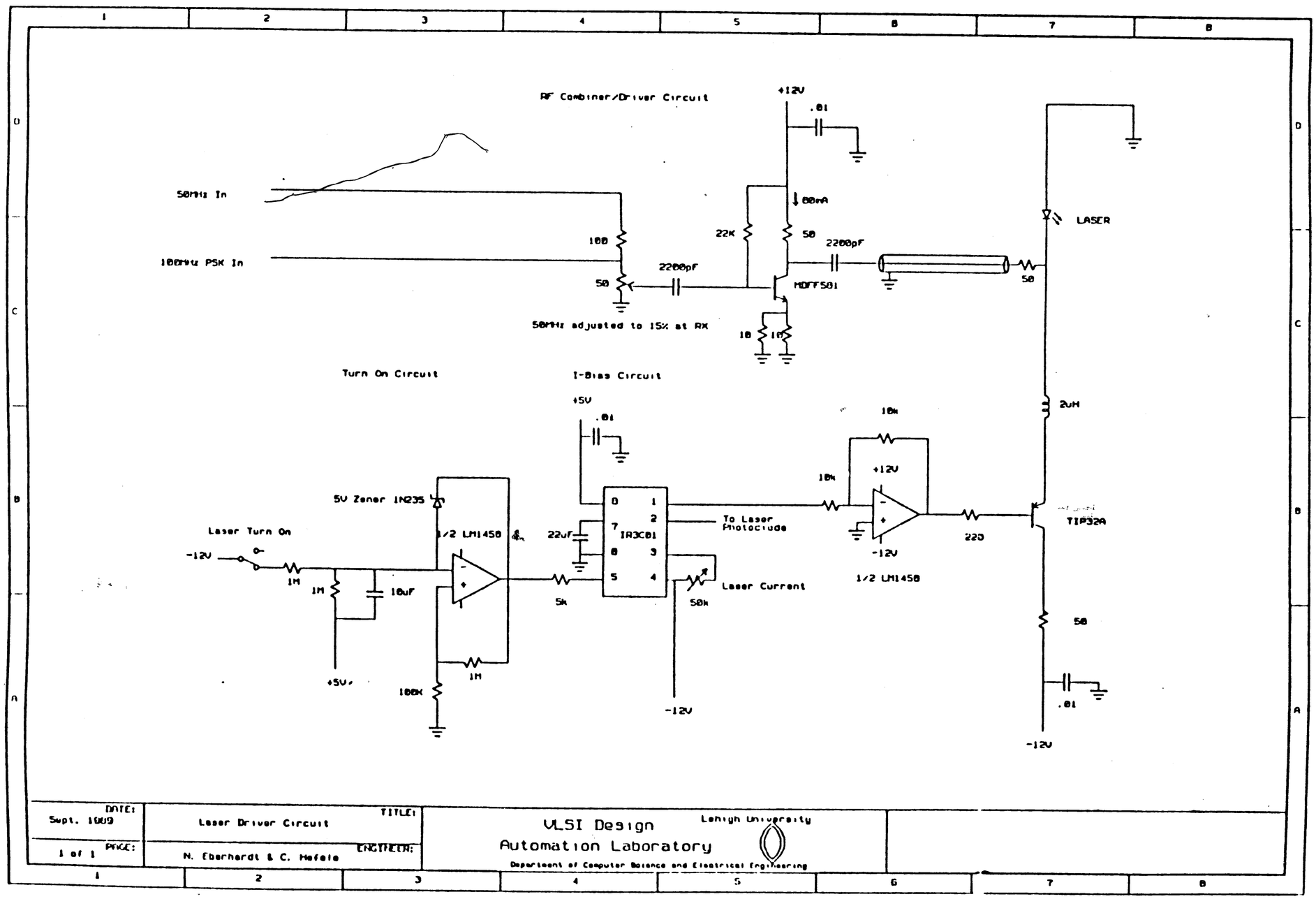

RF Combiner/Driver Circuit
+12V
.01
50MHz In
100MHz PSK In
100
50
2200pF
22K
50
2200pF
80mA
MOFF501
LASER
50
50MHz adjusted to 15% at RX
18
10
Turn On Circuit
I-Bias Circuit
+5V
.01
5V Zener IN235
1/2 LM1458
Laser Turn On
-12V
1M
1M
10uF
+5V
1M
100K
22uF
IR3C01
To Laser Photodiode
Laser Current
5k
50k
-12V
10K
+12V
-12V
1/2 LM1458
220
2uH
TIP32A
50
.01
-12V
DATE:
Sept. 1989
TITLE:
Laser Driver Circuit
PAGE:
1 of 1
ENGINEER:
N. Eberhardt & C. Hefele
VLSI Design
Automation Laboratory
Lehigh University
Department of Computer Science and Electrical Engineering

Figure 4-2

4.2.1 50 MHz Oscillator Circuit

The 50 MHz sinusoidal oscillator is used to provide a 50 MHz HF carrier source for the system. The oscillator is shown in the left section of figure 4-1. It is a modified Colpitts oscillator, with the 1M potentiometer and 56K resistor setting the DC bias current flowing through the 2N918 high frequency transistor. The collector output of the oscillator is fed back to the base through the crystal as well as through a capacitive voltage divider (formed by the 10pF and 33pF capacitors). The crystal, operating in it's third overtone, and tuned output transformer both force the oscillator to oscillate at 49.89 MHz. The oscillator output signal level at the output transformer is about 1.5V peak to peak, which is approximately +8dBm.

4.2.2 Frequency Doubler

The frequency doubler is used create the 100 MHz PSK carrier from the 50 MHz oscillator output. The 50 MHz oscillator output is fed to the Mini-Circuits SK2 frequency doubler, which operates from 1 to 500 MHz with a conversion loss of 13dBm. The output level of the doubler is -5dBm, and the signal must be amplified by the MAR1 integrated 50Ω microwave amplifier for use by the next stage.

4.2.3 50 MHz Carrier Signal

The 50 MHz carrier is injected at low level to provide the receiver with a carrier reference signal. Connected at the oscillator transformer output is a voltage divider consisting of a 220 ohm resistor and 100 ohm potentiometer. By adjusting the pot,

the level of the 50 MHz carrier ultimately added to the 100 MHz signal level can be adjusted from 1/10 to 1/5 of the 100 MHz signal's level.

4.2.4 Filter

In order to provide a clean sinusoid for the modulator, harmonics generated by the frequency doubler and harmonics generated by amplifier clipping must be removed. The pi-configuration passive filter to do this consists of a 39pF capacitor and .095μH inductor tuned to resonate at approximately 100 MHz. The output of the filter is a clean 100 MHz sinusoid which drives the frequency mixer in the next stage.

4.2.5 Mixer

In order to generate a binary phase shift keying signal, the 100 MHz signal must be multiplied/mixed with *bipolar* digital data, i.e. ±2.5 volts, not 0V/5V. This is because if we have a random string of bits with 1's and 0's represented as ±2.5V, their mean DC value is 0, thus there will be no carrier component in the output signal. However, in order to convert the normal 0 to 5V TTL levels to ±2.5V, the ground lead of the mixer must be 'floated' at 2.5V. Thus, a 0V TTL level is -2.5V relative to mixer ground, and 5V is +2.5V relative to mixer ground. The 2.5V mixer ground potential is obtained from the tap of the 500 ohm potentiometer, and is adjusted for best waveform symmetry at the output of the PSK modulator.

For test purposes, two bit sources were used. First, a 10 MHz square wave TTL clock was used as a 20 Mb test signal, as shown in the schematic. The second test

signal used consisted of a digital circuit to generate the 20 Mb sequence 1010101011111111. These bit sources drive the 7414 TTL inverter, which acts as a buffer amplifier between the digital source and the analog mixer. The 1.6K resistor limits the current through the mixer diodes to prevent damage to the mixer. The resonator consisting of the 27pF capacitor and .095μH inductor at the IF input short to ground any 100 MHz components in the modulating square wave. If present, the beat frequency between these 100 MHz components and the local oscillator would modulate the output signal undesirably.

The output of the mixer must be amplified and filtered in order to be used by the next stage. The band pass filter consisting of the .11μH inductor and 39pF capacitor eliminates signal harmonics created by the mixer. The MAR1 amplifier amplifies the RF output of the filter, which is a BPSK signal centered at 100 MHz.

4.2.6 Laser

The output of the modulator is used to modulate a laser which is collimated and directed at the optical receiver. The particular laser used here is a semiconductor laser, the Sony SLD202 V-3, a 50mW High Power Infra-Red Laser Diode with a radiant output wavelength of approximately 820 nm.

4.2.6.1 Laser Electrical Characteristics

A semiconductor laser is essentially an optical resonator created with semiconductor materials. Due to semiconductor characteristics, however, the amplitudes of the

different modes of this resonator are dependant on the currents passed through the laser junction. Below a certain threshold current, several modes exist; at higher current levels, a single mode dominates. (Jones) In terms of device terminal electrical characteristics, the laser acts just like a typical LED when currents through the laser are below a certain threshold current, here 90 mA. Above this threshold current, a single mode dominates, the device begins its' lasing action, and the radiant power output level rises quickly.

In practice, the laser output level will vary with temperature changes caused by environmental changes or device self-heating. In order to avoid this, a PIN photodiode is included in the laser package for monitoring the laser radiant power output. The output level sensed by the PIN diode can be used in a closed loop control system to automatically adjust the laser driving currents so as to give a constant light output from the laser and thus preventing "thermal runaway."

4.2.6.2 Laser Driver Circuit

The laser driver circuit provides bias and level control circuitry so as to properly modulate a semiconductor laser. The circuit has two main sections; the RF combiner/AC driver subcircuit, and the DC bias subcircuit.

4.2.6.3 DC Bias Subcircuit

The DC bias circuit consists of the turn-on circuit, the bias current circuit, and the bias transistor and transistor driver. These are shown in the lower half of figure 4-2.

The circuit centers around the Sharp IR3C01 chip, which is an integrated circuit specifically designed for closed-loop laser control. This chip monitors the internal laser PIN diode to detect the output light level, and adjusts the output current from pin 1 to keep the laser power output constant.

Pin 5 is a laser enable TTL input connected to a turn-on circuit which provides a delay of several when the power is first applied. This delay in turning on the laser prevents initial power surges from affecting the laser output, as well as allowing the user of the system to move away from the beam. Such a time delay is required by law, as is a key switch for turning on the laser system. Once the laser turn-on switch is thrown, the 10μF capacitor charges through the 1MΩ resistor. The following op amp stage is configured as a comparator, and the 5V zener diode in the feedback loop limits the output high voltage to 5 V. In several seconds, the rising capacitor voltage changes the comparator output +5V.

The variable resistor between pin 3 and 4 sets the desired level of laser current. The laser output at pin 1 connects to an inverting op amp, with unity gain. This amplifier is used to invert the positive voltage signal from the IR3C01 chip, so that it can properly drive the PNP TIP32A transistor. This transistor controls the laser's DC bias current, which is set at approximately 125mA.

In order to prevent the injected 100 MHz or 50 MHz AC components from changing the DC bias of the circuit, the 2 μH inductor in series with the laser acts as a RF choke, preventing the HF components from passing through the transistor. Since, at 100 MHz, the inductor's reactance is over 1KΩ, thus most of the RF currents will

flow through the 50 ohm load.

4.2.6.4 RF Combiner / AC Driver Subcircuit

Both the 50 MHz and 100 MHz signals are combined at the input of a MDF581 HF transistor, which can be seen in the top of figure 4-2. Once the DC bias current is set at 125mA, the transistor can fully modulate the laser between the laser threshold current of approximately 100mA and the laser maximum current of 150mA. The 50Ω variable resistor in the transistor base is used to set the AC output current level of the amplifier.

The output impedance of the transistor amplifier is 50 ohms. To minimize reflected power, the amplifier source impedance and laser load impedance must be matched. In order to provide a properly matched load for the amplifier, the laser circuit must be constructed so that it's equivalent impedance is 50Ω. Since the laser has a dynamic resistance only several ohms, a 50Ω resistor is placed in series with the laser in order to increase it's impedance to approximately 50Ω. A 50Ω coaxial transmission line connects the 50Ω output of the amplifier to the 50Ω resistor and laser.

4.3 Receiver

The function of the receiver circuitry is to detect, amplify, and demodulate the received optical signal. The circuitry to do this can be divided into the front-end subcircuits and the demodulator subcircuits.

4.3.1 Receiver Front End

Of most importance to the optical receiver are the photodetector and the following low noise preamplifier. Together, these two elements dictate the receiver sensitivity, noise characteristics, and overall performance. Sensitivity is a function of both the photodetector and the following preamplifier. The photodetector and the amplifier can compensate each other in performance. that is, a good photodetector will alleviate the low noise requirements on the amplifier, and vice versa. However, in order to achieve the ultimate performance in receiver sensitivity, both the photodiode and the amplifier have to be equally good (Van Muoi).

4.3.1.1 Avalanche Photodetector (APD)

Since only a small portion of the laser output is focused on the photodetector by a lens, a sensitive photodetector must be used. Avalanche photodetectors made from different materials provide high sensitivity, but have different noise and temperature characteristics that must be considered.

4.3.1.1.1 APD Fundamentals

APD photodetectors are diodes that convert incident light into current. When these diodes are reverse biased near their breakdown point, they amplify the small photocurrents by using the avalanche multiplication phenomenon within the photodiode. These devices are useful when the incident light, and thus the

photocurrent, is very small. From the circuit viewpoint, the APD looks like a current source (Personick).

APD's are similar in construction to PIN photodiodes, except that the doping of the junction material is altered, which changes the IV characteristic near the reverse-bias avalanche breakdown (Katzman). While reverse biased near the breakdown voltage, incident photons generate electron-hole pairs which in turn generate multiple secondary electrons-hole pairs in the PN junction.

If an electron has a probability P of creating a secondary electron hole pair, then the average total number of carriers, M, resulting from the initial electron is the sum total of

$$M = 1 + P + P^2 + P^3 + \ldots P^{\infty} = \frac{1}{1 - P}$$

where the first term is the initial carrier, the second is the probability that the injected carrier generates a secondary carrier, the third term is the probability of the secondary carrier generating a third, and so forth (Jones).

P is a very sensitive function of the voltage. In the dark limit it is given by

$$P = \left(\frac{V}{V_{br}} \right)^n$$

where V_{br} is the breakdown voltage, and n is an experimentally determined constant (Jones).

Thus,

$$M(V) = \frac{1}{1 - \left(\frac{V}{V_{br}} \right) n}$$

APD multiplication is similar to the action of a photomultiplier tube. The inherent gain of the APD is determined by the bias voltages; typical average avalanche gains range from 10 to 100. These gains are averages, since not every photon generates *exactly* M electrons. This variation causes what is known as *excess noise.* The *excess noise factor*, F(M), is defined as

$$F(M) = \frac{<M^2>}{<M>^2}$$

and is the ratio of the actual noise to that which would exist if the multiplication process was noiseless (i.e., all photons generate exactly M electrons). The avalanche excess noise factor F(M) is given by

$$F(M)=kM+(1-k)(2-1/M)$$

where k is the ionization constant for the device.

4.3.1.1.2 Dark Current and Multiplication Noise in APD's

With no optical signal present, the main source of shot noise in APD's is due to dark current. Dark current is the sum of the leakage current that flow through the

photodiode when no light is incident on the photodiode. Dark current, I_d, is the sum of the surface leakage currents, I_{du} which do not undergo APD multiplication, and the bulk leakage current, I_{dm}, which is subject to APD multiplication. Thus, the total dark current is given by

$$I_d = I_{du} + M \cdot I_{dm}$$

where M is the avalanche multiplication gain. Because of the avalanche multiplication process, the noise contributions of these two dark current components are significantly different. The power spectral density of the noise current due to dark current is given by

$$\sigma^2 = 2eI_{du} + 2eI_{dm}M^2F(M)$$

where e is the electron charge. F(M) is the avalanche excess noise factor, as above. If the APD is used with a moderate multiplication gain, the $2eI_{dm}M^2F(M)$ term dominates, and the $2eI_{du}$ component can be neglected (Van Muoi).

The avalanche excess noise factor F(M) is given by

$$F(M)=kM+(1-k)(2-1/M)$$

where k is ionization ratio constant of the device. For silicon detectors, a typical value of k is 0.02. For the range $1000>M>1$, and with $k=0.02$, F(M) can be well approximated by $M^{0.25}$. Thus, the power spectral density of the noise due to dark

current, $\sigma_{dm}{}^2$, is well approximated by

$$\sigma_{dm}{}^2 = 2\,e\,I_{dm}\,M^{2.25}$$

Dark current is highly temperature dependant; a good approximation for the temperature coefficient of dark current is that dark current doubles for every 10° C increase in operating temperature. Low temperature operation is desirable, since shot noise due to dark current is the main factor limiting the sensitivity of optical receivers (Van Muoi).

Multiplication of *signal photocurrents* are also subject to excess noise in the multiplication process. In fact, if the photocurrents are many times larger than the dark currents, the noise due to dark currents can be neglected, since the main source of noise in the APD will be due to excess noise in the multiplication of photocurrents. The spectral noise density, $\sigma_p{}^2$, caused by the multiplication of the photocurrent I_p is

$$\sigma_p{}^2 = 2\,e\,I_p\,M^{2.25}$$

where $I_p = R\,N_d$, R being the responsivity of the APD (in A/W), and N_d is the incident optical power (in Watts).

4.3.1.1.3 APD Materials and Device Performance

Currently, the APD's with the best performance are made from silicon. Silicon

devices respond best in the spectral region from approximately 0.5μm to 1.0μm. Significant progress had been made in improving long-wavelength (1.0μm to 1.6μm) detectors made from other materials, however dark currents still remain many times dark currents of silicon APD's. For example, a typical silicon APD may have a total dark current of 1nA, but an InGaAs APD may have a dark current of 10nA. Furthermore, Ge APD's may have dark currents of 100 to 1000nA. Since silicon APD's offer the lowest dark current and thus the least noise, they offer the best performance (by a minimum of 10dB) in the 0.5μm to 1.0μm region. The improvement of long wavelength (1.0-1.6μ) APD's using III-V materials is still under research and development. (Van Muoi)

4.3.1.1.4 APD Temperature Compensation

In practical use, an APD must be temperature compensated. This is because the gain of the APD is dependant on the breakdown voltage, and the breakdown voltage is dependant on temperature. Thus, with temperature compensation, any change in temperature will not change the APD gain if the APD bias voltage is adjusted accordingly. A compensation circuit to do this is shown in figure 4-4.

The 100k potentiometer is used to set the APD bias voltage for a certain temperature. If the temperature changes, the output of the LM3911 temperature sensor changes, and the difference between the actual temperature and the set temperature (the temperature 'error' signal) is inverted by the 1458 op amp and fed to the following stage. The second op amp of the 1458 adjusts the APD bias voltage by driving the high voltage transistor in a shunt-type regulator configuration. As the

temperature increases, the op-amp slowly turns off the regulator transistor. This results in less current being drawn through and less voltage dropped across the 58k series resistor, and increases the APD bias voltage as temperature increases. The voltage divider formed by the 200k and 10k resistors is used by the second op-amp to sense the voltage at the cathode of the APD, thus forming a negative feedback loop.

4.3.1.1.5 Selection of APD

The particular APD used for this application is the Mitsubishi Silicon Avalanche Photodiode PD1005. It has a breakdown voltage of approximately 150V, a dark current of approximately 0.3nA, a responsivity of 0.4 A/W, and a cutoff frequency of 400 MHz.

4.3.1.2 Preamplifier Design Techniques

Since the noise performance of a system is mostly determined by the noise performance of the detector and first stage of amplification, careful design of the preamplifier is crucial. The goal of designing a low noise front end is minimizing the contributions of various sources of noise, from both the APD and the preamplifier. Three designs of the first amplification stage can be used; resistive termination, high impedance (integrating) preamplifiers, and transimpedance preamplifiers.

4.3.1.2.1 Resistive Termination

The most straightforward method is to terminate the photodetector in a load resistor R. APD photocurrents will then create a voltage across the resistor R, with R acting as a current-to-voltage converter. While this approach has the advantage of being simple, it has been shown to be excessively noisy in comparison to other techniques (Smith & Personick). For example, a 75Ω resistor has 100 times *more* thermal noise and 100 times *less* gain than a typical 7KΩ transimpedance amplifier.

4.3.1.2.2 High Impedance Preamplifiers

An alternative approach is often referred to as the high impedance or integrating front end. Considerable improvement in noise performance can be achieved over the simple resistive termination (Personick). However, the result of optimizing the preamplifier for best noise performance is an undesirably narrow bandwidth. This situation is remedied by subsequent equalization which can be performed with little or no noise penalty (Smith). The degree of equalization will depend on the parasitics of the circuit, thus "tuning" may be required of each unit to match the cutoff and pick-up frequencies of the amplifier and equalizer, respectively (Hullet and Muoi). Typically, these preamplifiers must be equalized over many decades of frequency. The most serious drawback of this approach, however, is the loss of dynamic range resulting from equalization and the high input impedance (Smith). In addition, baseline wander effect is more severe due to the integration effect of long strings of "ones" and "zeros" in the input data stream, thus necessitating the use of a line coding scheme that is DC balanced. The high impedance preamplifier approach is

used extensively in nuclear engineering and other applications where individual photons must be counted.

4.3.1.2.3 Transimpedance Preamplifiers

The third approach, and the one employed here, is to use a transimpedance amplifier, which is essentially a current-to-voltage converter. When compared to an optimized, equalized high-impedance amplifier, a transimpedance amplifier/ APD combination requires only approximately 1 dB more optical power to obtain the given error rate. Circuit simplicity, eliminating the need to employ equalization, and increased dynamic range are the advantages gained by accepting this 1 dB difference in performance (Smith).

A simplified diagram of the transimpedance amplifier is shown in figure 4-3. In the limit of large gain, the output voltage V_0 is related to the input current, i, by the relation

$$V_0 = -Z_f \cdot i$$

where Z_f is the feedback impedance. Transimpedance amplifiers are less noisy for a given bandwidth, or alternatively, have more bandwidth for a given noise level.

In principle, the preamplifier gain can be increased by increasing the feedback resistance R_f. However, the feedback resistance thermal noise of a transimpedance amplifier is normally a significant portion of the total noise, thus there is a

Simplified Model of Transimpedance Amplifier

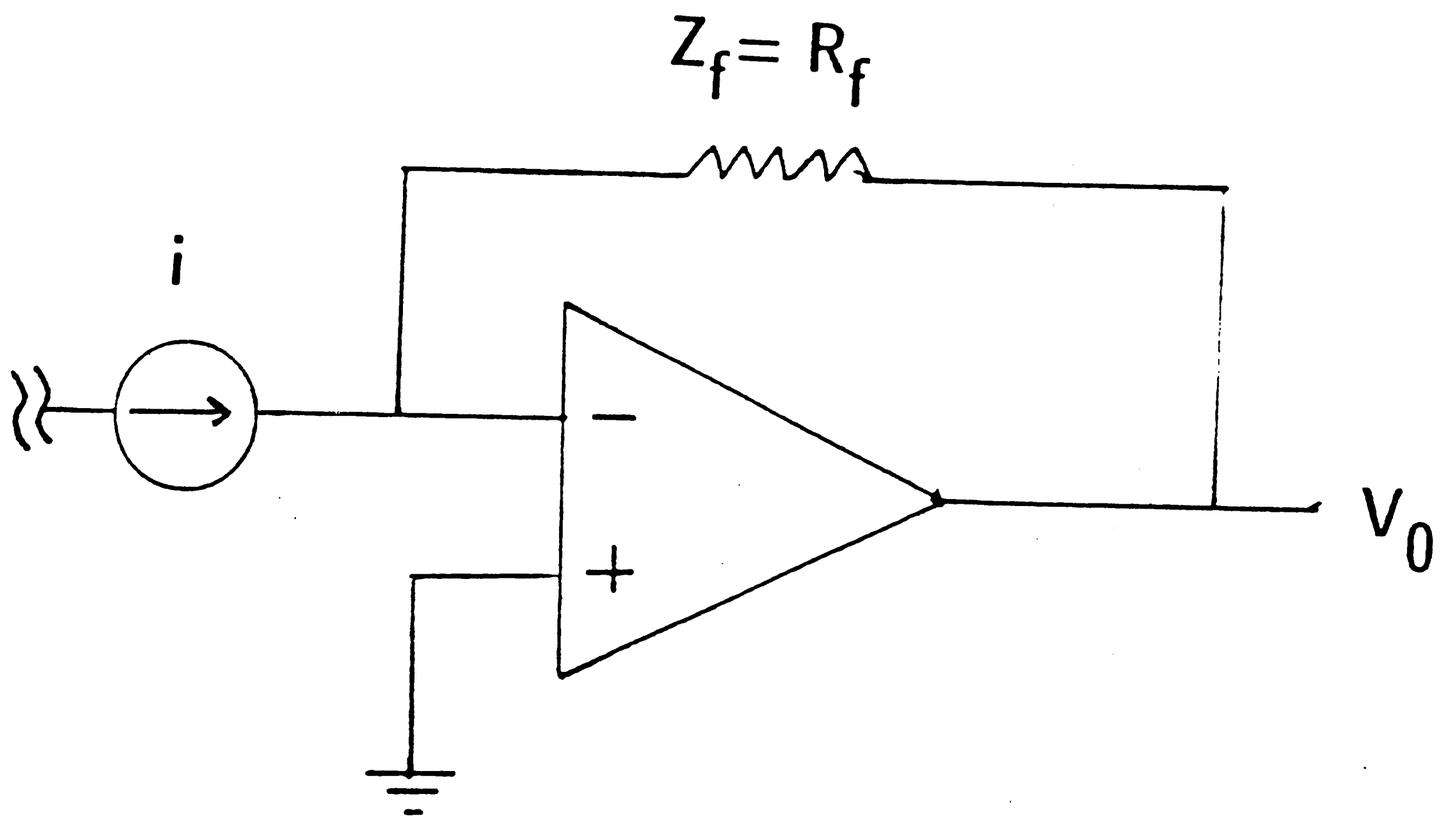


Figure 4-3

compromise between noise performance and gain. Since the APD looks like a current source with infinite impedance (Personick), it is best matched to a high-impedance load. If wide-band 50Ω amplifiers are used, the sensitivity of such systems are limited due to the low amplifier input impedance (Kasper et al).

4.3.1.3 Use of Transimpedance Amplifier

In the system design, a Signetics NE5212N transimpedance amplifier is used. This amplifier has been optimized at the factory for high performance. The resulting amplifier has an equivalent transimpedance of 7KΩ, a bandwidth of 150 MHz, and an input noise current of $\sqrt{\overline{i_n^2}}$=2.5pA/Hz$^{1/2}$.

The use of the transimpedance amplifier is shown in figure 4-4. The input is AC coupled to the APD, and the output is connected to the demodulator circuit. In the completed communication link operating over a 3 km. distance, the input to the transimpedance amplifier is approximately 10μA P-P, and the output is approximately 100 mV P-P.

4.3.1.4 Optimization of Noise Performance

The signal-to-noise ratio of the APD/preamplifier combination can be maximized with the proper choice of the APD multiplication rate. For the system under consideration, the optimum multiplication rate is in the 50-100 range.

At first, it may seem that the signal to noise ratio can be increased to arbitrarily

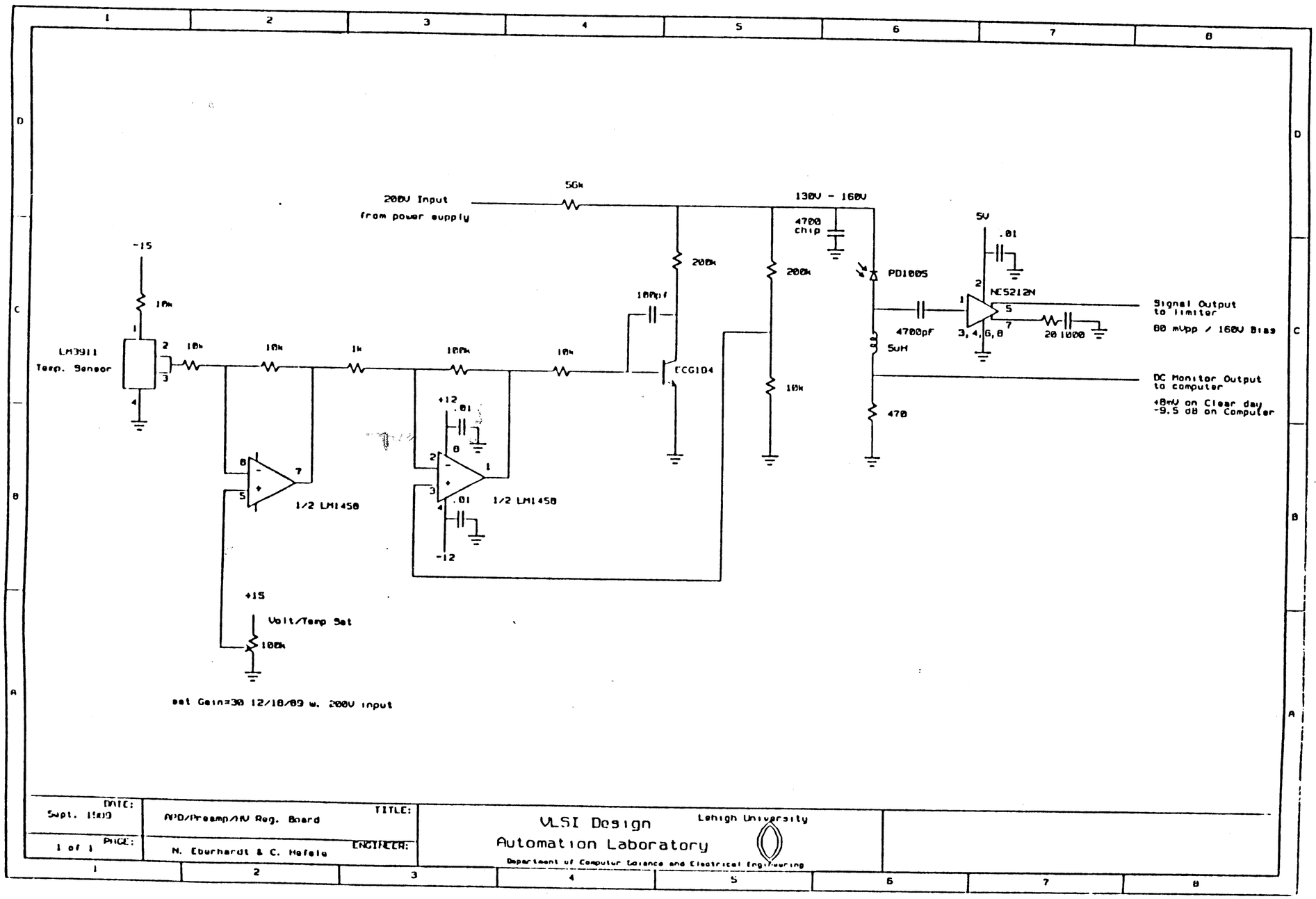

200V Input
from power supply
5Gk
130V - 160V
4700
chip
200k
200k
PD1005
5V
.01
NE5212N
Signal Output
to limiter
80 mVpp / 160V Bias
4700pf
5uH
3, 4, 6, 8
20
1000
DC Monitor Output
to computer
+8mV on Clear day
-9.5 dB on Computer
470
10k
-15
10k
LM3911
Temp. Sensor
10k
10k
1k
100k
10k
100pf
ECG104
+12
.01
-12
.01
1/2 LM1458
1/2 LM1458
+15
Volt/Temp Set
100k
set Gain=30 12/18/89 w. 200V input
DATE:
Sept. 1989
TITLE:
A/D/Preamp/HV Reg. Board
PAGE:
1 of 1
ENGINEER:
N. Eberhardt & C. Hefele
VLSI Design
Automation Laboratory
Lehigh University
Department of Computer Science and Electrical Engineering

Figure 4-4

high levels by increasing the APD gain without bound. However, increasing the APD gain adds more noise to the system due to the increased multiplication of dark currents and increased multiplication excess noise. There exists an optimum APD gain where the signal-to-noise ratio is maximized for a given incident optical power.

In general, the optimum avalanche gain is such that the detector shot noise is comparable to the receiver amplifier noise. Thus, the better the photodetector (lower dark current and excess noise), the higher the optimum avalanche gain. Similarly, the better the amplifier (lower noise current i_n^2), the lower the optimum avalanche gain. (Van Muoi)

To analyze this quantitatively, we must consider both the signal and the noise photocurrents in turn. With an optical signal incident on the APD, the *signal* photocurrent, I_S, is

$$I_s = R\ M\ N_d$$

where R is the responsivity of the APD (in Amperes/Watt), M is the multiplication rate, and N_d is the optical power incident on the photodetector.

The noise photocurrents are the sum of multiplication noise in the APD due to dark currents and photocurrents, and noise generated by the transimpedance amplifier. In this example, we assume that the excess noise due to multiplication of dark currents is negligible, and thus the main noise component from the APD is due to multiplication noise of the photocurrents.

The spectral power density (power per unit Hertz) of the APD multiplication noise due to photocurrents is approximately (as above),

$$\sigma^2 = 2e\, I_p\, M^{2.25}$$

where $I_p = R\, N_d$.

If the transimpedance amplifier has an equivalent noise power spectral density $\overline{i_n^2}$ at it's input, then the RMS total of the *noise* photocurrents, $\sqrt{I^2_{tot}}$, is the RMS sum of the amplifier noise, $\overline{i_n^2}$, and the excess noise due to APD multiplication of the photocurrent:

$$\sqrt{I^2_{tot}} = \sqrt{\left(2\, e\, R\, N_d\, M^{2.25} + \overline{i_n^2}\right)B}$$

where e, the electron charge, is 1.6×10^{-19}, and B is the bandwidth of the system in Hertz.

The signal to noise ratio (SNR) is the ratio of the signal and noise photocurrents, thus

$$SNR = \frac{I_s}{\sqrt{I^2_{tot}}} = \frac{R\, M\, N_d}{\sqrt{\left(2\, e\, R\, N_d\, M^{2.25} + \overline{i_n^2}\right)B}}$$

A plot of SNR vs. the multiplication rate M is shown in figure 4-5 for optical power levels and noise levels modeled for the constructed atmospheric laser communication

SNR vs. APD Gain and Optical Power

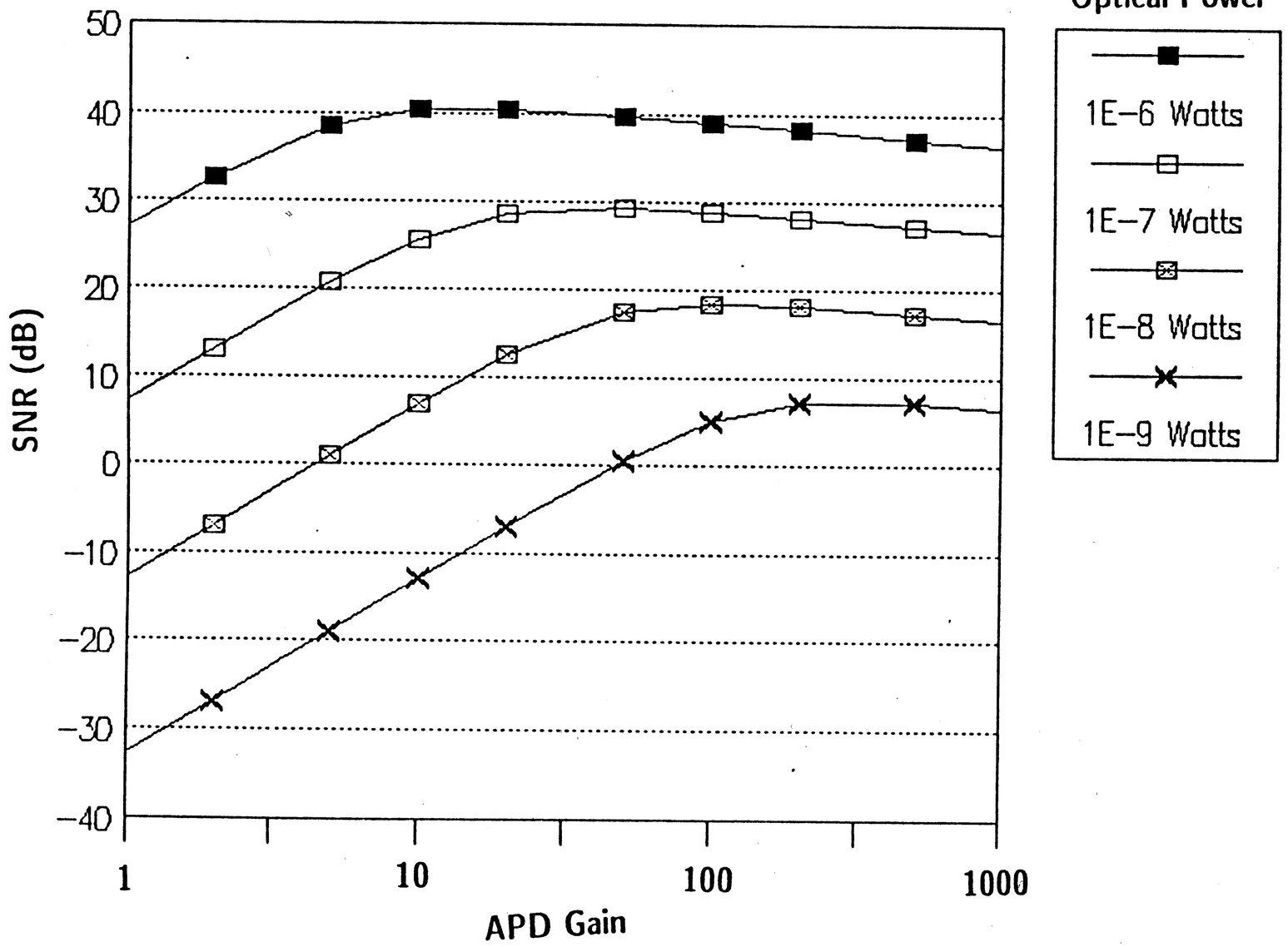


Figure 4-5

link. From these plots, once can see that the optimum multiplication rate is approximately 50-500, depending on the incident optical power. The transimpedance amplifier in this example is a Signetics SE5212, with $\sqrt{\overline{i_n^2}} = 2.5\text{pA}/\sqrt{\text{Hz}}$. In clear weather the optical power is approximately $N_d = 10^{-6}$ Watts; this decreases under adverse weather conditions. The responsivity of the APD modeled is 0.45 Amps/Watt, and the bandwidth of the system is 50 MHz.

As the figure clearly shows, the system has a rapidly rising SNR which then peaks and then slowly decreases with further increases in M. Even though the maximum SNR is obtained at one value M, APD gains above M produce SNR's that are only a few dB's from the optimum. For APD gains below M, the system SNR's decreases dramatically. Thus, it is desirable to have the APD multiplication gain at or above the optimum gain, M, once it is found.

To derive an expression for the optimum gain M, the partial derivative $\frac{\partial SNR}{\partial M}$ can be taken and set equal to zero to find the maximum:

$$\frac{\partial SNR}{\partial M} = 0 = \frac{\partial}{\partial M} \frac{R\ M\ N_d}{\sqrt{\left(2\ e\ R\ N_d\ M^{2.25} + \overline{i_n^2}\right)B}}$$

The result is

$$M = \left[\frac{4\ \overline{i_n^2}}{N_d\ e\ R}\right]^{\frac{4}{9}}$$

which is the expression for the optimum multiplication rate for maximized SNR. Note that the optimum gain is independent of the bandwidth of the system.

We can use this formula to find the optimum gain for the designed system under adverse weather conditions. If we assume $N_d=10^{-8}$ watts (typical 20dB loss under adverse weather conditions), and using the values above for $\overline{i_n^2}$ and R, we obtain an optimum gain M of 110, which is the peak of the 10^{-8}W curve in figure 4-5.

In order to ensure good performance over a wide range of received optical power, N_d, the system SNR should be optimized for the *lowest* optical power levels expected. At these levels, where the SNR is the lowest, optimization of the SNR is critical. On the other hand, at high optical power levels, even suboptimal choices of M will result in SNR's higher than those obtainable with low optical power levels.

4.3.2 Demodulator Circuitry

A functional diagram of the BPSK demodulator circuitry was shown in figure 3-1. Once the optical signal is detected and amplified by the APD and transimpedance amplifier, the signal is passed to a limiter. This limiter removes any amplitude variations in the signal. The nonlinear limiting creates harmonics which are removed by the following 50 MHz and 100 MHz filters. The 50 MHz reference signal is detected by a PLL, and doubled in frequency to obtain a 100 MHz carrier used for demodulation. The information contained in the sidebands centered at 100 MHz are mixed down to baseband, and the carrier removed with a low pass filter. From the output of the low pass filter, a threshold detector determines which of two levels have been sent.

4.3.2.1 Limiter

The limiter following the transimpedance amplifier allows a wide dynamic range of signals to be used by the following demodulator without using an automatic gain control. Furthermore, a limiter removes any AM fluctuations on the BPSK signal, and can improve the SNR by 3dB. The limiter circuit design, shown in figure 4-6, is based around two differential pair amplifiers made up of a pair of high frequency transistors. These differential pairs have one 'input' shorted to AC ground, thus making a single ended amplifier. This differential pair of transistors can be shown to begin limiting when the input (at the base) is changed ±70mV from it's bias point (Smith). Thus, the differential pair is easily driven into saturation. This allows limiting at a lower drive level than if a pair of back-to-back diodes were used, which clip at approximately ±700mV.

This constructed limiter begins limiting at approximately ±75mV (-12dBm), in good agreement with theoretical value of ±70mV. Since the transimpedance amplifier and the MAR monolithic amplifiers are limited in their drive capability, the dynamic range of each limiting stage is restricted to a range of 17 dB, from -12dBm to 5 dBm. Two limiting stages are used to give a total limiter dynamic range of over 30 dB. These stages are coupled together with MAR1 and MAR4 amplifiers, which give a total of 25 dB of amplification. If the signal passes unlimited through the first limiter stage, it is amplified sufficiently so that it will be limited by the second limiter stage (if, of course, it is in the 30dB clipping range of the entire limiter).

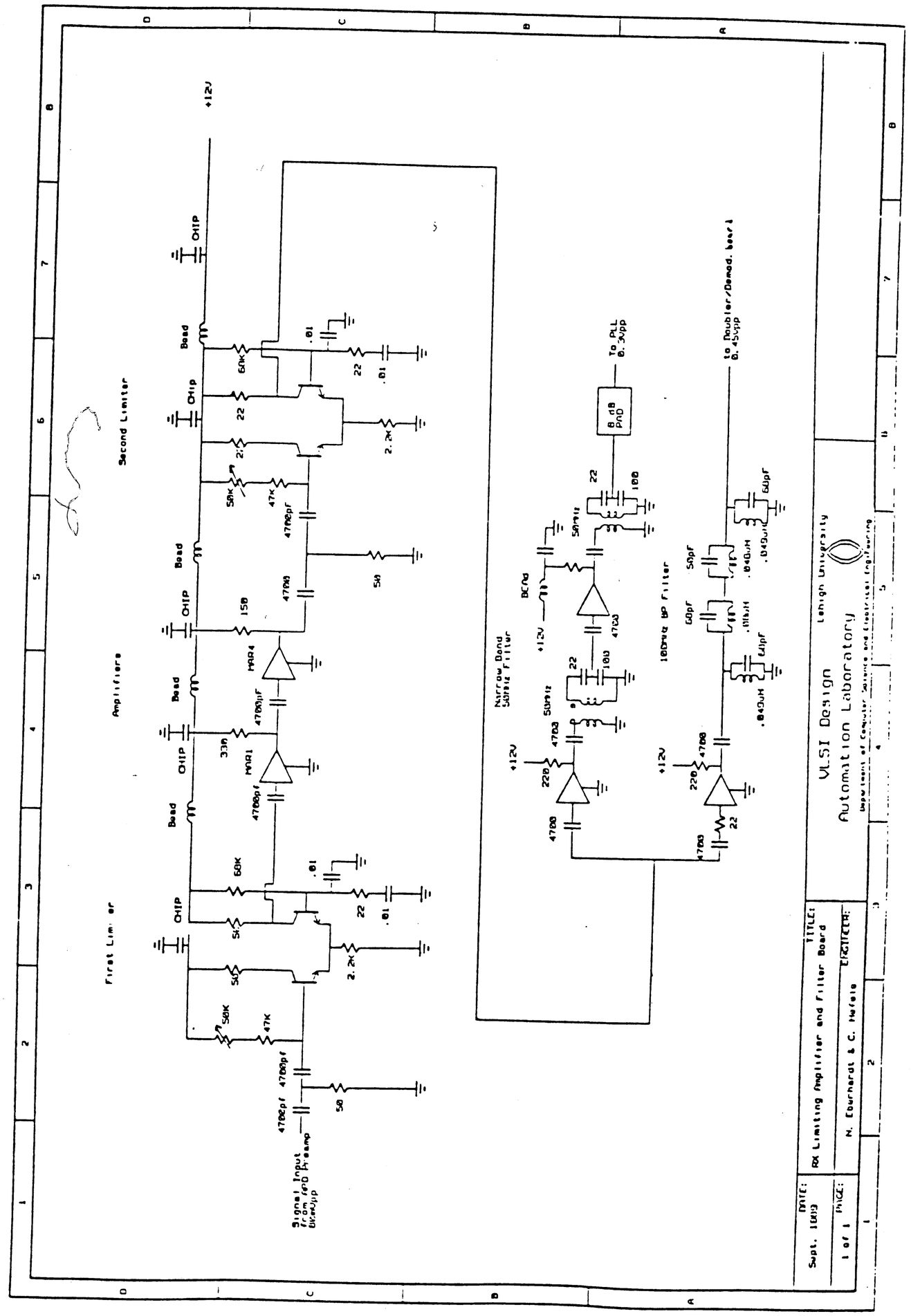

Amplifiers
Second Limiter
First Limiter
Narrow Band 50MHz Filter
180MHz BP Filter
+12V
Bead
CHIP
MAR4
MAR1
47000pF
4700pF
60K
50K
47K
22
.01
2.2K
150
330
50
To PLL
8.3Vpp
8 dB PAD
to Doubler/Demod. board
8.45Vpp
Signal Input from IF Amp
50pF
.040uH
.011uH
.049uH
Sept. 1989
1 of 1
RX Limiting Amplifier and Filter Board
N. Eberhardt & C. Hafele
Lehigh University
VLSI Design
Automation Laboratory

Figure 4-6

4.3.2.2 Filters

The filters following the limiter remove the harmonic and intermodulation components generated by the limiting process. There are two filters: a narrowband filter at 50 MHz to isolate the reference signal, and a 40 MHz wide bandpass filter centered at 100 MHz to isolate the BPSK sidebands.

The 100 MHz bandpass filter has a symmetrical Butterworth response with a bandwidth of 40 MHz. Four parallel resonators form this filter. The two end resonators have a resonant frequency of 100 MHz, and serve to short to ground any signal not near the filter center frequency and also pass those that are. The center resonators, on the other hand, resonate 65 MHz and 145 MHz and prevent signals from passing through the filter near the passband edges, but pass signals that are not near the pass band edges. The result of these resonators combined action is a bandpass response. The individual capacitor and inductor values are chosen to give the desired resonant frequencies and to match the 50Ω impedance level in the passband.

The 50 MHz filter is constructed from two tuned transformer resonators. These two stages are separated by a MAR1 buffer amplifier, so that the first resonator does not load the second. This makes for a high Q filter about 50 MHz.

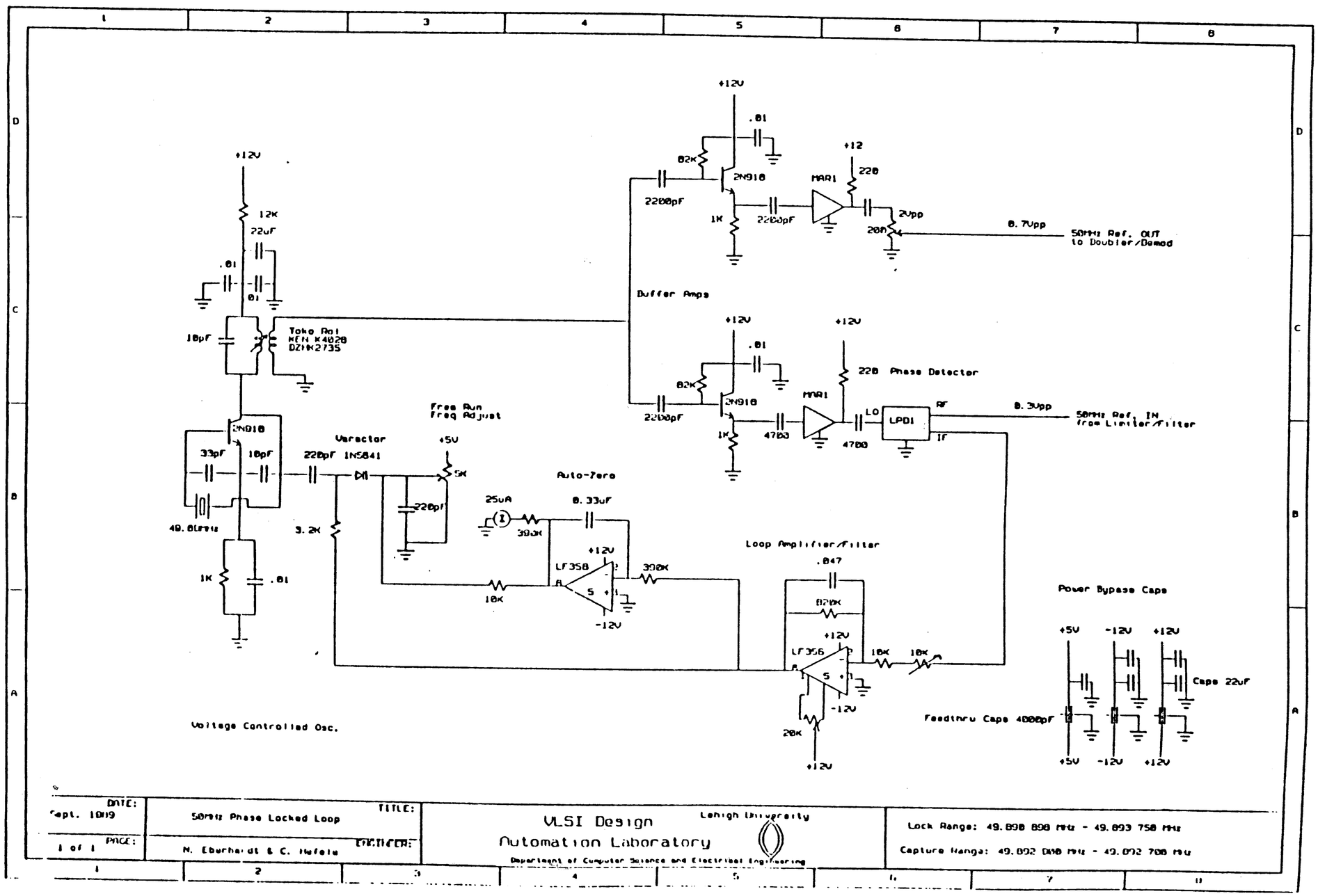
4.3.2.3 50 MHz PLL Construction

A phase-locked loop (PLL) following the 50 MHz filter is used to lock on to the 50

MHz reference signal. The PLL acts as a very narrow bandwidth (200 Hz) filter, thus improving the 50 MHz reference signal SNR by rejecting out-of-band noise. The 50 MHz PLL constructed, shown in figure 4-7, consists of a VCO, a low pass filter, phase detector, and autozero circuit. The voltage controlled oscillator is almost identical to the oscillator used in the BPSK modulator, except a varactor diode has been added so that the oscillator frequency can be changed with a control voltage.

The output of the 50 MHz oscillator is amplified by two buffer amplifiers to avoid loading the oscillator. The each amplifier consists of a 2N918 high frequency transistor in an emitter follower configuration that provide a high impedance input and low impedance output. The output of one amplifier is the VCO frequency, which is then used by the demodulator board as a 50 MHz reference signal; the other amplifier output drives the Mini-Circuits LPD1 phase detector. The LPD1 is simply a mixer optimized for use as a phase detector. It's output is the low frequency product of the 50 MHz VCO and the input signal.

The low frequency output of the phase detector passes to the active loop filter, which is an op amp in an inverting low pass filter configuration with a -3dB bandwidth of 20 Hz and variable gain from 40 to 80. The offset trim pot is used to null any op amp offset voltage. The output of the active filter drives the VCO's varactor diode connected in series with a 3.2K current limiting resistor. The varactor diode is AC coupled to the oscillator, and by varying the reverse bias voltage across the diode the diode's capacitance can be changed by several picofarads. This change in capacitance can in turn change the oscillator frequency by up to 2.75 kilohertz, the lock range of the PLL. The 5K pot connected to the cathode of the varactor



DATE: Sept. 1989	TITLE: 50MHz Phase Locked Loop	VLSI Design Automation Laboratory, Lehigh University, Department of Computer Science and Electrical Engineering	Lock Range: 49.898 898 MHz - 49.893 758 MHz
PAGE: 1 of 1	ENGINEER: N. Eberhardt & C. Hefele		Capture Range: 49.892 048 MHz - 49.892 708 MHz

Figure 4-7

diode is used to set the free running frequency of the oscillator. Furthermore, the auto-zero circuit, also connected to the cathode of the varactor diode, forces the PLL steady state phase error to zero by integrating the DC output of the loop filter. The DC output of the filter is a result of op amp offset voltages as well as a constant phase difference between the 50 MHz input and output waveforms. This level, which can be read on the 25μA meter connected to the integrator/Auto-Zero circuit, also influences the reverse bias voltage across the varactor diode.

4.3.2.4 PSK Demodulator

The demodulator circuit, shown in figure 4-8, recovers the digital data stream from the BPSK modulated signal centered at 100 MHz. Since the 50 MHz and BPSK signals pass through different subcircuits on their way to the demodulator, any resulting phase difference is corrected by using a coaxial cable delay line. Since the BPSK signal centered at 100 MHz consists only of two sidebands and no carrier, a 100 MHz carrier must be reconstructed at the receiver for demodulation. To do this, the 50 MHz reference signal is amplified in order to drive the Mini-Circuits SK2 frequency doubler. The output of the doubler is a 100 MHz signal which is in exact synchronization with the transmitter's carrier oscillator. Then, the 100 MHz BPSK sidebands and the reconstructed carrier are mixed by the ASK1 Mixer. This mixes the BPSK signal down to baseband. The output of the mixer is an analog bipolar signal, representing the level of each bit. The NE527 voltage comparator converts the analog bipolar level representing each bit into two complementary TTL levels for data output. An example test signal received over this a 3km test link is shown in figure 4-9.

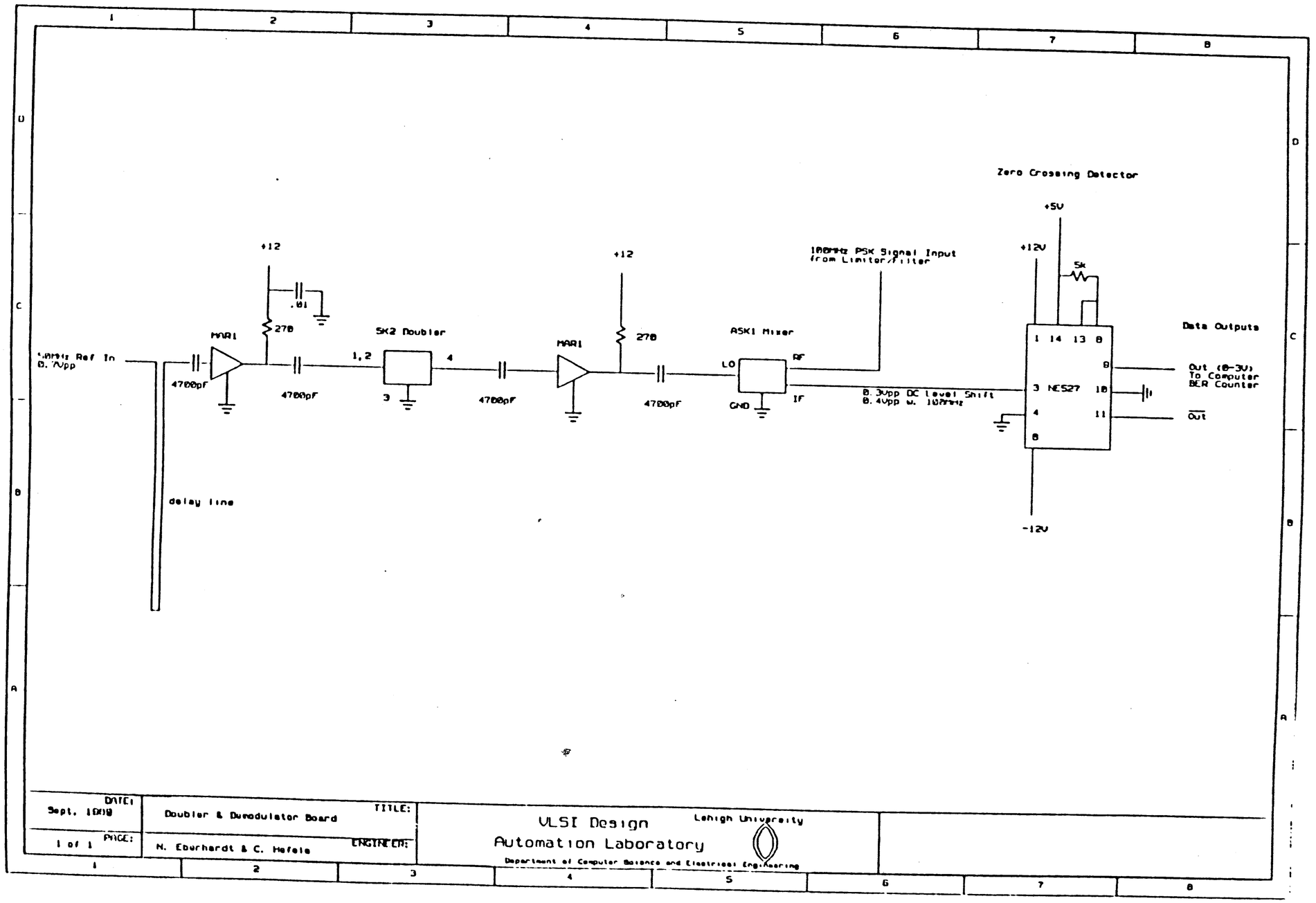
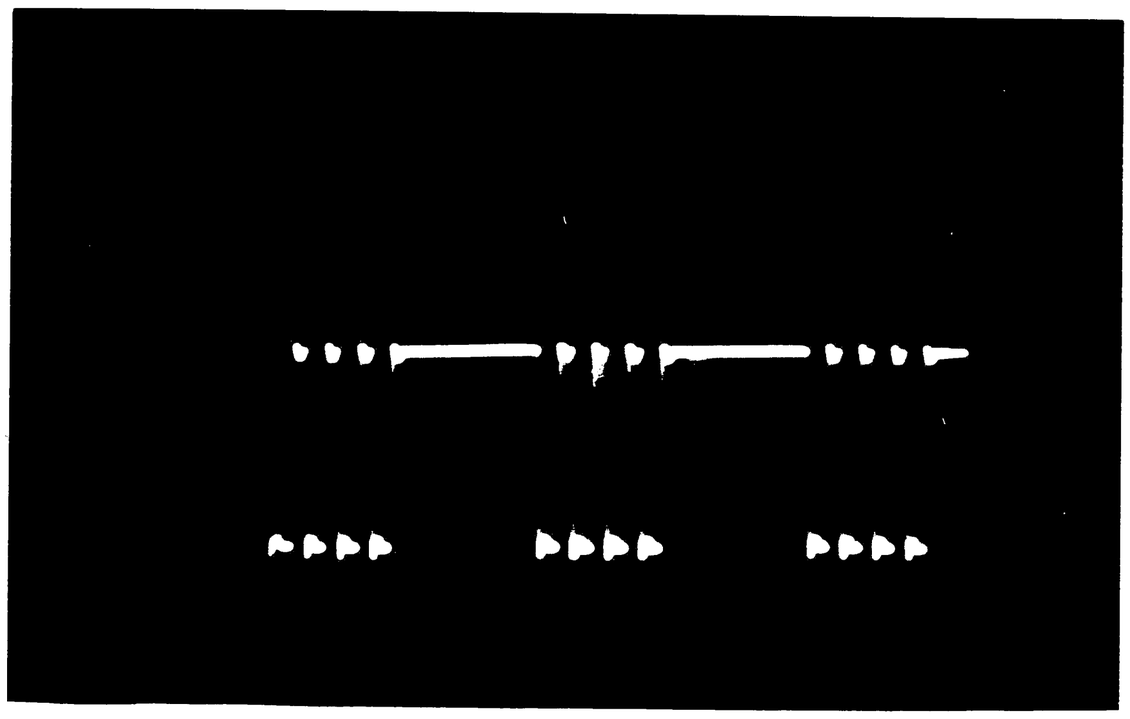

Zero Crossing Detector
+5V
+12V
5k
100MHz PSK Signal Input
from Limiter/Filter
+12
.01
270
MAR1
SK2 Doubler
1, 2
4
3
4700pF
ASK1 Mixer
LO
RF
IF
GND
8.3Vpp DC level Shift
8.4Vpp w. 100MHz
Data Outputs
1 14 13 8
3 NE527 10
4
8
11
Out (0-3V)
To Computer
BER Counter
Out
-12V
delay line
DATE:
Sept. 1990
TITLE:
Doubler & Demodulator Board
PAGE:
1 of 1
ENGINEER:
N. Eberhardt & C. Hefele
VLSI Design
Automation Laboratory
Lehigh University
Department of Computer Science and Electrical Engineering

Figure 4-8

Test Signal Received with 3km System



Digital test sequence received over the
3 km atmospheric laser communication link.
(2V/div. vertical, 0.2μs/div. horizontal)

Figure 4-9

5. OPTICS

5.1 Introduction

The functions and optical design considerations for the transmitter and receiver optics are distinctly different. In designing the transmitter optics, which forms a collimated beam from the laser output, there are design compromises between beamwidth and pointing accuracy, beam shape, and laser astigmatism. In designing the receiver optics, which focuses a small portion of the beam on the photodetector, there are compromises between signal level, acceptance angle, lens diameter, focal lengths, lens correction, and minimum spot size. Weighing these factors, the transmitter and receiver optics were assembled and tested. The transmitter has an optical power output of 40mW (peak), with a 1mR by 2mR beamwidth. The receiver has a 3mR acceptance angle, with a focused spot size much smaller than the detector.

5.2 Transmitter Optics

The goal of the transmitter optics is to collimate the modulated laser output and direct this small divergence beam at the receiver. The transmitter lens must be optically matched to the characteristics of the laser source, and an appropriate beamwidth collimator must be selected and constructed.

5. OPTICS

5.1 Introduction

The functions and optical design considerations for the transmitter and receiver optics are distinctly different. In designing the transmitter optics, which forms a collimated beam from the laser output, there are design compromises between beamwidth and pointing accuracy, beam shape, and laser astigmatism. In designing the receiver optics, which focuses a small portion of the beam on the photodetector, there are compromises between signal level, acceptance angle, lens diameter, focal lengths, lens correction, and minimum spot size. Weighing these factors, the transmitter and receiver optics were assembled and tested. The transmitter has an optical power output of 40mW (peak), with a 1mR by 2mR beamwidth. The receiver has a 3mR acceptance angle, with a focused spot size much smaller than the detector.

5.2 Transmitter Optics

The goal of the transmitter optics is to collimate the modulated laser output and direct this small divergence beam at the receiver. The transmitter lens must be optically matched to the characteristics of the laser source, and an appropriate beamwidth collimator must be selected and constructed.

5.2.1 Semiconductor Laser

Since semiconductor lasers are rugged, small, inexpensive, and can be easily modulated, they are the laser of choice. The output beam of a semiconductor laser is a Gaussian beam with elliptical cross section. Along the major and minor axes of the beam cross section, the Gaussian intensity distribution is given by

$$P(r) = \frac{2\,P_0}{\pi\,a^2}\,e^{2r^2/a^2}$$

where P_0 is the total output power, P(r) is the intensity at distance r from the beam center, and a is the radius where the intensity falls to $1/e^2$ along the major/minor axis. Typically, specifications for such lasers give the beamwidth between the $1/e^2$ points along the major and minor axes of the output beam's elliptical cross section.

The focal point along the major or minor beam axis is the point where the $1/e^2$ beamwidth lines intersect within the laser. The major axis and the minor axis focal points are different, typically by several microns. This is referred to as the laser astigmatism, and it can be corrected with a cylindrical lens which moves the focal point in one plane, but leaves the focal point in the perpendicular plane unaffected.

The particular laser selected was a SLD202 V-3 50 mW High Power Semiconductor Laser Diode made by Sony. It has a wavelength of 820 nm, in the near-infrared region. The beamwidths measured parallel and perpendicular to the beam minor axis are $\phi_{\perp} = 28^{\circ}$ and $\phi_{||} = 13^{\circ}$. The laser has an astigmatism of $Z_{||}$-$Z_{\perp} = 40\mu$m.

5.2.2 Limits on Collimation and Beam Diameter

Under the laws of geometrical optics, a Gaussian beam diverging at an angle of θ from a point should be able to be perfectly collimated. In reality, a Gaussian beam either diverges from a region where the beam is smallest, called the beam waist, or converges to one. The amount of divergence or convergence is measured by the full angle beam divergence θ, which is the angle subtended by the $1/e^2$ diameter points for distances far from the beam waist.

The far field beam divergence of a Gaussian beam is given by

$$\theta = \frac{2\lambda}{\pi a}$$

where θ is the full angle beam divergence measured at the $1/e^2$ intensity points. λ is the wavelength of the radiation. and a is the beam radius at the beam waist. From this equation, one can see that zero divergence is impossible, and that in order to obtain a divergence of 1 mR, the beam diameter must be at least 1 mm. Smaller divergences require larger output beam diameters.

5.2.3 Collimation Requirements

Although it is physically possible to obtain beamwidths of a few tens' of microradians for use in an atmospheric laser communication system, it is usually not practical to do so for several reasons. First, a small divergence requires a large beam waist, and thus larger, more expensive lenses. For example, a beamwidth of

10μR requires at least a 100mm lens. Next, the transmitter may shift due to vibration or thermal expansion. Also, the initial alignment of systems with very narrow beamwidths can be time consuming. Furthermore, atmospheric turbulence over paths of many kilometers can deflect the transmitted beam by up to 100 microradians.

In order to keep the transmitted beam on the receiver, systems with beamwidths on the order of one milliradian are the most practical compromise between receiver signal strength and pointing accuracy. Narrower beamwidths would only become practical if an automatic pointing circuit were used to continuously compensate for pointing errors.

5.2.4 Summary of Transmitter Optical Considerations

First, the laser astigmatism must be corrected with a cylindrical lens. Second, the entire laser output should be directed at the collimation lens. Next, the output beamwidth should be approximately 1mR. Furthermore, the output beam should be larger than 1mm in diameter, which is the beam waist size for a Gaussian beam with 1mR divergence.

5.2.5 Transmitter Optics Construction

A diagram of the assembled transmitter optics is shown in figure 5-1. It consists of a custom cylindrical lens which corrects for the laser astigmatism, and a collimator lens, which collimate the laser output to a 1 mR by 2mR beam.

Transmitter Optics

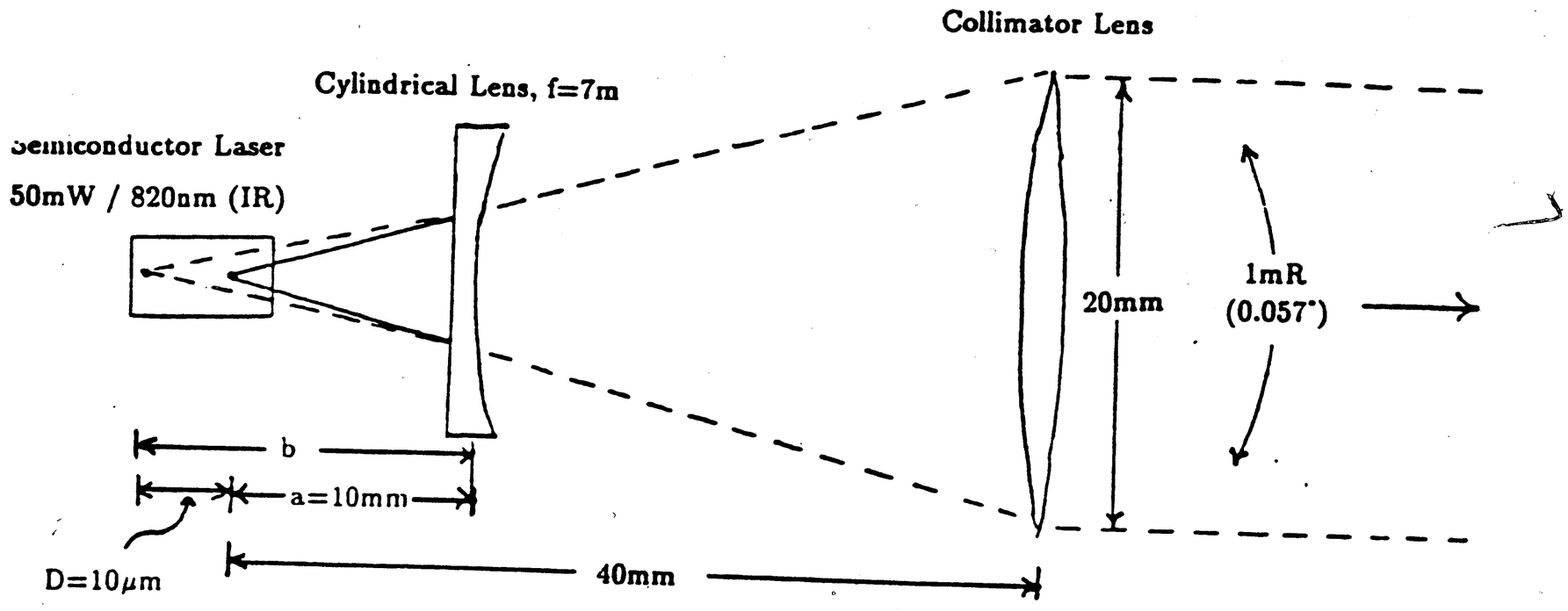


Figure 5-1

The custom cylindrical lens was manufactured by Optics for Research. This lens obeys the thin lens equation

$$\frac{1}{f} = \frac{1}{a} - \frac{1}{b}$$

From figure 5-1, one can see that the astigmatism of the laser, D, is

$$D = b - a$$

Substituting b=D+a into the thin lens equation, one can obtain

$$\frac{1}{f} = \frac{D}{a(D+a)}$$

Using this equation, one can verify that f=7000mm and a=10mm will correct the laser's 10μm astigmatism.

The spherically corrected collimator lens is also manufactured by Optics for Research. The collimator lens has a diameter of 20mm, and a focal length of 40mm. With a laser beamwidth of 28° and lens focal length of 40mm, the lens is not overfilled. The output of the collimator is a 1mR by 2mR beam, reflecting the elliptical output of the laser source.

5.3 Receiver Optics

There are four major design factors to consider while designing the receiver optics. First, due to the large path losses in a typical atmospheric laser communication

system, the receiver optics should collect as much of the transmitted beam as is practically possible. Second, any stray light must be eliminated with filters. Third, spherical aberration and the diffraction limit determine the focused spot size, which in turn influences the amount of light falling on the photodetector. Lastly, the receiver acceptance angle is influenced by the receiver lens focal length, as well as by any additional lenses. The constructed receiver optical assembly is shown in figure 5-2.

5.3.1 Path Losses

Typical path losses may be many 10's of dB's. For example, if the transmitter beamwidth is set at approximately 1 mR, then after 3 km the beam will have a diameter of approximately 1 mR × 3 km = 3 meters. If the receiver lens diameter is 50mm, only a small portion of the transmitted beam is collected. The loss can be roughly approximated by finding the ratio of the receiver lens area to the approximate area of the total beam at the receiver. If this loss, L_{dB} is converted into decibels, we have

$$L_{dB} = 10 \text{ Log } \frac{\pi d_1^{\,2}}{\pi d_2^{\,2}} = 10 \text{ Log } \frac{0.050^2}{3.0^2} = -36 \text{ dB}$$

where d_1 and d_2 are the receiver lens and received beam diameters, respectively. Adverse weather conditions, such as snow and fog, can increase losses by up to another 40dB per kilometer, in some instances.

Receiver Optics

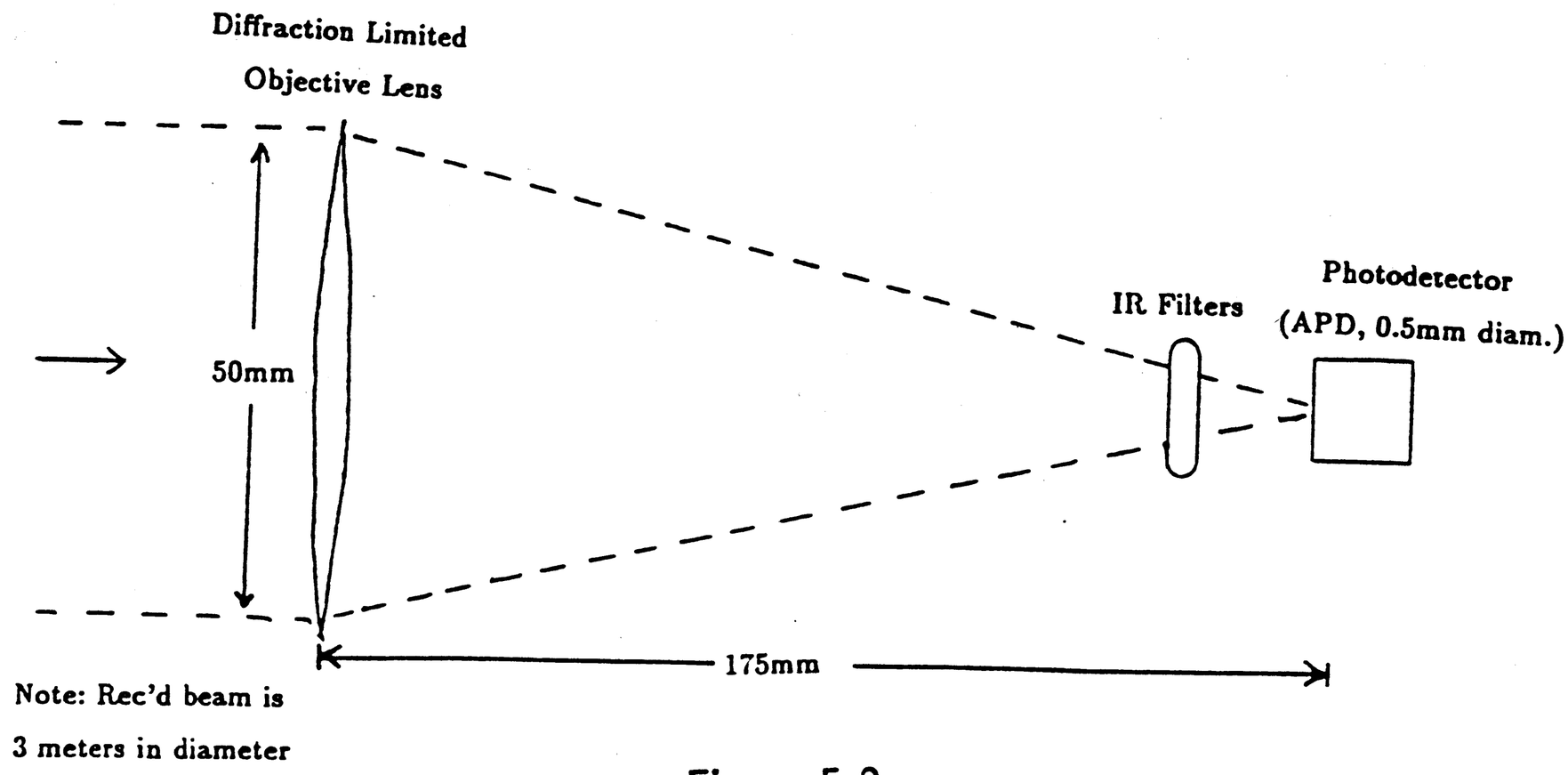


Figure 5-2

Receiver Optics with Elliptical Lens

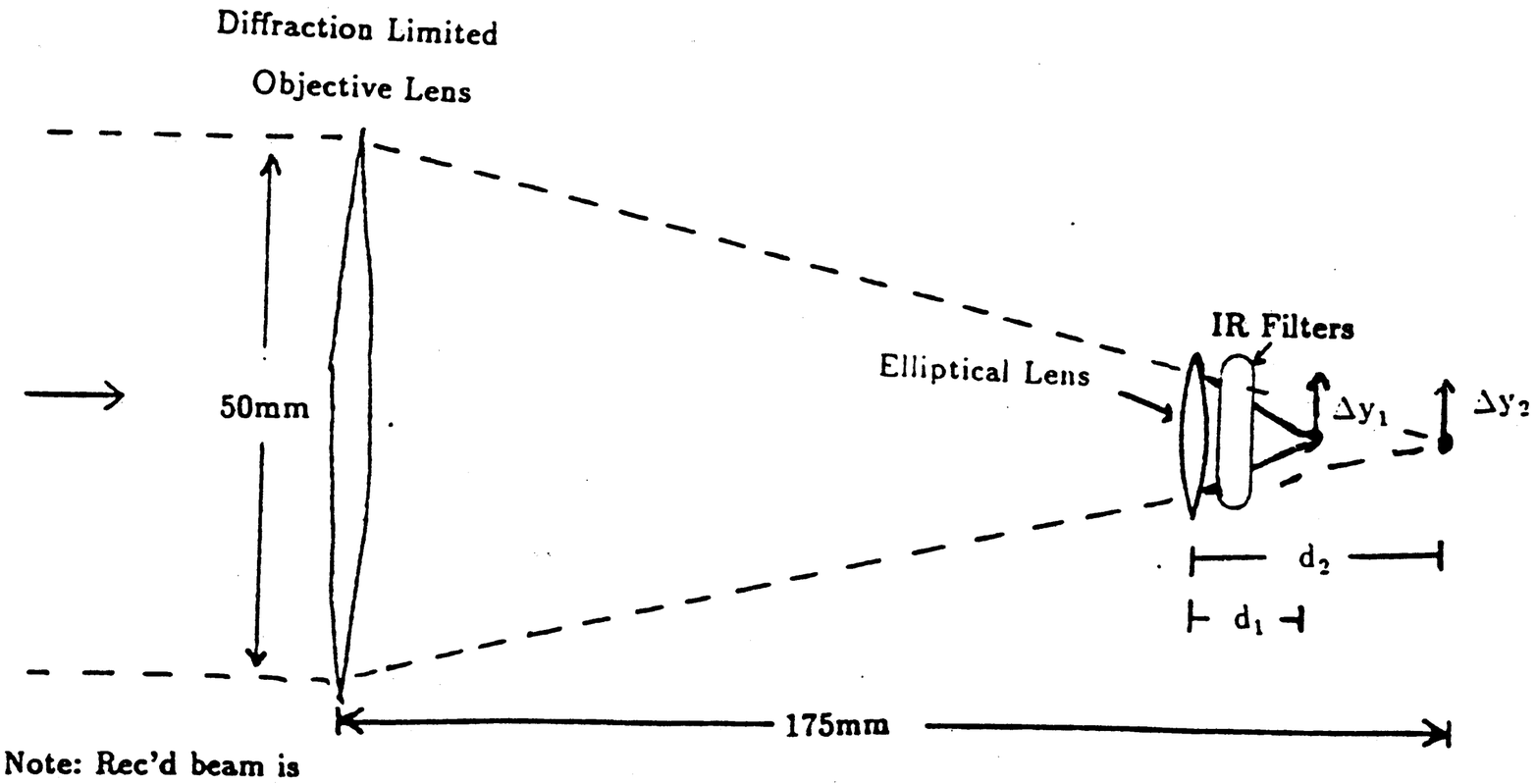


Figure 5-3

5.3.2 Filters

In order to eliminate stray background light effects, an optical filter was placed directly in front of the photodetector. This optical "bandpass" filter is centered around the 820nm wavelength, and attenuates stray background light of wavelengths other than the laser wavelength. This "bandpass" filter is actually the combination of two optical filters, one "lowpass", and one "highpass". The filters used are an RG9, and a RG780, respectively.

5.3.3 Diffraction Limited Optics

When a lens is irradiated at its aperture with a uniform intensity distribution, a diffraction pattern is formed at the focus. This pattern consists of a bright central spot, known as the Airy disk, surrounded by rings of rapidly diminishing intensity. With perfect lenses, 84% of the optical power is within the central spot, with the rest contained in the rings. The diameter, S, of the central spot is approximated by

$$S = 2.44\ \lambda\ F^{\#}$$

where S is the diameter of the central spot, λ is the wavelength of the source, and $F^{\#}$ is the ratio of the focal length to the lens aperture. If the lens used is not perfect, then there is a difference between the longest and shortest optical paths leading to the focus. For a path difference of one quarter wavelength ($\lambda/4$), the size of the central spot is essentially unchanged, but the energy in the central spot decreases to 68%. Traditionally, when this $\lambda/4$ limit on optical paths is not

exceeded, the lens system is considered to be *diffraction limited.* This term means that the diffraction, and *not* the lens defects, determines the smallest spot size possible.

5.3.4 Spherical Aberration

Using a geometrical optics approach, it can be shown that when using a spherical lens with parallel incident rays, not all incident rays are focused to the same focal point. This effect is known as spherical aberration, and is a result of the spherical shape of the lens. The result is that incident light can not be focused to a point, but only to a minimum spot size, called the blur circle. The diameter of the blur circle, S, is given by

$$S = \frac{Kd^3}{f^2}$$

where d is the diameter of the lens, f is the focal length of the lens, and K is a constant reflecting the index of refraction of the lens material (here, K=0.067 for n=1.5, 0.018 for n=2.4, 0.0087 for n=4.0).

Spherical aberration can be corrected with the proper lens design so that the lens can perform to the diffraction limit. This correction is expensive, so for inexpensive lenses the most important factor that determines spot size is spherical aberration.

5.3.5 Spot Size and Detector Size

The spot size is the larger of the diffraction limit and the blur circle diameters. To minimize lost energy, it is desirable to have the focused spot completely on the detector. Since the detector, the PD1005 detailed in Chapter 4, has a photosensitive surface 0.5mm (500μm) in diameter, the focused spot must be smaller than this. Actually, the spot size should be considerably smaller than the detector size, in order to give good acceptance angle characteristics (see below). With diffraction limited lenses, this is possible; with light at 0.820μm, the diffraction limited spot size is 7μm, which is much smaller than the detector 500μm diameter.

5.3.6 Acceptance Angle

The acceptance angle is the angle subtended by the photodetector in the image plane of the objective lens. Alternatively, with the receiver pointed exactly at the transmitter, the receiver acceptance angle is the full angle through which the receiver unit can move and still have the transmitter's image focused on the receiver photodetector. The angle of acceptance arises from the fact that the photodetector is of a finite size. The larger the detector, the larger the angle of acceptance, given a focused spot image. If the focused spot is of comparable size to the detector, then the acceptance angle decreases. The acceptance angle θ for a single lens is given by

$$\theta = \frac{d_p - d_s}{f}$$

where d_p is the photodetector diameter, d_s is the spot size ($d_s < d_p$), and f is the lens

focal length. Obviously, it is desirable to have a small focal length, large photodetector diameter and small spot size to increase the angle of acceptance.

The angle of acceptance can be improved by using an additional lens near the detector, as shown in figure 5-3. The lateral magnification of the second lens is given by

$$\frac{y_1}{y_2} = \frac{d_1}{d_2}$$

If the original image moved an amount Δy_2, then the new image only moves an amount Δy_1 given by

$$\Delta y_1 = \frac{d_1}{d_2} \Delta y_2$$

Since $d_2 < d_1$, the new focus moves less than the old focus, and thus the angle of acceptance is *increased* by a factor d_2 / d_1.

5.3.7 Summary of Receiver Optical Considerations

The two goals of the receiving optics are to first collect as much optical power as possible, and second to transfer it to the photodetector over a large angle of acceptance. The first goal, to collect as much of the transmitted beam as possible, can be met by using an objective lens with a diameter as large as is economically and physically possible. The second goal can be met by using a "fast" speed lens.

The speed of a lens, $F^{\#}$, is defined as the ratio of the focal length to the lens aperture,

$$F^{\#} = \frac{f}{d}$$

"Fast" lenses have small $F^{\#}$, and "slow" lenses have large $F^{\#}$. Fast lenses have several advantages. First, diffraction limited fast lens has a smaller spot size than a diffraction limited slow lens. Second, fast lenses yield larger acceptance angles than slower lenses, given a fixed detector size. However, fast diffraction limited lenses are costly since they must be corrected for spherical aberration. Slow lenses, on the other hand, need little correction for spherical aberration, but have larger spot size due to diffraction.

Since correction of spherical aberration is typically the main factor determining spot size, it is the most important factor. Thus, the best lens to obtain would have a large diameter and short focal length, and be corrected so that the spot size is at most half the diameter of the of the photodetector.

5.3.8 Receiver Optics Construction and Testing

The receiver optics are shown in figure 5-2. The objective lens used in the receiver is manufactured by Optics for Research. The lens is optimized at the factory for high performance. It is corrected to the diffraction limit, and has a 50 mm diameter and 175 mm focal length. The speed of the lens is $F^{\#} = 175/50 = 3.5$, which is moderately "fast".

Since the lens is diffraction limited, the spot size is given by

$$S = 2.44\,\lambda\,F^{\#} = 2.44 \cdot 0.82 \times 10^{-6} \cdot 3.5 = 7\mu m$$

The spot size is 7μm, which is much less than the 500μm photodetector diameter. The angle of acceptance with the PD1005 detector is thus

$$\theta = \frac{d_p - d_s}{f} = \frac{500\mu m - 7\mu m}{175mm} = 2.8mR$$

An experimental plot of received signal amplitude vs. receiver angle is shown in figure 5-4. The detector is at the focal point of the objective lens. The steep edges of the plot indicate the points where the 7μm focused spot moves completely off of the detector surface. As can be seen from this plot, the angle of acceptance is approximately 2.8mR, as predicted.

If the photodetector is moved away from the focal point, the spot size increases, and this alters the angle of acceptance plot. Experimental plots of the receiver signal amplitude vs. receiver angle with the photodetector 1mm and 2mm back from the focal point are shown in figures 5-5 and 5-6 respectively. From the triangles in figure 5-2, the spot size for the 1mm plot is obviously 50mm/175mm x 1mm = 0.28mm, half of the detector size. Similarly, the spot size for the 2mm plot is 50mm/175mm·2mm = 0.57mm, which is the detector size. As shown in figure 5-7, the light reaching the detector is in the intersection of the focused spot area and the photodetector area. If the spot increases in size to 0.28mm or 0.57mm, then as the spot scans across the detector there can be a large region where there is only *partial*

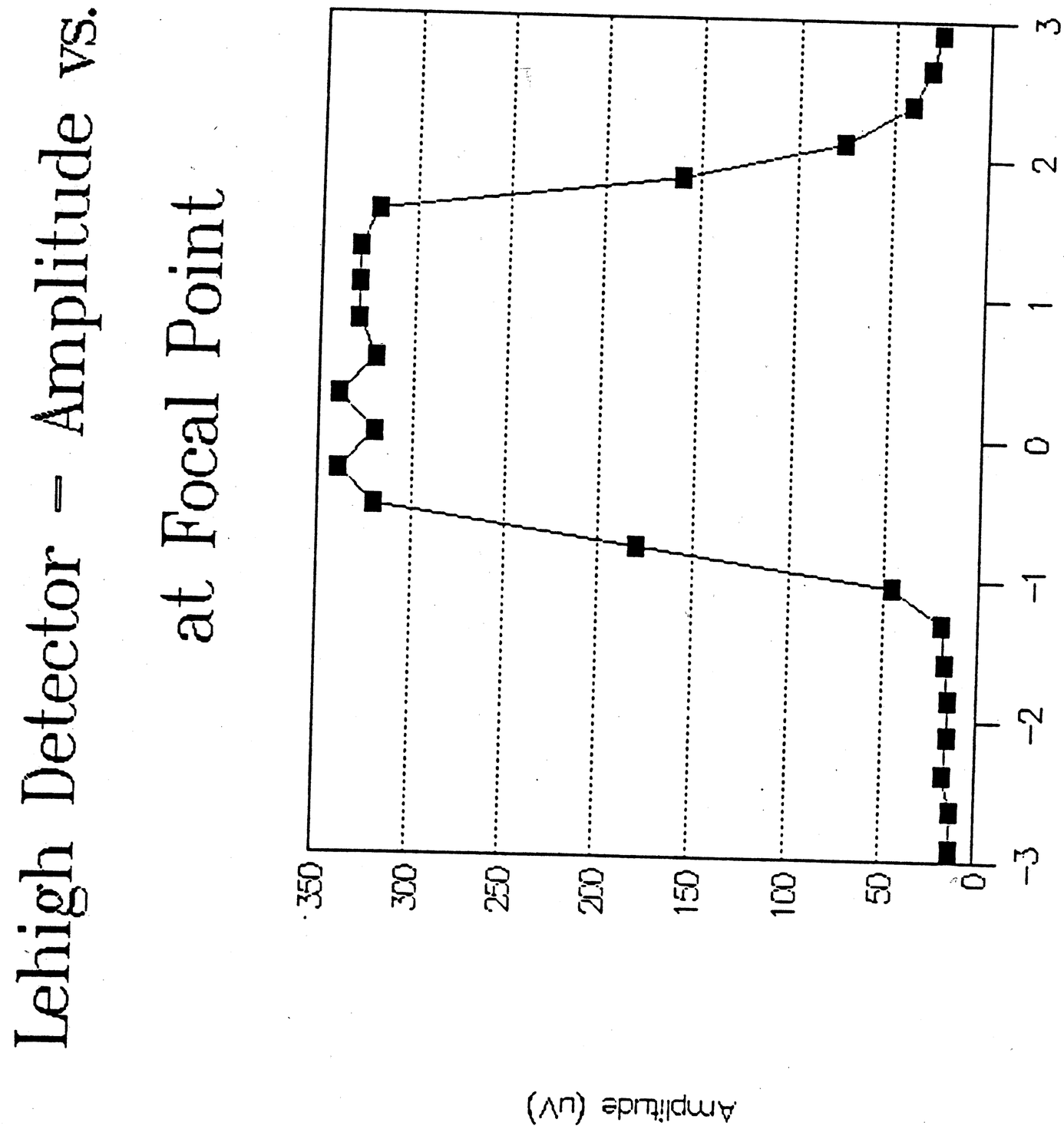

Lehigh Detector – Amplitude vs. Angle
at Focal Point
Amplitude (uV)
350
300
250
200
150
100
50
0
−3
−2
−1
0
1
2
3
Angle (mR)

Figure 5-4

Lehigh Detector – Amplitude vs. Angle
–1mm from Focal Point

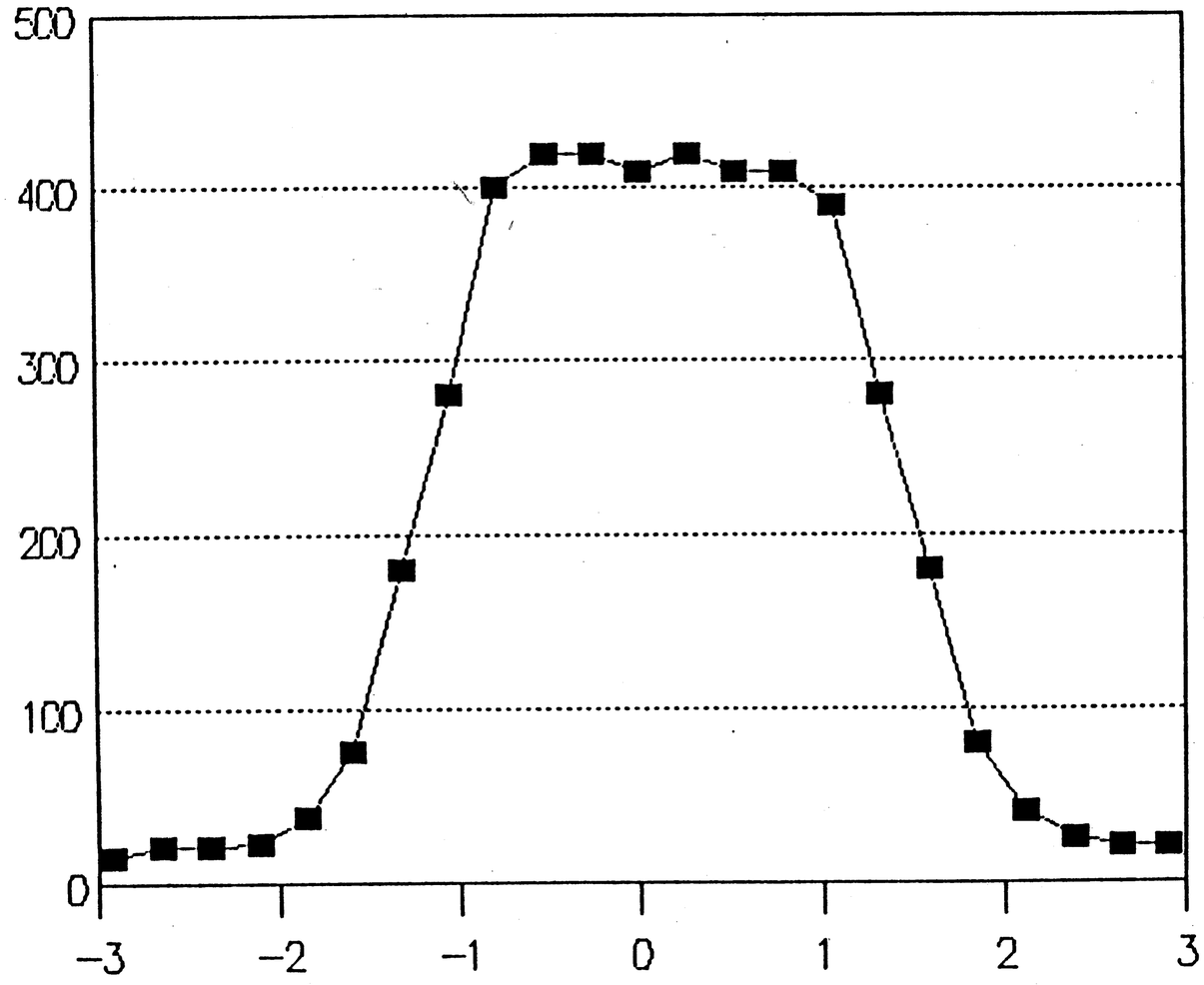


Figure 5-5

Lehigh Detector – Amplitude vs. Angle
–2mm from Focal Point

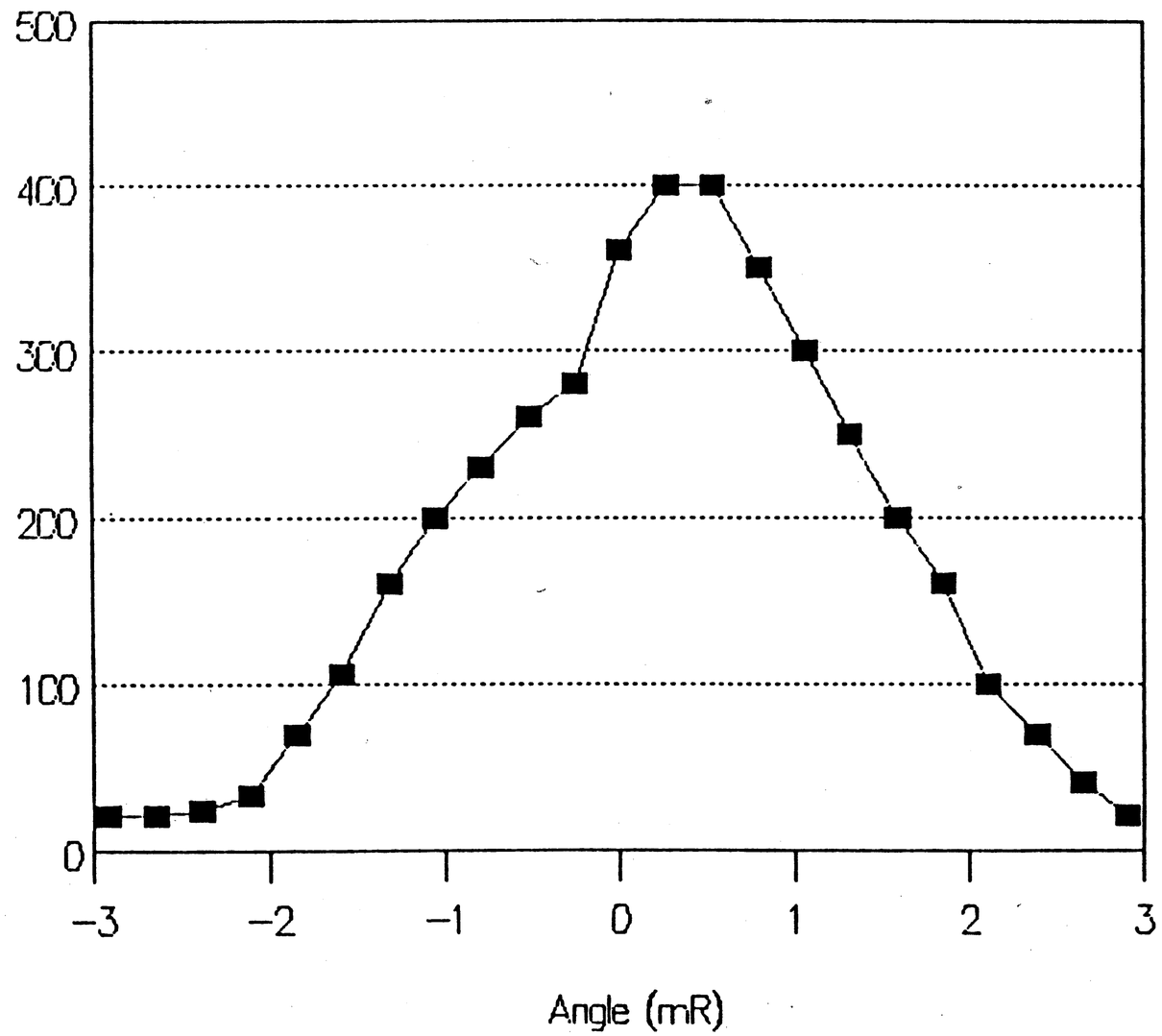


Figure 5-6

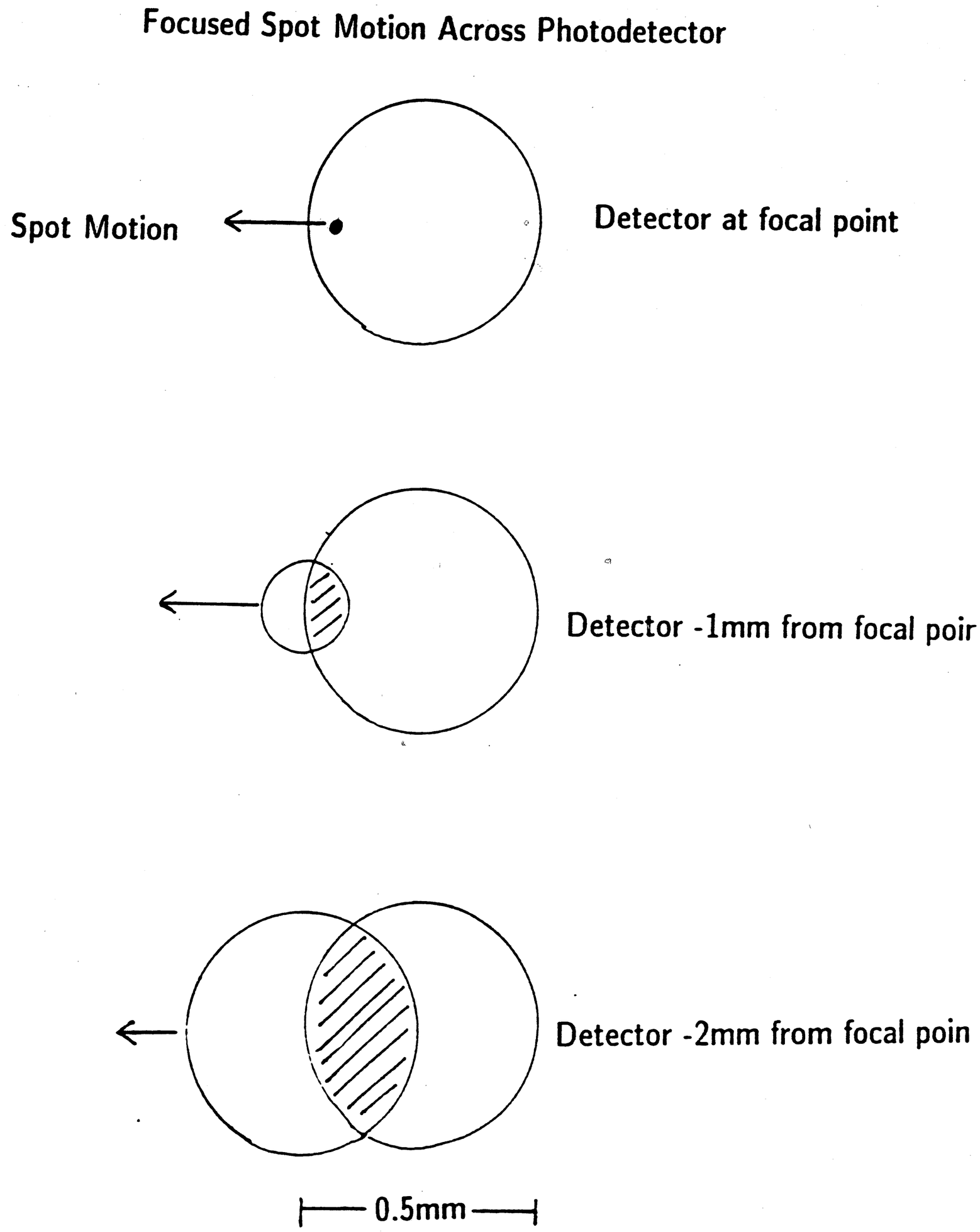

Focused Spot Motion Across Photodetector
Spot Motion
Detector at focal point
Detector -1mm from focal poir
Detector -2mm from focal poin
0.5mm

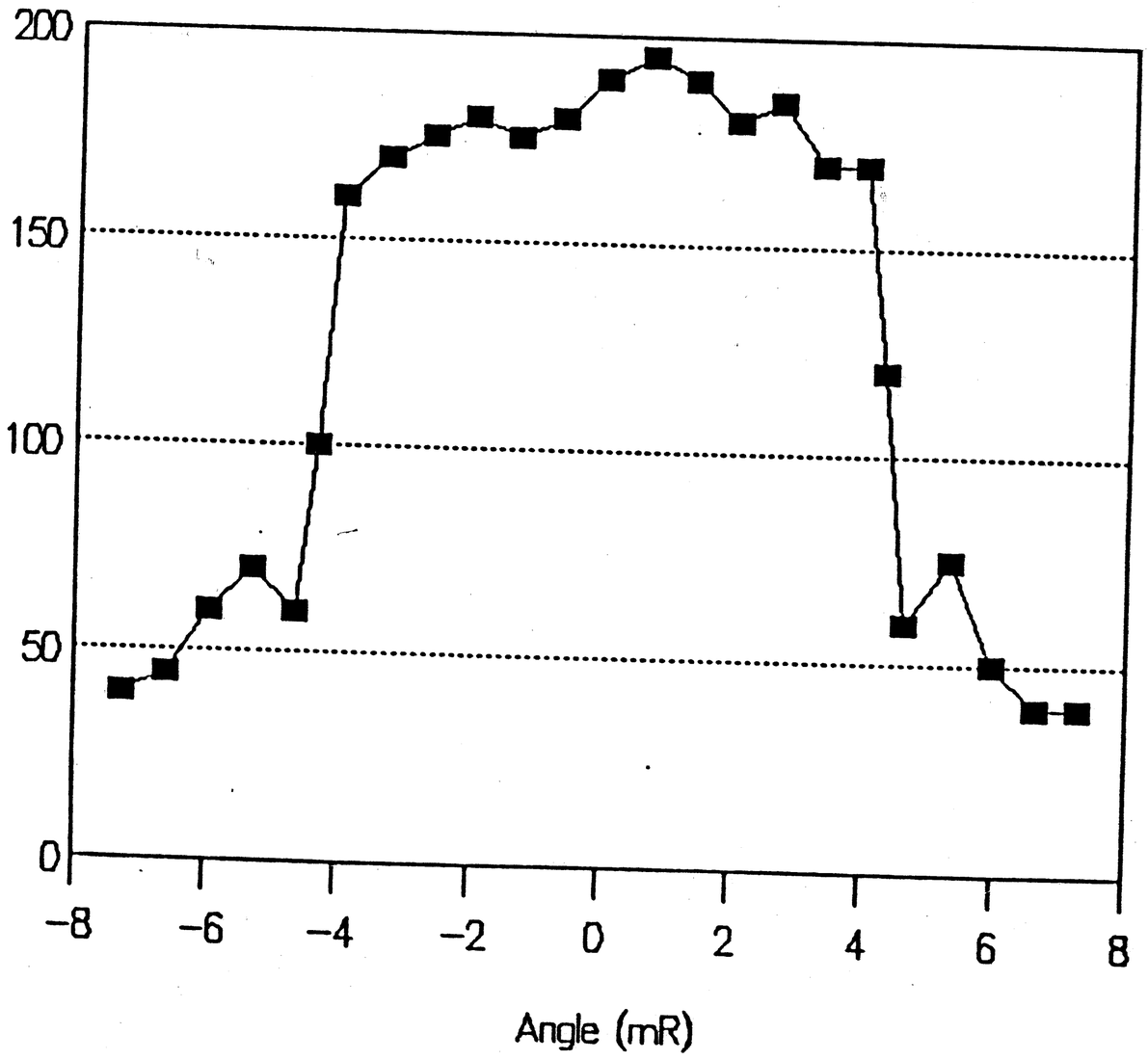
Figure 5-7

overlap of detector and spot areas. This yields less steep edges on the amplitude vs. angle plots in figure 5-5 and 5-6. This is in contrast to the case shown in figure 5-4, where the spot was essentially either *on* or *off* of the detector.

As was shown in section 5.3.6, the angle of acceptance can be increased by using a small collecting lens in front of the detector. A small plastic elliptical lens with focal length 4.5mm was placed 0.9mm from the photodetector. The photodetector and collector lens were then adjusted so as to focus the spot on the detector. With this lens in place, the signal amplitude vs. angle was measured. From the results shown in figure 5-8, one can see that the acceptance angle has been tripled. While either a collecting lens or shortening the objective focal length will increase the angle of acceptance, the collecting lens is the least expensive option of the two.

Lehigh RX with Elliptical Lens

Signal Amplitude vs. Angle



LP04 Elliptical Lens

Focal Length: 4.5 mm

Working Dist: 1.9mm

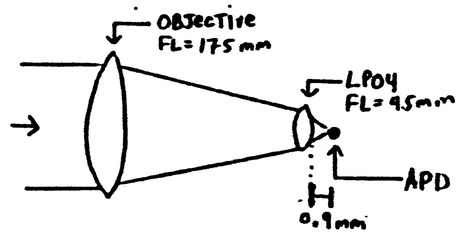


Figure 5-8

6. MODELS OF ATMOSPHERIC PROPAGATION

6.1 Introduction

An atmospheric laser communication system's performance is highly dependant on the state of the atmosphere between the transmitting and receiving stations. The two most significant atmospheric effects are extinction and atmospheric turbulence. Extinction causes energy to be removed from the beam as it propagates between the transmitter and receiver. Extinction is mainly due to scattering by particles in the atmosphere, such as rain, snow, or fog. Atmospheric turbulence causes slight changes in the optical properties of air, which in turn effects optical propagation over long paths. Turbulence causes the received beam to fluctuate in intensity, and causes beam steering over long paths.

6.2 Extinction

As a parallel beam of light passes through the atmosphere, the beam is attenuated due to scattering and absorption by particles in the atmosphere. This attenuation is termed extinction. Extinction is due to two factors; absorption, and scattering. Absorption is due to molecular constituents such as water vapor, carbon dioxide, and oxygen. For water particles at optical wavelengths, absorption is typically negligible, and scattering effects dominate.

Extinction creates an exponential decline in the optical intensity at any point along the beam center, i.e.

$$I = I_0 \, e^{-\alpha x}$$

where α is the extinction coefficient, x is the distance from the transmitter, and I_0 is the initial intensity (optical *power*) at x=0. To convert α to decibels, one can use the formula

$$\alpha_{dB} = 4.34 \, \alpha$$

6.2.1 Extinction Measures

Visual range and meteorological range measurements are used by the National Weather Service to categorize atmospheric visibility conditions and extinction. The concept of visual range is derived from ideas of contrast attenuation and the visual threshold at the wavelength at which the eye has the greatest sensitivity, that is 0.55μm. Visual range is the distance, under daylight conditions, where the contrast between a specified object and it's sky background is just above the threshold contrast of an observer (McCartney). The visual range is a function of the atmospheric extinction, as well as the target and the observer's threshold contrast. Values of visual range can be found from

$$R_V = \frac{1}{\alpha} \ln \frac{C}{\epsilon}$$

where C is the inherent contrast of the target against the background, α is the extinction coefficient, and ϵ is the threshold contrast of the observer.

By specifying a fixed target contrast and threshold, the subjective factors in visual

range measures can be eliminated. Meteorological range, R_m, assumes a visual threshold with ϵ=0.02, and specifies a black target against the sky, thus giving unity contrast. Using the above formula for visual range, the meteorological range, R_m, is defined as

$$R_m = \frac{1}{\alpha} \ln \frac{1}{0.02} = \frac{3.912}{\alpha}$$

where R_m is in kilometers. This is a very useful relationship which can be used to quickly find, for example, loss per kilometer under various weather conditions. Table 6-1 below shows the meteorological range, extinction coefficients, and loss in dB/km over a range of common weather conditions. As can be seen, the losses have a very wide range.

TABLE 6-1 Extinction and Meteorological Ranges in Various Weather Conditions

Visibility Code	Meteorological range	α (Extinction, km^{-1})	Loss
Dense Fog	40-70m	48	-210 dB/km
Thick Fog	70-250m	23	-100 dB/km
Moderate Fog	250-500m	11.5	-50 dB/km
Light Fog	500m-1000m	6	-26 dB/km
Thin Fog	1-2 km	3	-13 dB/km
Haze	2-4 km	1.38	-6 dB/km
Light Haze	4-10 km	0.57	-2.5 dB/km
Clear	10-20 km	0.27	-1.2 dB/km
Very Clear	20-50 km	0.10	-0.45 dB/km
Extremely Clear	50-150 km	0.01	-0.04 dB/km
Scattering Limit	310 km		

6.2.2 Scattering Theory

When a beam of light intercepts a particle, some energy is absorbed and some is redirected by diffraction. This results in extinction of the beam. Scattering theory provides methods to calculate the extinction coefficient given some knowledge about the size and number of particles per unit volume. There are two types of scattering: Rayleigh scattering, and Mie scattering. The type of scattering depends on the relative size of the particle to the wavelength. A complete analysis of scattering theory is beyond the scope of this section. Here, only the main results and applications are discussed. A complete analysis can be found in (van de Hulst).

6.2.2.1 Rayleigh Scattering

Rayleigh has derived results for scattering by particles much smaller than the wavelength. Thus, scattering by molecular sized particles ($r \ll \lambda$) is often termed Rayleigh scattering. The attenuation due to Rayleigh scattering, α_{SR}, is

$$\alpha_{SR} = \frac{0.827\ N\ A_p{}^3}{\lambda^4}$$

where N is the number of particles per unit volume in the path, and A_p is the cross sectional area of the scattering particle (Pratt). In the formula, the particle area and wavelengths are in centimeters. Note that the scattering coefficient is inversely proportional to λ^4, thus short wavelength light is scattered much more than longer wavelength light. This accounts for the blue color of the daylight sky, since shorter

wavelength components passing through the atmosphere will be scattered by air molecules to the eye more than longer wavelengths.

6.2.2.2 Mie Scattering

A rigorous theory of scattering from spheres of arbitrary size, derived from Maxwell's equations, has been derived by Mie in 1908. Scattering by particles large compared to the wavelength is termed Mie scattering, even though the theory is valid for small particles as well.

6.2.2.3 Mie Scattering Analysis Assumptions

In Mie's analysis, there are three main assumptions. First, that the incident light is monochromatic. Second, that there is single scattering by independent, spherical particles. Single scattering means that the particles are far from one another, so that each is illuminated by the incident light, and that no particle is illuminated by the light scattered from another particle. Lastly, the observer is assumed to be an infinite distance from the particles; thus, any scattered light is considered lost, even if it is scattered at small angles.

6.2.2.4 Mie Scattering Formula

In a medium with N particles per unit volume, the intensity scattered per unit volume is simply N times the intensity intercepted by the cross section of a single spherical particle of radius r. The extinction coefficient α is given by (van de Hulst)

$$\alpha = N \pi r^2 Q_{ext}(r)$$

The term $Q_{ext}(r)$ is the scattering efficiency, described below, which accounts for diffraction effects. If the particles have different radii, with N(r) dr particles with radii between r and r+dr per unit volume, then

$$\alpha = \int_0^\infty N(r) \pi r^2 Q_{ext}(r) dr$$

6.2.2.5 Scattering Efficiency

The scattering efficiency factor for extinction, Q_{ext}, is defined as the ratio of the effective scattering cross sectional area of a particle to the actual geometrical cross sectional area. Here, the effective scattering cross sectional area is equal to the area needed to intercept as much incident energy as is scattered by the particle. Using a combination of geometrical optics and Huygen's principle, one can obtain the following formula for the scattering efficiency (Van de Hulst):

$$Q_{ext}(r) = 2 - 4e^{-\rho \tan\beta}\left(\frac{\cos\beta}{\rho}\right)\sin(\rho-\beta) - 4e^{-\rho \tan\beta}\left(\frac{\cos\beta}{\rho}\right)^2\cos(\rho-2\beta) + 4\left(\frac{\cos\beta}{\rho}\right)^2\cos(2\beta)$$

where $\rho = \frac{4\pi r}{\lambda}(n-1)$, $x=\frac{2\pi r}{\lambda}$ as above, and $\beta = \tan^{-1}\left(\frac{n'}{n-1}\right)$, where n' and n are the imaginary and real parts of the refractive index, respectively. At 0.63μm, water and ice have an index of refraction of $1.33 + i\cdot10^{-6}$.

A plot of $Q_{ext}(r)$ versus particle size r for λ=0.830μm, n=1.33+i·10^{-6} is shown in figure 6-1. As can be seen from the plot, a particle scatters the highest amount of energy when the wavelength is approximately the same as the particle radius. Furthermore, from this plot, one can easily see that if the particle size is equal to or greater than the wavelength, then the scattering efficiency is fairly constant, but if the particle size is less than the wavelength, scattering efficiency falls off rapidly. If the wavelength is increased so that $\lambda \gg r$, then scattering efficiency and scattering losses decrease dramatically.

6.2.2.6 The Scattering Paradox

Note that the scattering efficiency of a large sphere is Q=2, that is, it scatters twice the energy falling on it's geometrical cross section. This is a noteworthy paradox, that a particle can remove from the incident beam twice the optical power that it can intercept. Common sense dictates that a flower pot in a window prevents only sunlight falling on it from entering the room, not twice this amount. Recall, however, that the analysis assumes that the observer is infinitely distant, so that all the scattered light, including that diffracted (scattered) at small angles, is counted as removed from the beam. Thus, at a large to infinite distance, an observer will see that the pot will screen out twice the light, half due to diffraction effects, half intercepted by the pot itself (Van de Hulst).

6.2.3 Scattering by Atmospheric Particles

Now, using simple scattering theory and models of rain, snow, and fog particle sizes,

Efficiency Factor Q vs Particle Radius

For 0.830 Micron Wavelength

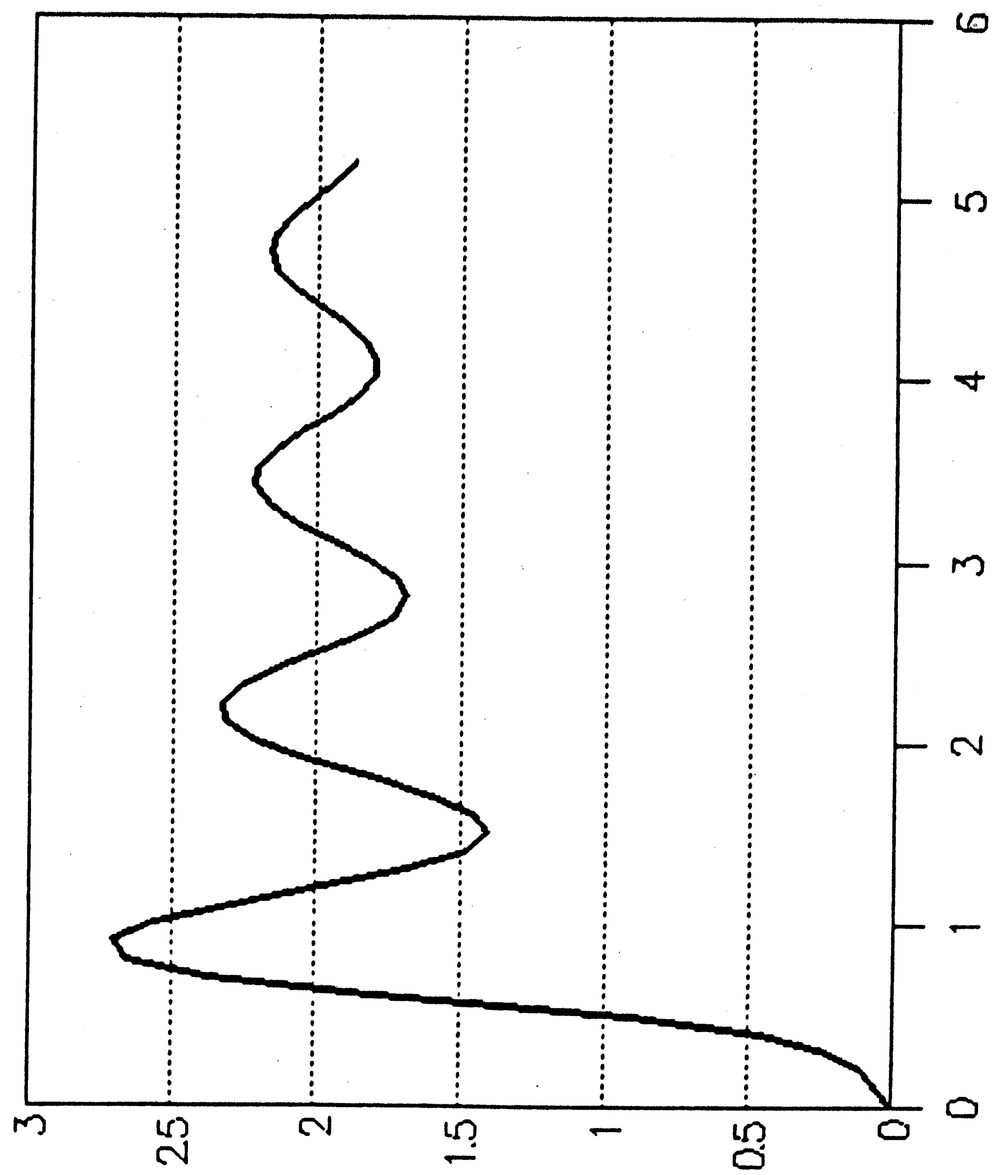


Figure 6-1

optical losses per kilometer through the atmosphere are easily derived.

6.2.3.1 Atmospheric Particles

The size ranges of particles encountered in the atmosphere are shown in table 6-2 (Houghton). Atmospheric particles can be divided into two categories: aerosols and hydrometeors. Aerosols are particulate matter suspended in the atmosphere, including smog, smoke, haze clouds, fog, and fine soil particles. These particles are generally less than one micron. Hydrometeors are water particles in solid or liquid form, such as rain, mist, snow, ice pellets, hail, ocean spray, clouds, and fog. Generally, the particles that cause the most scattering are hydrometeors, since they are occur in the largest numbers in the atmosphere.

TABLE 6-2 Atmospheric Particle Sizes

Hydrometeors		Aerosols	
Particle	Approximate Sizes	Particle	Approximate Sizes
Hail	1mm - 1cm	Dust	0.1μm - 1mm
Rain	0.1mm - 1cm	Haze	1nm - 10μm
Mist	0.1mm	Smoke	1nm - 1μm
Ocean Spray	1μm - 1mm	Smog	1nm - 1μm
Fog & Clouds	0.1μm - 100μm		

6.2.3.2 Extinction due to Water Particles

Since the particles that cause the most atmospheric scattering are hydrometeors, it

is desirable to calculate the scattering losses due to these particles. Water is practically lossless at visible wavelengths, so most of the losses due to water particles are due to scattering effects, not absorption.

Although there are never consistently particles of a single radius in the atmosphere, it is useful to calculate the attenuation coefficient for this ideal case. Given w mg/m^3 of spherical water particles, the number of particles for a given radius r is

$$N(r) = \frac{w}{\frac{4}{3}\pi r^3}\delta(r)$$

where $\delta(r)$ is the Dirac delta function. Substituting N(r) in

$$\alpha = \int_0^\infty N(r)\ \pi\ r^2\ Q_{ext}(r)\ dr$$

and normalizing by dividing through by w, we obtain:

$$\alpha = \frac{3\ Q_{ext}(r)}{4\ r}\left[\frac{\text{nepers/km}}{mg/m^3}\right] = 3.25\ \frac{Q_{ext}(r)}{r}\left[\frac{dB/km}{mg/m^3}\right]$$

where the radius of the water sphere r is in microns (Chu & Hogg).

Rain, typically with $1000mg/m^3$, has particle radii from approximately 0.1mm to 10mm. Fog, on the other hand, with $100mg/m^3$, has particle radii typically 1/10,000th that of a raindrop. Using these figures in the above equation, it is easy

to show that fog has a higher attenuation. This is because even though rain has a higher water content than fog, fog has a greater surface area available for scattering, and thus yields higher losses. (Chu & Hogg)

6.2.3.3 Extinction due to Rain

Recall from above that for water spheres of radius r, the extinction coefficient is

$$\alpha_{dB} = 3.25 \frac{Q_{ext}(r)}{r} \left[\frac{dB/km}{mg/m^3} \right]$$

Define d as the number of milligrams of water particles of radius r per cubic meter. Then, the loss due to these particles is

$$\alpha_{dB} = 3.25\ d \frac{Q_{ext}(r)}{r}\ dB/km$$

Here, d can be found by forming the ratio of the volume of water collected in area A to the distance swept out by an area A falling at the speed of the raindrops:

$$d = \frac{R\ A}{v \cdot t \cdot A}$$

R is the rainfall rate (mm/hr), A is a suitable cross sectional area (1 m^2), and the v·t product is the distance a raindrop falls in a unit time period (m/hr).

Rain droplets come in a spectrum of sizes, and the distribution of sizes depends on the rainfall rate. Laws and Parsons have modeled the distribution of drop sizes from experimental data. The precipitation model is expressed as

$$n(r,R) = \frac{2}{\sqrt{\pi}} \cdot \frac{0.4}{\sigma(r)} e^{-[r-r_0(R)]^2/\sigma^2(R)}$$

where n(r,R) is the normalized percentage distribution, R is the rain rate in mm/hr, r_0 is the mean particle radius in cm, σ is the variance of the distribution in cm, and 0.4 is the normalizing constant. Also.

$$r_0(R) = 0.045 \log_{10}(R) + 0.05 \text{ cm}$$

and

$$\sigma(R) = 0.05 \log_{10}(R) + 0.05 \text{ cm}$$

Plots of the drop size distribution for rain rates of 1, 10, and 100 mm/hr are shown in figure 6-2.

In order to simplify analysis, a single droplet size is assumed. From empirical studies by Laws & Parsons (Houghton, p. 1046), the mean diameter of raindrops is given by

$$r = 0.619 R^{0.182}$$

where r, the drop radius, is given in mm, and R, the rainfall rate, is given in mm/hr. Empirical studies have also shown (Wang) that the terminal velocity of a raindrop can be approximated by

$$v=200\sqrt{r}$$

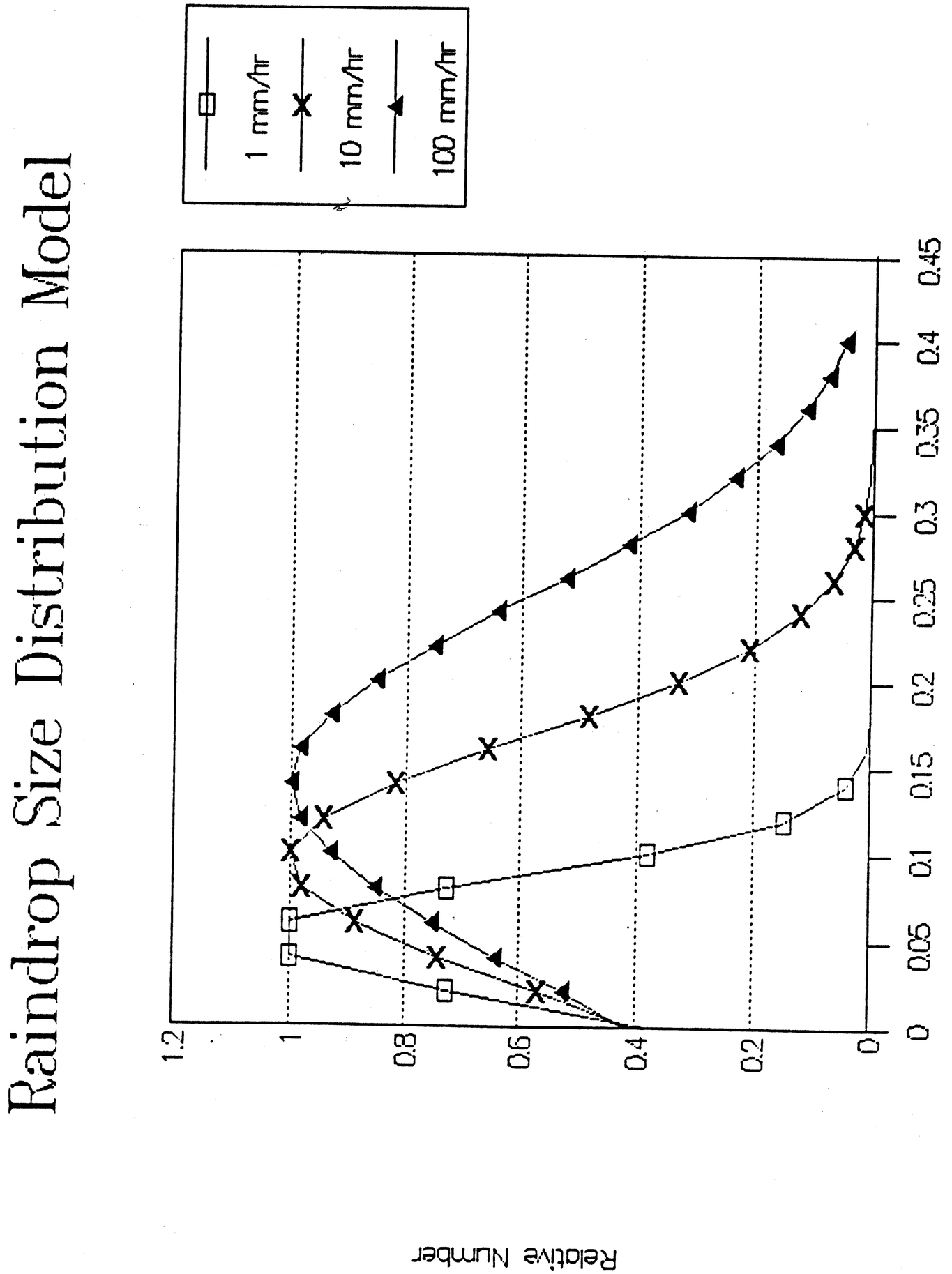

Raindrop Size Distribution Model
1 mm/hr
10 mm/hr
100 mm/hr
Relative Number
1.2
1
0.8
0.6
0.4
0.2
0
Drop Size (cm)
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45

Figure 6-2

where v is in meters/second, and r is in meters.

By assuming $Q_{ext}(r)=2.0$ (since $r \gg \lambda$ for raindrops at optical wavelengths), and using the above three expressions, we can solve for α_{dB} as a function of R. The result is

$$\alpha_{dB} = 0.5862\ R^{\,0.727}\ \text{(dB/km Loss)}$$

where α_{dB} is in dB/km, and R is in mm/hr. A graph of loss versus rainfall rate is shown in figure 6-3. Although the graph is not linear, a linear approximation over the range 10 to 100mm/hr gives a loss of 0.15 dB/km per mm/hr.

Using more realistic raindrop size distributions, such as the Laws & Parsons distribution (Laws & Parsons), it has been shown (Chu & Hogg) that the loss at 0.63μm is approximately 0.22 dB/km per mm/hr. However, a more sophisticated model that also includes forward scattering effects (Chu & Hogg) yields losses of 0.155 dB/km per mm/hr.

Typical values of R, the rainfall rate, are 0.25mm/hr (drizzle), 1mm/hr (light rain), 4mm/hr (moderate rain), 16mm/hr (heavy rain), and 100 mm/hr (extremely heavy rain).

6.2.3.4 Extinction due to Fog

The size of a fog particle is typically in the range of 1μm to 100μm. Assuming a mean of r=10μm as a mean fog particle radius, and recalling that

Attenuation due to Rain

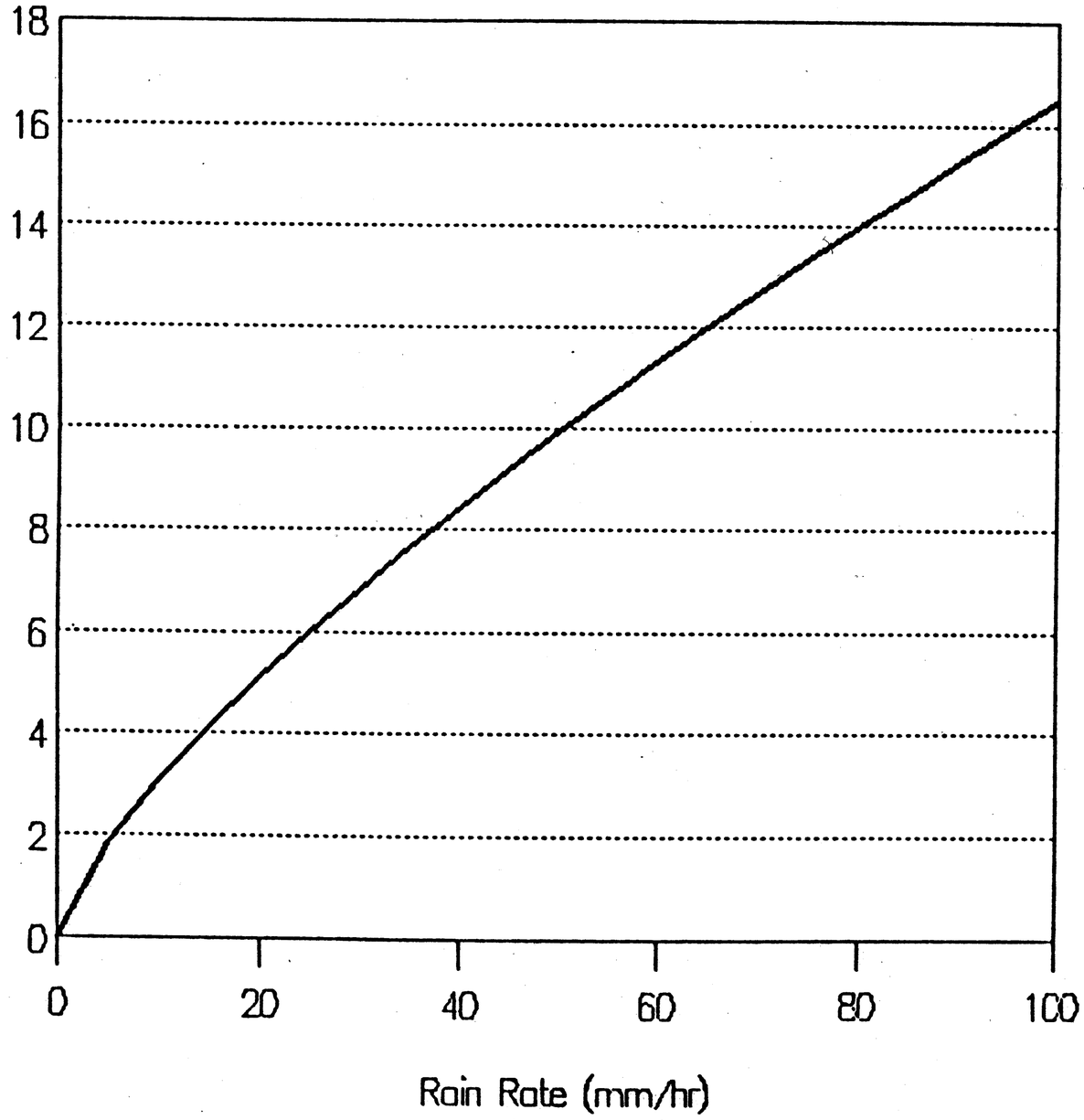


Figure 6-3

$$\alpha_{dB} = 3.25 \frac{Q_{ext}(r)}{r} \left[\frac{dB/km}{mg/m^3} \right]$$

one can find that α_{dB}=0.65 dB/km per mg/m^3 using this simple single-radius model. Using a more sophisticated model using a range of droplet sizes it has been shown (Rensch & Long) that over the wavelengths between 0.3μm and 2.0μm, extinction is approximately constant at 0.4$\pm$0.15 dB/km per mg/m^3. Since water densities in fogs are typically between 10mg/m^3 to 1000mg/m^3, fogs can easily create 100 dB/km losses.

Experimental difficulty and natural variety of fog types has caused much variation in published data on fog particle sizes. The following proposed model (Chu & Hogg) has β and γ parameters which are selected to best match observed fog droplet distributions:

$$N(r/r_c) = A \cdot (r/r_c)^{\beta} \cdot e^{-B \cdot (r/r_c)^{\gamma}}$$

Here, r is the droplet radius, r_c is the radius of drops with the maximum number density, B=β/γ, and A is given by

$$A = (\beta/\gamma)^{(\beta+1)/\gamma} \frac{\gamma}{\Gamma\left(\frac{\beta+1}{\gamma}\right)} \cdot \frac{N}{r_c}$$

where N is the number of drops per unit volume. Typical values of r_c range from

1μm to 30μm. A graph of a typical fog drop size distribution with β=1, γ=1, and r_c=10μm is shown in figure 6-4.

As can be seen from the graph, there is a wide range of particle sizes much larger than optical wavelengths of 0.5μm-1μm. This means that wavelengths in this range are easily scattered by the fog droplets. Only if the wavelength was larger than the majority of the fog droplets would scattering losses decrease. For droplet radii less than 5.0μm, infrared wavelengths do not necessarily guarantee smaller values of extinction compared to shorter wavelengths (Rensch & Long). Observations have shown (Arnulf) that small drop fogs are can be frequency selective, but these fogs are rare. Attenuation by fogs is for the most part not frequency selective in the 0.35 to 2μm region, because most droplets are larger than the wavelength.

The question arises of how to quantitatively describe a fog in the field. By finding a fog's water density, or by using the meteorological range, one can quantitatively measure a fog. Of course, meteorological range is more subjective, but it is more easily measured than fog water density.

6.2.3.5 Extinction due to Snow

Snowfall rates are measured by liquid water content, not by unmelted accumulation. In general, the size and density of snowflakes depends on many factors, including the type of snow storm, which varies by region. Snowflake size and density can even vary over time during the same snowstorm (Gray & Male). Snow densities can range from 50kg/m^3 to 200kg/m^3. A generally accepted average of snow density is

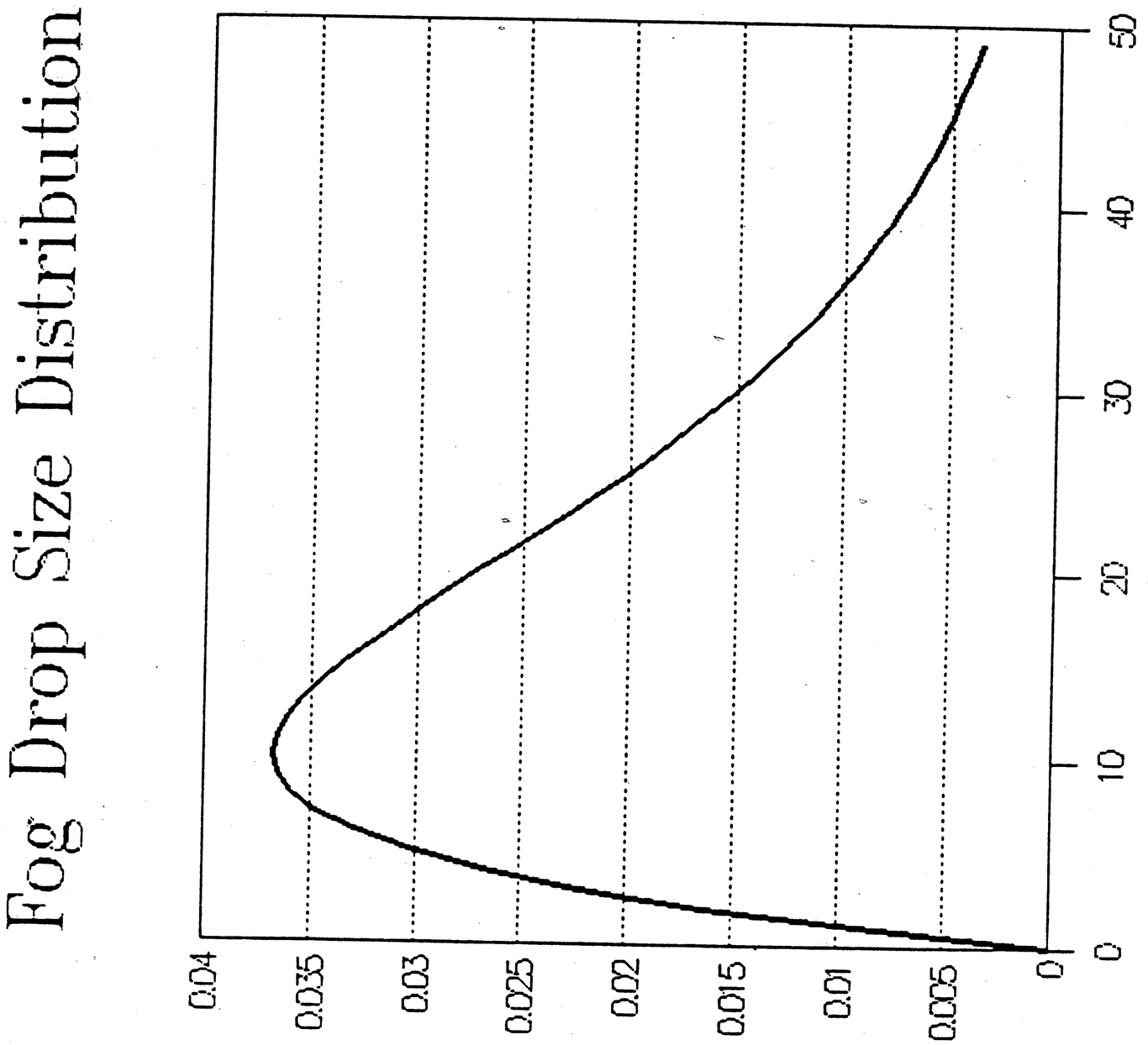

Fog Drop Size Distribution
Rel Density N(r)
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0
10
20
30
40
50
Drop Radius (microns)

Figure 6-4

100kg/m^3. Since the density of pure water is 1000kg/m^3, snow has density is 0.1 relative to water. Thus, for example, 10mm of "dry" snowfall will yield 1mm of water when melted.

Snowflake size distribution is such that relatively few flakes have dimensions greater than 10 mm; for the most part, the particle distribution is in the range between 1mm and 6mm (Lillesaeter). Crystals of falling snow are large in comparison to optical wavelengths, so the scattering efficiency Q, and thus the extinction, should be wavelength independent.

A number of investigators have derived empirical relationships between measured attenuation and snowfall rate (Seagraves) For example, attenuation measurements made at 0.45μm by (Lillesaeter) indicate the relationship

$$\alpha_{dB}=1.693\ R\ (dB/km)$$

where α_{dB} is in dB/km, and R is the dry snow rate in millimeters per hour.

Similarly, Warner's measurements at 0.45μm obtained

$$\alpha_{dB}=1.098\ R\ (dB/km)$$

and O'Brien's results are equivalent to

$$\alpha_{dB}=2.552\ R^{\,0.693}\ (dB/km)$$

These models are plotted in figure 6-5 for *unmelted* snowfall rates up to 50mm/hr. From the graph, one can see that there is some disagreement between the models; however, it is apparent that snow can easily create wavelength independent losses of many 10's of dB's.

6.3 Atmospheric Turbulence Effects

In an atmospheric laser communication link over a long ($\geq$1km) path, atmospheric turbulence causes slight changes in the optical index of refraction of air between the transmitter and receiver. These changes, in turn, cause amplitude fluctuations in the received beam. In order to understand the effects of turbulence on the system, one must first investigate the cause, structure, and effects of atmospheric turbulence. Then, the probability distribution, frequency spectrum, and variance of these fluctuations can be described.

Here we do not attempt to present the derivations of the theoretical results in any complete manner. Only the main results and applications are discussed. Furthermore, the theory uses several approximations, so it is meaningless to worry about high accuracy in calculations. The theory gives a rough quantitative estimate of the phenomena involved.

6.3.1 The Cause and Development of Atmospheric Turbulence

The atmosphere is in a constant state of turbulent motion. This turbulence is

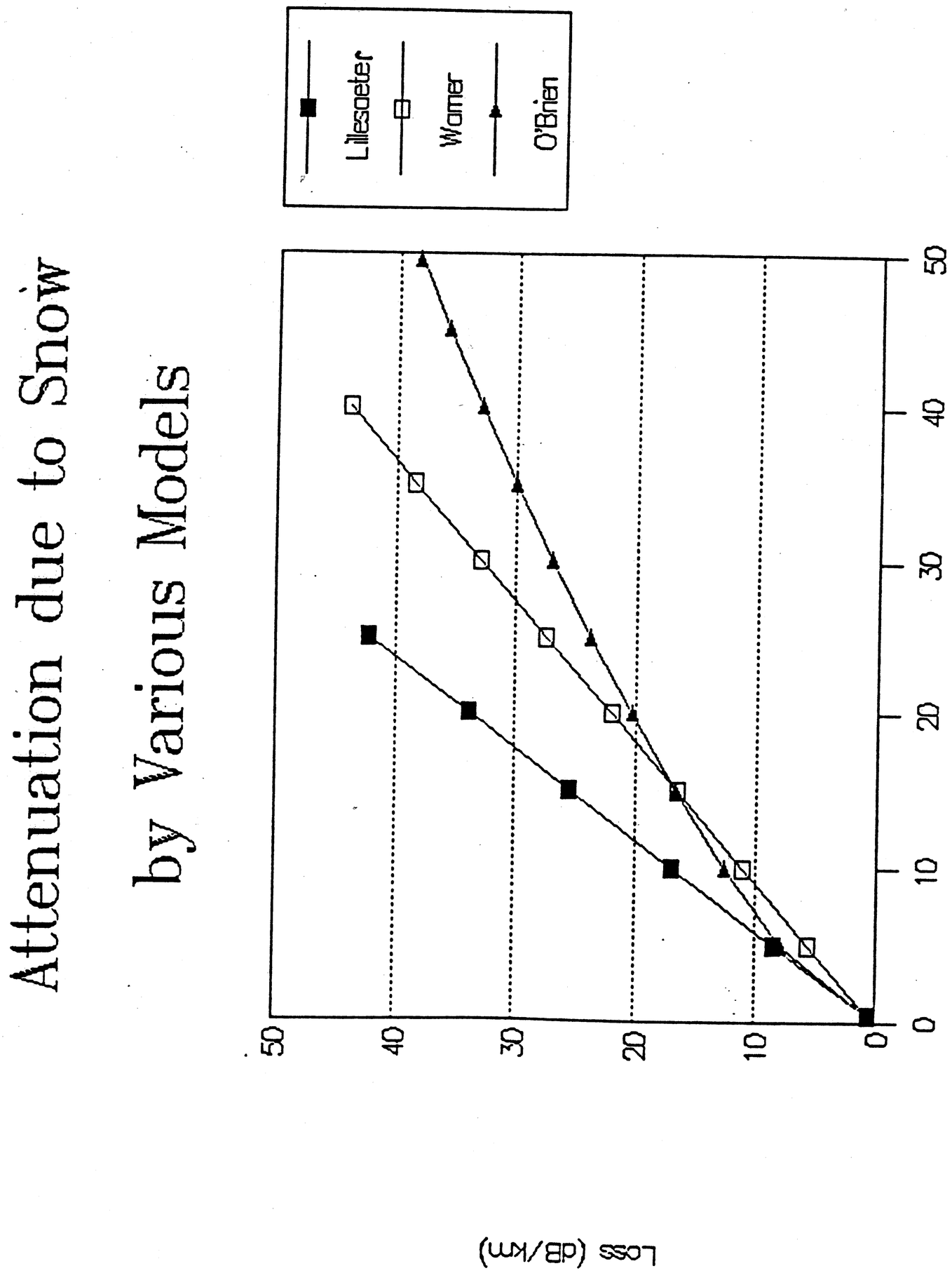

Attenuation due to Snow
by Various Models
Lillesoeter
Wamer
O'Brien
Dry Snow Rate (mm/hr)
0
10
20
30
40
50
Loss (dB/km)

Figure 6-5

introduced by wind or convection from solar heating of the ground, which forms large eddies of heated air. Spatial and temporal measurements of these eddies using microwaves indicate that the eddies move a distance of 0.4 times their size before breaking up. Larger turbulent eddies break up and transfer their energy to smaller eddies. These eddies then form and transfer energy to still smaller eddies, and so on. Eventually, for very small eddies, viscous effects become important, and energy is dissipated.

The lengths l_0 and L_0 are the size of the smallest and largest eddies or "blobs" of turbulence. Usually, l_0 and L_0 are referred to as the inner and outer scale sizes of the turbulence. They can be estimated by

$$l_0 = (10^{-9} \, h)^{1/3}, \; l_0 > 2mm$$

and

$$L_0 = 2 \sqrt{h}$$

where all units are in meters (Brookner). Typical minimum values for l_0 are approximately 1 or 2mm near ground. Typical maximum values for L_0 are approximately 100 meters, and are found in the upper atmosphere.

6.3.2 Optical Effects of Turbulence

Turbulence induced temperature variations cause changes in the index of refraction of air, which in turn effects beam propagation through the atmosphere. The optical index of refraction of air is given by

$$n = 1 + 10^{-6} \cdot \frac{79\,P}{T}\left(1 + \frac{7800\,Q}{T}\right)$$

n=index of refraction T=Absolute Temperature,

P=Pressure in Millibars, Q=Relative Humidity

A rule of thumb is that the index of refraction changes by 10^{-6} for every °C. Refractive index fluctuations are caused almost exclusively by fluctuations in temperature, since pressure variations are relatively small and are rapidly dispersed (Lawrence). These temperature variations change in space and time.

6.3.3 Structure and Measures of Turbulence

It is convenient, from a theoretical viewpoint, to express statistical spatial variations of the index of refraction of the atmosphere in terms of a *structure function*, defined as:

$$D_n(\bar{r}) = \overline{[\, n(\bar{r}_2) - n(\bar{r}_1)\,]^2}$$

where $n(\bar{r}_i)$ denotes the index of refraction at a position represented by the vector $\bar{r}_i$, and $\bar{r}=\bar{r}_2-\bar{r}_1$. Structure functions were introduced to deal with an isotropic random variable that varies in space and time, such as the atmospheric index of refraction. Theoretically, the structure function follows the relation (Tatarski)

$$D_n(|\,\bar{r}\,|) = C_n^{\,2}\,|\,\bar{r}\,|^{\,2/3} \quad \text{for } l_0 \ll r_0 \ll L_0$$

where C_n is the structure constant of the turbulence.

6.3.4 The Structure Constant

The parameter C_n is called the structure constant of the turbulence. It provides a measure of the strength and structure of the atmospheric turbulence. The value of the structure constant varies with altitude and time of day. Typical daytime values at ground level are given in table 6-3. There is no simple rule for estimating C_n; it was originally derived from experimental data. The division into categories of turbulence is somewhat subjective, but it provides a basis for estimating quantitatively the degree turbulence.

TABLE 6-3

Structure Constants for Various Turbulence Levels

Weak Turbulence	$C_n = 8 \times 10^{-9}\ m^{-1/3}$
Intermediate Turbulence	$C_n = 4 \times 10^{-8}\ m^{-1/3}$
Strong Turbulence	$C_n = 5 \times 10^{-7}\ m^{-1/3}$

6.3.5 Turbulence and Received Beam Fluctuations

Turbulence causes amplitude fluctuations (often called scintillations) in a beam that

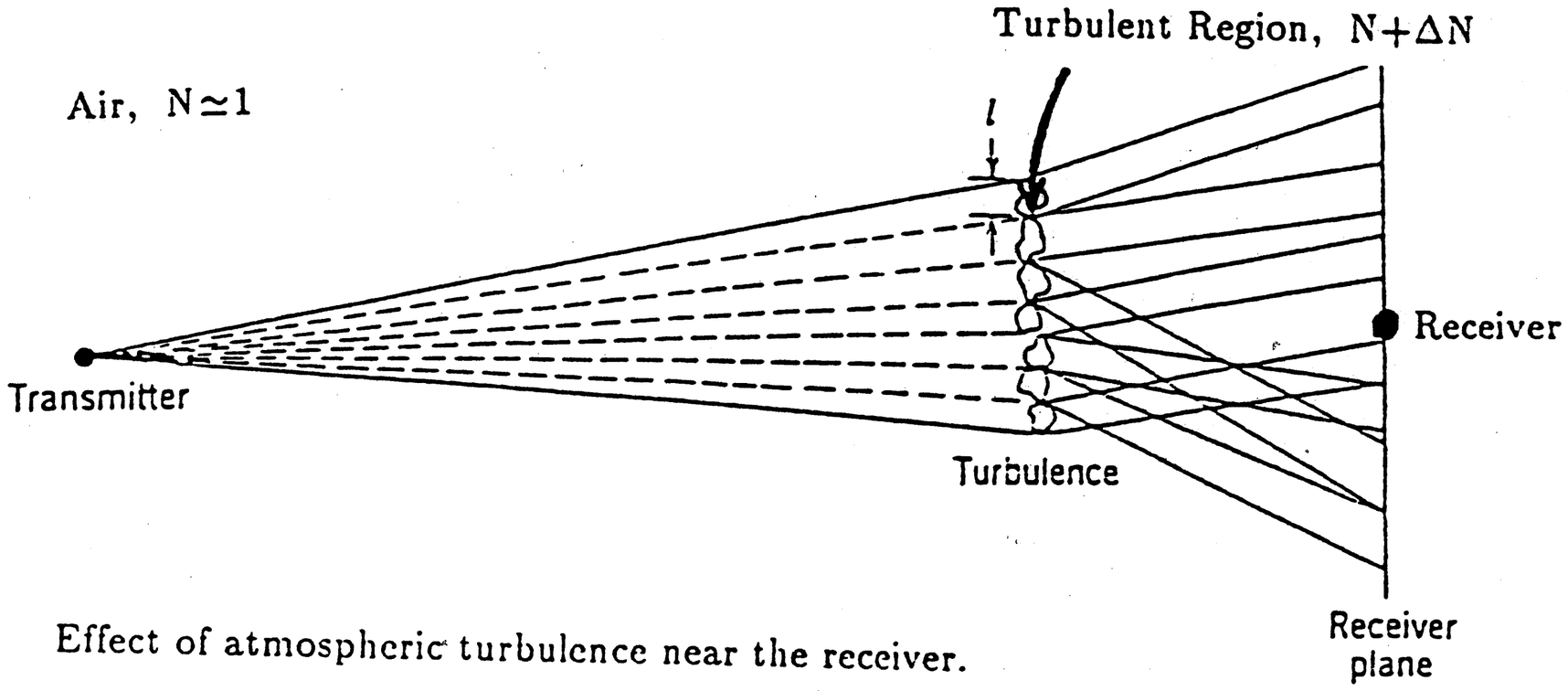
has passed through the atmosphere. A formal treatment of wave propagation in a turbulent medium, such as that found in (Tatarski), is beyond the scope of this section. Instead, simplified models and results are given. These allow us to readily describe the cause of the fluctuations, as well as the variance, spectrum, probability distribution, and other characteristics of the fluctuations.

6.3.5.1 Cause of Fluctuations

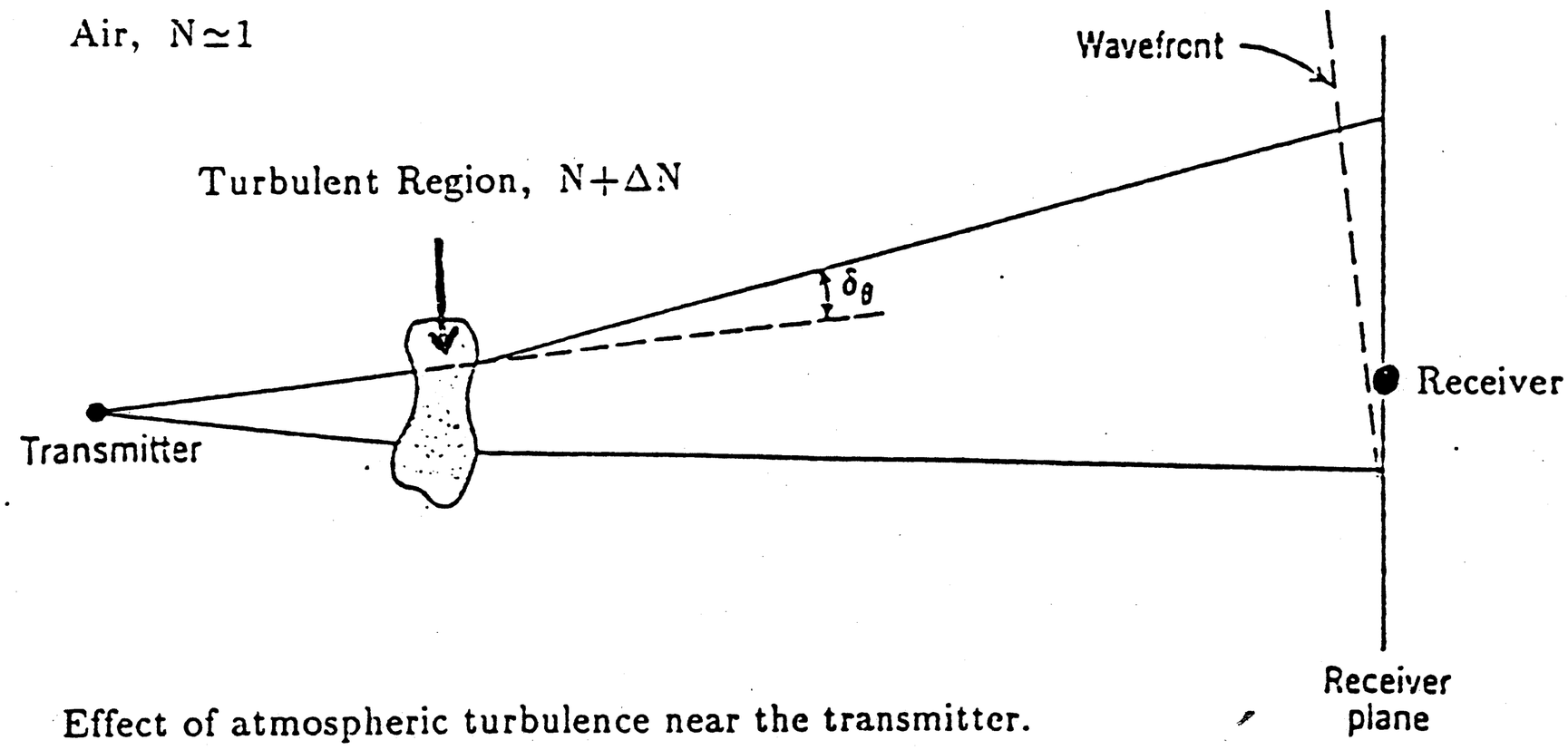
It is convenient to consider the turbulent atmosphere to be composed of eddies of index of refraction $n+\Delta n$ in a medium of index n. Here, n is the index of refraction of the surrounding air, and Δn is the change in index of refraction due to the temperature variation of the eddy. If the eddy has a dimension l, where $l_0<l<L_0$, then the effect of the eddy depends on the relative sizes of the beam diameter d_B, to the eddy dimension l. If $d_B \ll l$, then the effect of the eddy is to deflect the beam as a whole. These angular deflections are typically on the order of 1 to 10 microradians per kilometer. At the receiver, the beam appears to follow a two dimensional random walk in the receiver plane. For $d_B \simeq l$. the turbulent eddies act as lenses which focus or defocus all or parts of the beam, causing a granular structure to the cross section of the beam. If $d_B \gg l$, then small portions of the beam are independently diffracted, and the phase fronts are badly distorted. The effects of $d_B \gg l$ and $d_B \ll l$ are shown in figure 6-6 (Pratt).

In the receiver plane, the received beam intensity varies in time and space. The resulting spacial amplitude variations occur on a variety of scales, however the predominant scale size is the Fresnel-zone size $\sqrt{\lambda L}$, where λ is the wavelength

Optical Effects of Atmospheric Turbulence



Effect of atmospheric turbulence near the receiver.



Effect of atmospheric turbulence near the transmitter.

Figure 6-6

and L is the path length. Predominant scale sizes are approximately 2 cm for a 1 km path at optical wavelengths.

6.3.5.2 Probability Distribution of Fluctuations

For plane wave propagation it has been shown (Tatarski), by use of a perturbation approximation to the wave equation, that the amplitude fluctuations introduced over an optical path have a lognormal distribution with a variance σ^2. A simplified explanation of this lognormal distribution can be obtained if one considers the model depicted in figure 6-7. The transmitted beam crosses an optical path between the transmitter and receiver. The path is divided into N "slabs" of width d, where d is larger than outer scale size of the turbulence. The turbulence in each slab prevents a portion of the beam from reaching the receiver; thus, each slab has an optical transmission coefficient $T_i<1$. The transmission coefficient for the entire path, T_{Path}, is

$$T_{Path} = \prod_{i=1}^{N} T_i$$

Taking the log of each side of this equation,

$$\text{Log}\,(T_{Path}) = \sum_{i=1}^{N} \text{Log}(T_i)$$

Now, since d is larger than the outer scale size of the turbulence, the transmission coefficients of each slab are independent random variables. Thus, $\text{Log}(T_{Path})$ is the

Beam Transmission Model

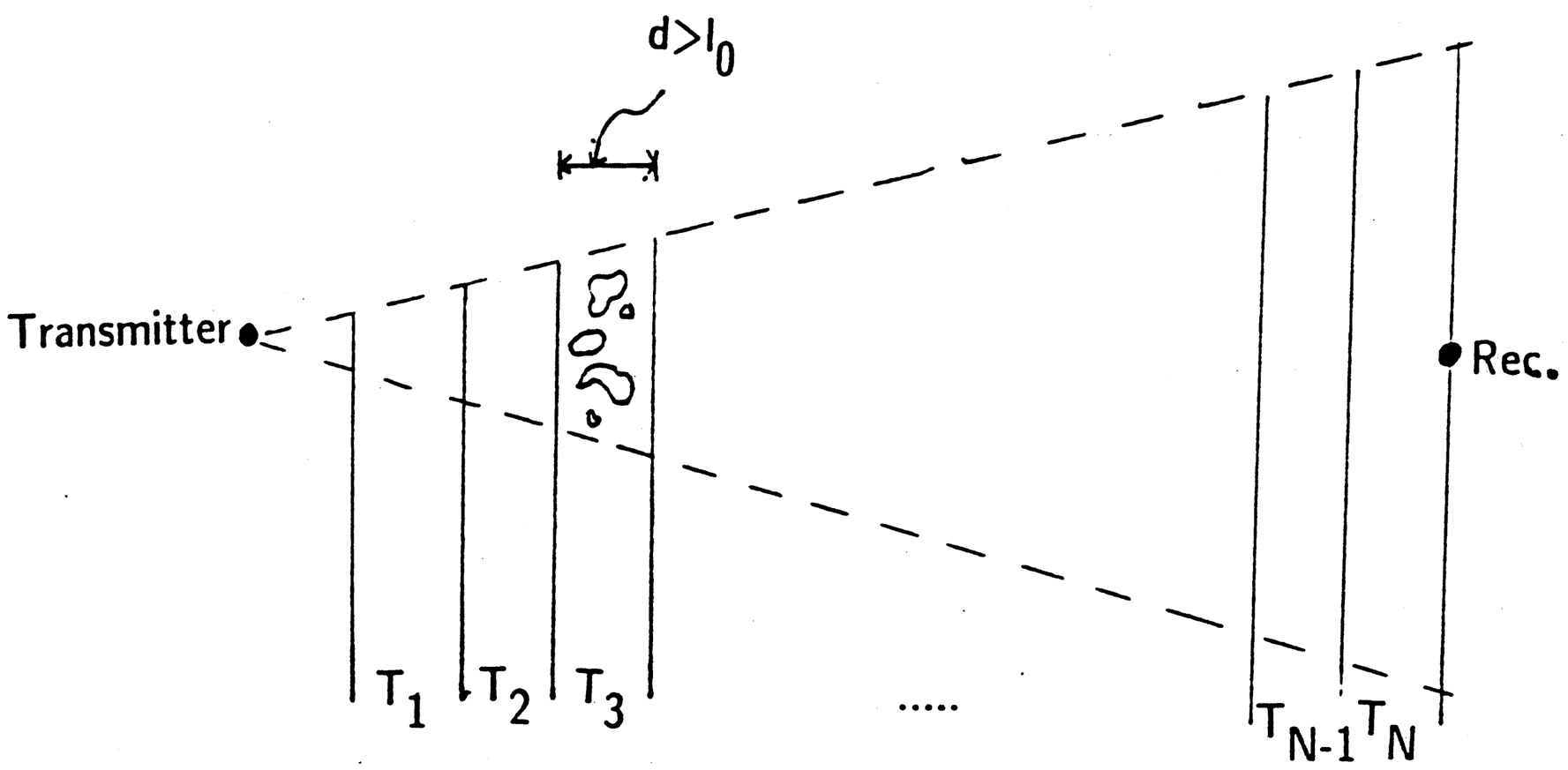


Figure 6-7

sum of N uncorrelated terms, where N is large. Using the mean value theorem, the result is a normal distribution of $\mathrm{Log}(T_{Path})$.

6.3.5.3 Variance of Atmospheric Induced Fluctuations

For propagation along a horizontal path, the variance of the amplitude fluctuations is given by (Tatarski):

$$\sigma^2 = \overline{\left[\ln\left(\frac{A}{A_0}\right)\right]^2} = 0.31\, C_n^{\,2}\, (2\pi/\lambda)^{7/6}\, L^{11/6} \quad \text{for} \quad L_0 \gg \sqrt{\lambda L} \gg l_0$$

Here, A is the signal amplitude, A_0 is the mean amplitude, C_n is the structure constant of the turbulence, λ is the wavelength, and L is the path length. A plot of standard deviation ($\sqrt{\sigma^2}$) vs. distance for typical intermediate turbulence is shown in figure 6-8.

6.3.5.4 Saturation of Fluctuations with Distance

As the path length L becomes very large, as in observing a star through the atmosphere, the above formula predicts large fluctuations. Clearly, this result seems inadequate, since experience shows that stars do not 'twinkle' with extremely large amplitudes.

Experimental results show that the variance of the lognormal amplitude fluctuations increases with distance and turbulence strength up a certain threshold. Beyond the threshold, one observes what has been termed a saturation effect, where σ remains

Amplitude Variations vs. Path Length for Intermediate Atmospheric Turbulance

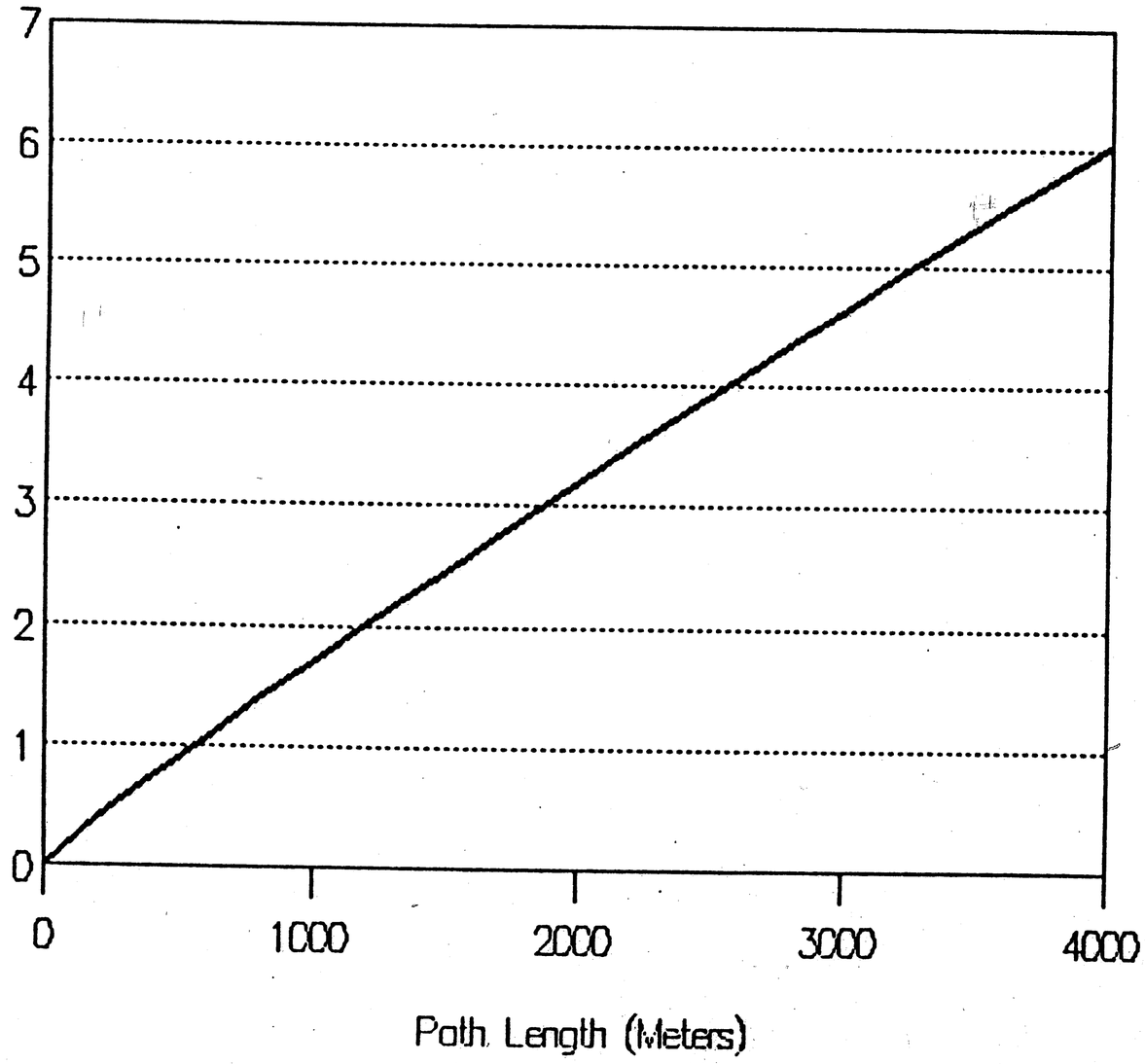


Figure 6-8

essentially constant as either the turbulence strength or path length increases. Typically, for paths above ground, saturation may occur within a few kilometers (Lawrence).

The saturation effect has not been sufficiently explained as yet by theoretical models. The only worker that appears to derive a saturation like curve for the variance of the log intensity fluctuations is (deWolf), who forwards the following model:

$$\sigma_d^2 = \sum_{n=0}^{\infty} \frac{1}{(n+m)^2}$$

$$m = \frac{1}{1 - e^{-\sigma^2}}$$

$$\sigma^2 = 1.23\, C_n^2 (2\pi/\lambda)^{7/6}\, L^{11/6}$$

where σ^2 is the variance calculated in (Tatarski). Using this model, the standard deviation, $\sqrt{\sigma^2}$, is plotted in figure 6-9 with a value of $C_n = 5\times10^{-8}$. As can be seen from the figure, the fluctuations begin to saturate after a few kilometers, at a standard deviation of approximately 5 to 6dB.

6.3.5.5 Fluctuations and Aperture Diameter

The results thus far have been derived for a point receiver. If the aperture size, d, is increased, the fluctuations remain unchanged while $d < \sqrt{L\lambda}$ (Hohn). For a L=1km

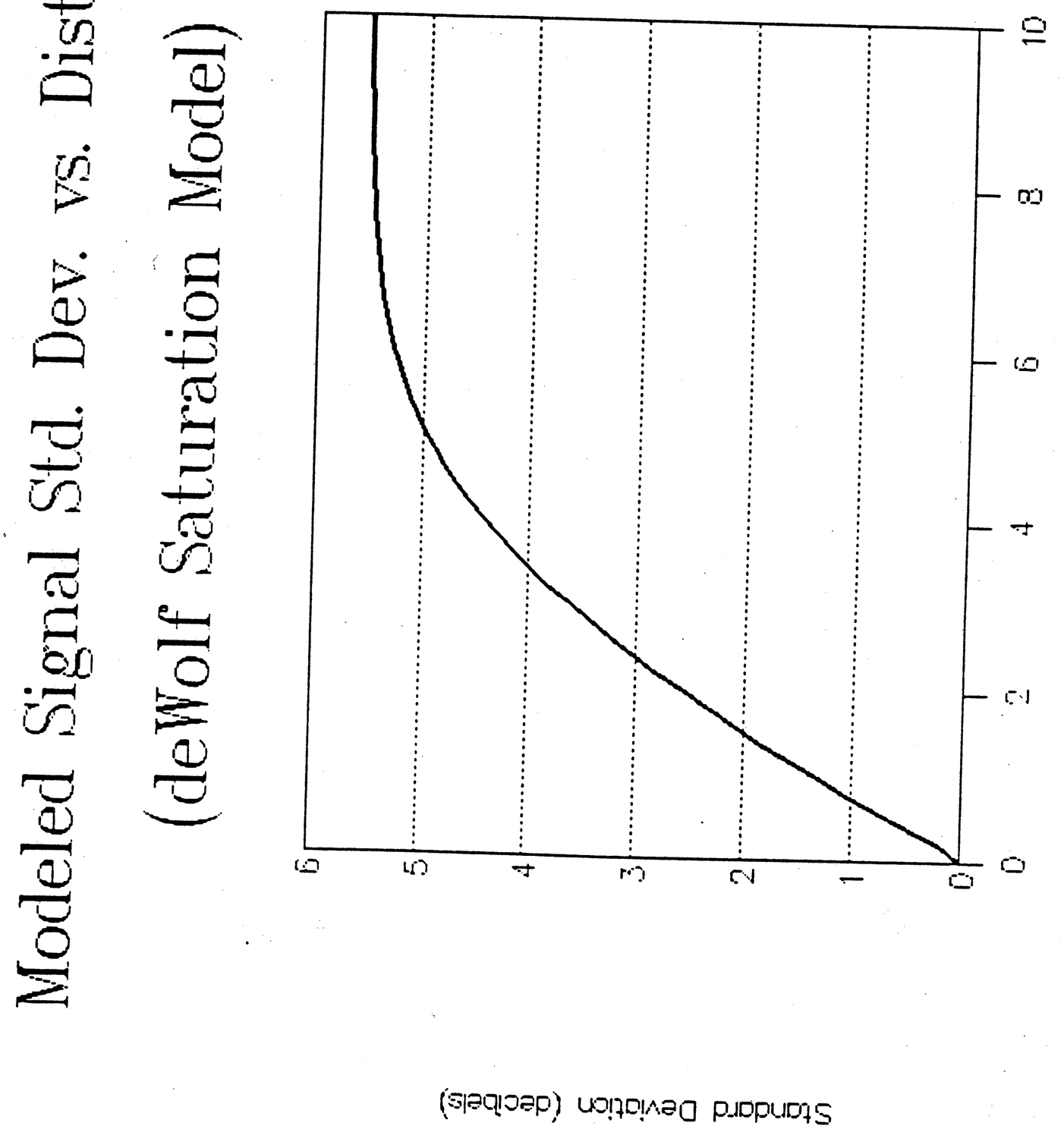

Modeled Signal Std. Dev. vs. Distance
(deWolf Saturation Model)
Standard Deviation (decibels)
0
1
2
3
4
5
6
0
2
4
6
8
10
Distance (Meters) (thousands)

Figure 6-9

link at λ=0.83μm, this means d<29mm. This is the case in many practical systems.

When, however, the receiver aperture is much larger than the characteristic length of the turbulence elements, i.e. d $\gg\sqrt{L\lambda}$, then increasing the size of the collecting aperture decreases the variance of the atmospheric induced fluctuations. This decrease is due to what is termed "aperture averaging". This averaging occurs because the light focused on the photodetector is the sum of all the spatially varying light intensities across the lens' surface. If the above condition is are met, then the standard deviation of the lognormal fluctuations, $\sqrt{\sigma^2}$ can be shown (Fried) to follow

$$\sqrt{\sigma^2} = \frac{K}{d}$$

where K is a constant determined by the signal variance and aperture diameter.

Since changing the aperture diameter does not effect the time scales on which turbulence develops, the spectrum of the fluctuations is expected to be independent of the receiving aperture diameter. This has been shown to be the case (Subramanian).

6.3.5.6 Spectrum of Fluctuations

Time variations in the received signal strength are caused by turbulent eddies moving through the beam with the wind. A complete analysis (Tatarski) of the spectral characteristics of the log amplitude fluctuations due to turbulence predicts an increasing spectrum up to a "breakpoint" frequency f_0, after which the power

spectrum decreases as $f^{-8/3}$. The breakpoint frequency f_0 is found by dividing the wind velocity perpendicular to the beam by the size of the turbulence eddies that have the greatest contribution to amplitude fluctuations. Since the predominant turbulence scale size is $\sqrt{\lambda L}$ (Davis), one can calculate f_0 by

$$f_0 = \frac{v_n}{\sqrt{\lambda L}}$$

Here, v_n is the wind velocity normal to the beam path, λ is the wavelength, and L is the path length. This formula is valid if the fluctuations are not in the saturation regime. For λ=0.830μm, L=1 km. and v_n=2 m/s, the breakpoint frequency is f_0=70Hz.

For frequencies below and above the breakpoint frequency, the relative amplitude of the spectrum W(f) is approximated by (Lawrence):

$$W(f) = \frac{1}{f_0}[\ 1 + 0.48(f/f_0)^{4/3} + \ldots\ ] \quad \text{for } f \ll f_0$$

$$W(f) = \frac{7.6}{f_0}\,(f/f_0)^{-8/3} \quad \text{for } f \gg f_0$$

Typically, $f_0 \simeq 100$Hz, and most of the signal energy is below 500Hz.

6.3.6 Beam Deviation on Long Paths

Atmospheric turbulence also causes what is called *beam steering*, where the entire

beam is deflected. This steering has been estimated at 1 to 10 microradians per kilometer along the horizontal and vertical axes. However, over long distances of many tens of kilometers, the beam has been noted to move primarily in the *vertical* direction. For example, over a 45 km path passing over Boulder Colorado, (Ochs et al) observed beam wandering in the vertical direction by several hundred microradians. Experimental studies on a 36 km path (King et al) show that the beam can move vertically by 1mR at a rate of 0.1μR/s. These changes are most notable at sunrise, presumably due to ground warming and rising heated air. This explains the fact that no horizontal movement was noted on either link.

7. MEASUREMENTS OF ATMOSPHERIC PROPAGATION

7.1 Introduction

In order to determine the reliability of atmospheric laser communication links, weather effects were studied on two links in the field. Using the models in chapter 6 along with the collected data, there are two main results. First, that rain, snow, and fog can increase link optical losses by many 10's of dB's per kilometer. Second, atmospheric turbulence can cause the received signal to fluctuate by many dB's. These losses decrease the SNR at the receiver.

7.2 Experimental Link Description

In order to measure the effects of weather, two experimental laser communication links were established. To account for all weather types, the links were monitored for one full calendar year. The received signal levels on these links are continuously monitored by a computer monitoring system. The system consists of a logarithmic preamplifier interface between the receivers and the computer, the computer system itself, and the software for monitoring and analysis. Collected data can be correlated with weather information from the National Weather Service in order to quantitatively determine performance under various weather conditions.

7.2.1 Link Location

There were two atmospheric laser communication links that were tested. The

transmission path for both links passes over the city of Bethlehem, PA, through the Lehigh Valley, and over the Bethlehem Steel works and the Lehigh River.

The first link used the electronics and optics described in the previous chapters. The transmitter, with it's 1mR beamwidth, was located on the roof of Bethlehem Steel's Martin Tower. For stability, the transmitter was mounted on a 55 gallon barrel full of water. The receiver was located 3km away, at Packard Lab, Lehigh University. The receiving station was mounted indoors, just inside of a third floor window.

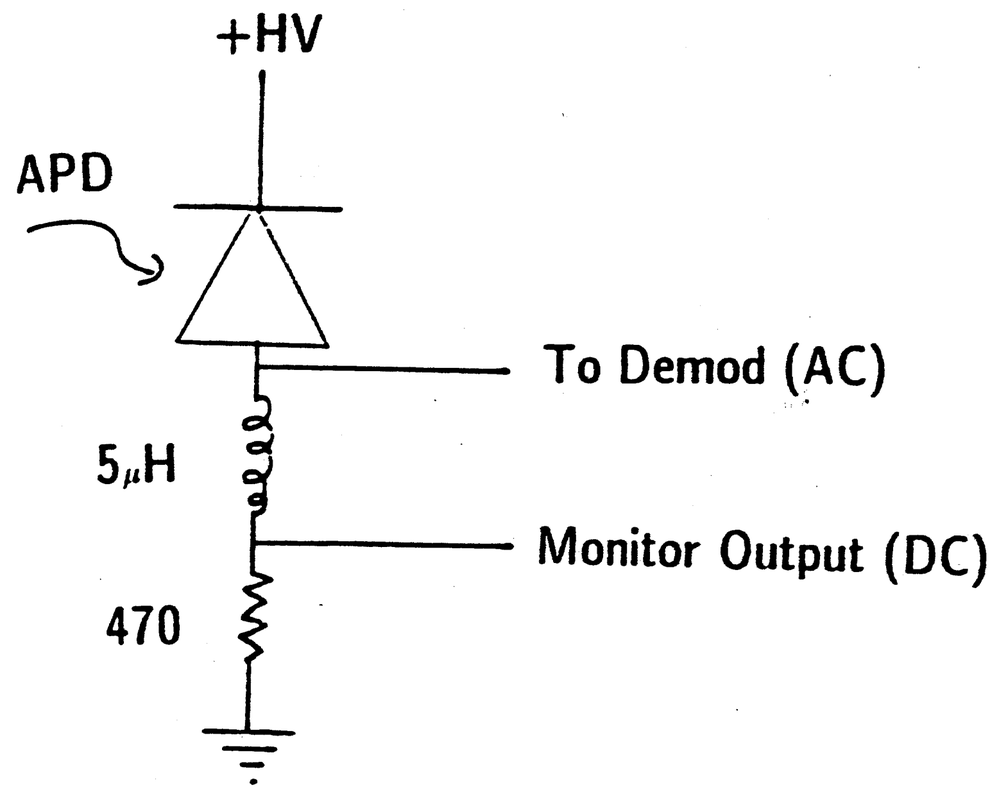
The second link used an industrial unit designed by Laser Communications, Inc. The unit is designed for use over a 1km distance. The transmitter uses a semiconductor laser of the same wavelength and power as the laser in the 3km link, but the transmitter has a 3mR beamwidth. The transmitter was located in the attic of the Moravian College Music Building. The receiving station, 1km distant, was located in Packard lab next to the 3km link receiver.

7.2.2 Receiver Signal Output

In order to monitor the average light level on the detector, a tap was made in each receiver to monitor the DC photocurrent. The locations of the taps are shown in figure 7-1. In each circuit, the DC component of the photocurrent (the average level) passes through a resistor to ground, thus creating a small voltage (1-10mV). This voltage is passed to the logarithmic preamplifiers.

Photodetector Circuit Outputs

3 km Link



1 km Link

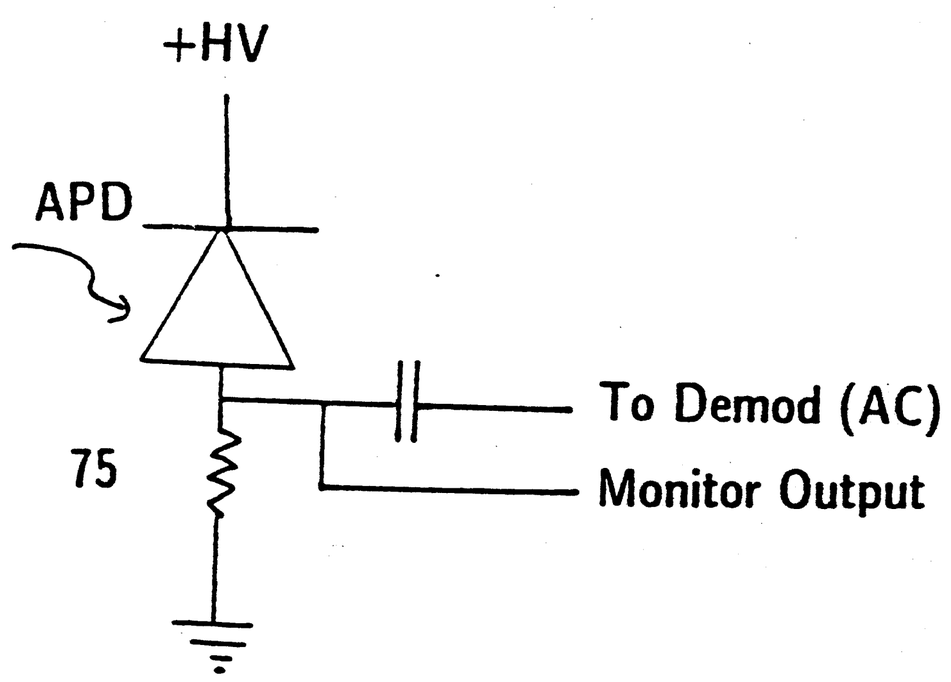


Figure 7-1

7.2.3 Logarithmic Preamplifier

Two preamplifiers were built to interface the receivers to the computer. The logarithmic property of the preamplifier allows a 3 decade range of input levels to remain within the input range of the computer A/D. The preamplifiers perform three functions. The first function is to amplify the small voltage from the receivers up to the 0V to 5V range needed by the computer's A/D converter. Second, the logarithm of the amplified signal is taken. An Intersil ICL8048 log amplifier is used that has a 3 decade range. Offset potentiometers are used to set the zero point and null any offsets. Third, the logarithmic signal is filtered to remove high frequency components that would result in aliasing when used with the computer's A/D converter. The filter is a three pole active Salen-Key low pass filter, with a -3dB frequency of 800 Hz. The filter response is 30 dB down by the Nyquist frequency of 2400 Hz. The output of the filter is passed to the computer A/D input.

7.2.4 Computer and A/D Converter

The computer used to monitor the link is an Intel PC-compatible computer. It is equipped with a 12 bit A/D converter card with 16 channels. Two of the channels are used for monitoring the two links. Software limitations set the A/D sampling rate to 4800 samples per second.

7.2.5 Monitoring Software

Custom software was developed to automate the monitoring of the links. The

developed program consists of 2000 lines of code developed in Turbo Pascal. The program has been named WXMON, short for "Weather Monitor." All data recorded by the program can be saved to disk for further analysis or plotting. The monitor software has the following functions:

- Short Term Signal Sampling: One link's received signal level can be sampled at 4800Hz for up to 2 seconds. The samples can be written to disk in order to study short term fluctuations in the received signal due to atmospheric turbulence.

- Long Term Link Logging: The mean signal level and standard deviation of the signal from the mean can be recorded over time scales of minutes, hours, or days. The time spacing between the log entries is adjustable from seconds to hours. Multiple links can be recorded simultaneously, if needed. Logs of the recorded data are automatically written to disk. and are used to study long term losses (snow, rain, fog) in the atmosphere.

- Spectrum Calculation: An FFT is used on sampled data to find the spectral components of the atmospheric induced signal fluctuations. The number of samples, and thus the spectral resolution, is user specified. The maximum number of points is 2048, which yields a spectral resolution of 2.4Hz. Multiple FFT's can be averaged to obtain a mean spectrum.

- Signal Histogram: In order to calculate the distribution of the signal about the mean, a histogram of sampled values can be taken over many minutes. The number of points to use in the histogram is user specified.

• Dip Statistics: Dip times, that is, times spent below a certain threshold, can be recorded. A histogram of these dip times can be created, both for times below a certain threshold, as well as times between dips below a threshold.

• Self-Calibration: Due to the real-time nature of link monitoring, software timing considerations are crucial. The software is capable of timing the execution speed of it's time-critical segments. Timing measurements are used to optimize the software's performance, as well as give an absolute time reference for the program internals.

7.3 Experimental Link Measurements

Using the test links and the weather monitoring software, several measurements of atmospheric optics effects were taken. Fluctuation distribution, spectra and time variations were recorded, as well as mean signal levels in order to measure extinction. Data from the links is correlated with weather information obtained from the National Weather Service office at A.B.E. Airport, Allentown, which is 6 kilometers distant. Using this information, the measurements are used to confirm the models developed in chapter 6. These models can be used calculate the loss in SNR at the receiver due to extinction of the received beam.

Note that measurements made in decibels represent *optical* power at the receiver. For example, a 1 decade drop in photocurrent due to a one decade (10dB) drop in optical power is recorded as a 10dB drop in optical power, and *not* as a 20dB drop in

photocurrent.

7.3.1 Amplitude Fluctuations

Even under constant weather conditions, the received signal level fluctuates about a mean value due to the motion of turbulent eddies, snowflakes, and raindrops through the atmosphere. The fluctuations in the received beam's intensity occurs on the scale of milliseconds as well as seconds, as collected data shows.

7.3.1.1 Short Term Amplitude Fluctuations

Typical short-term amplitude fluctuations on the two test links under sunny, cloudy, rain, and snow conditions are shown in figures 7-2 and 7-3. These plots show the amplitude fluctuations, sampled every 0.21ms, over a 50ms period. Under sunny conditions, wind and rising eddies of hot air increase atmospheric turbulence, which in turn, cause signal fluctuations at the receiver (see Chapter 6). Therefore, it is not surprising to see that under cloudy conditions, when there is less wind an less ground warming, that there is not as much variation in the signal. As can be seen from the plots, these variations can be as much as ±3dB under sunny conditions, and ±1.5dB under clouds, snow and rain.

With rain or snow in the atmosphere, signal fluctuations are caused not only by turbulence, but also by snow or rain falling through the atmosphere. Raindrops move faster than wind or turbulent eddies of rising warm air. Therefore, raindrops passing through the beam create fluctuations that vary faster than those due to

Signal Amplitude Fluctuations on 3 Kilometer Link

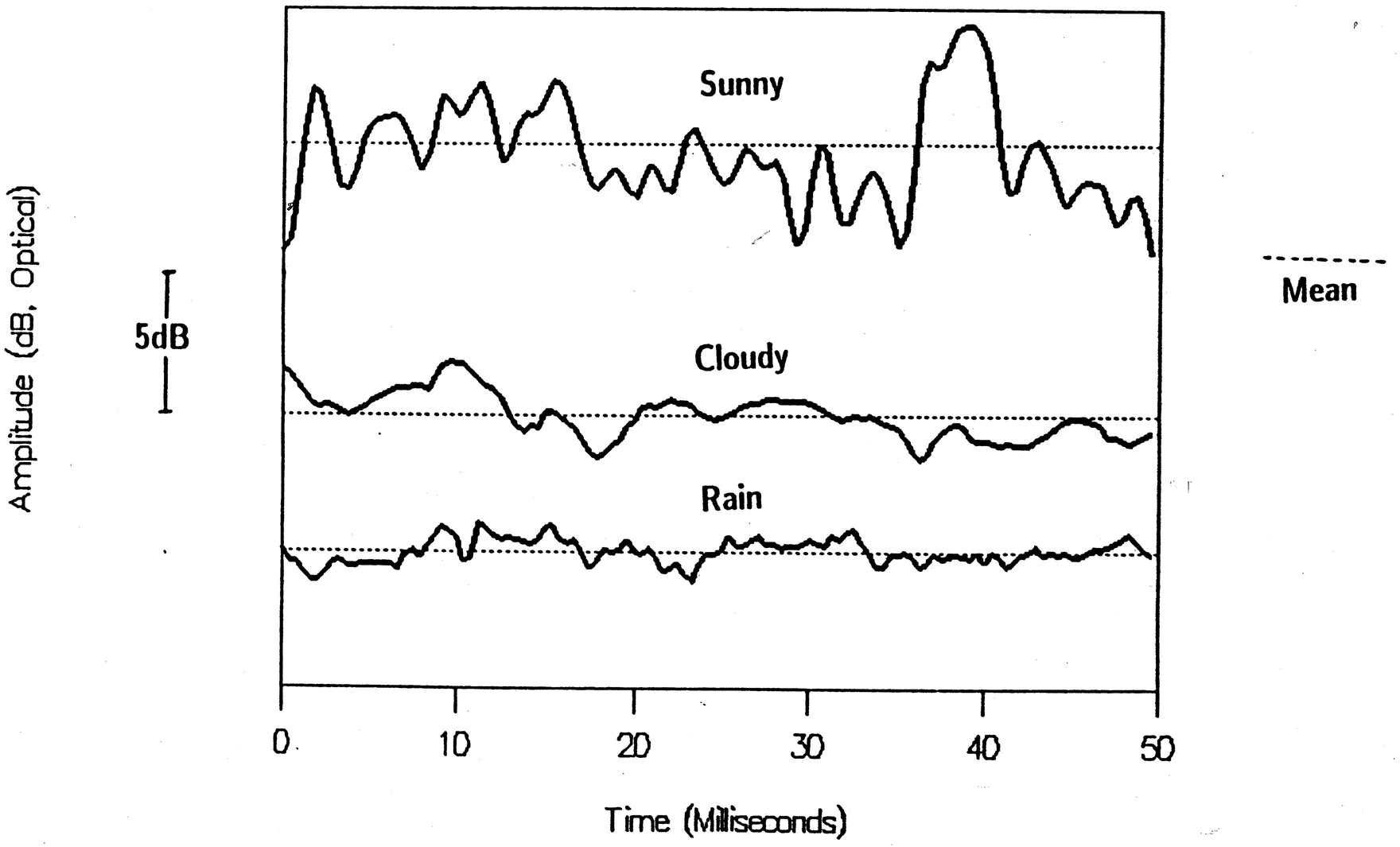


Figure 7-2

Signal Amplitude Fluctuations on 1 Kilometer Link

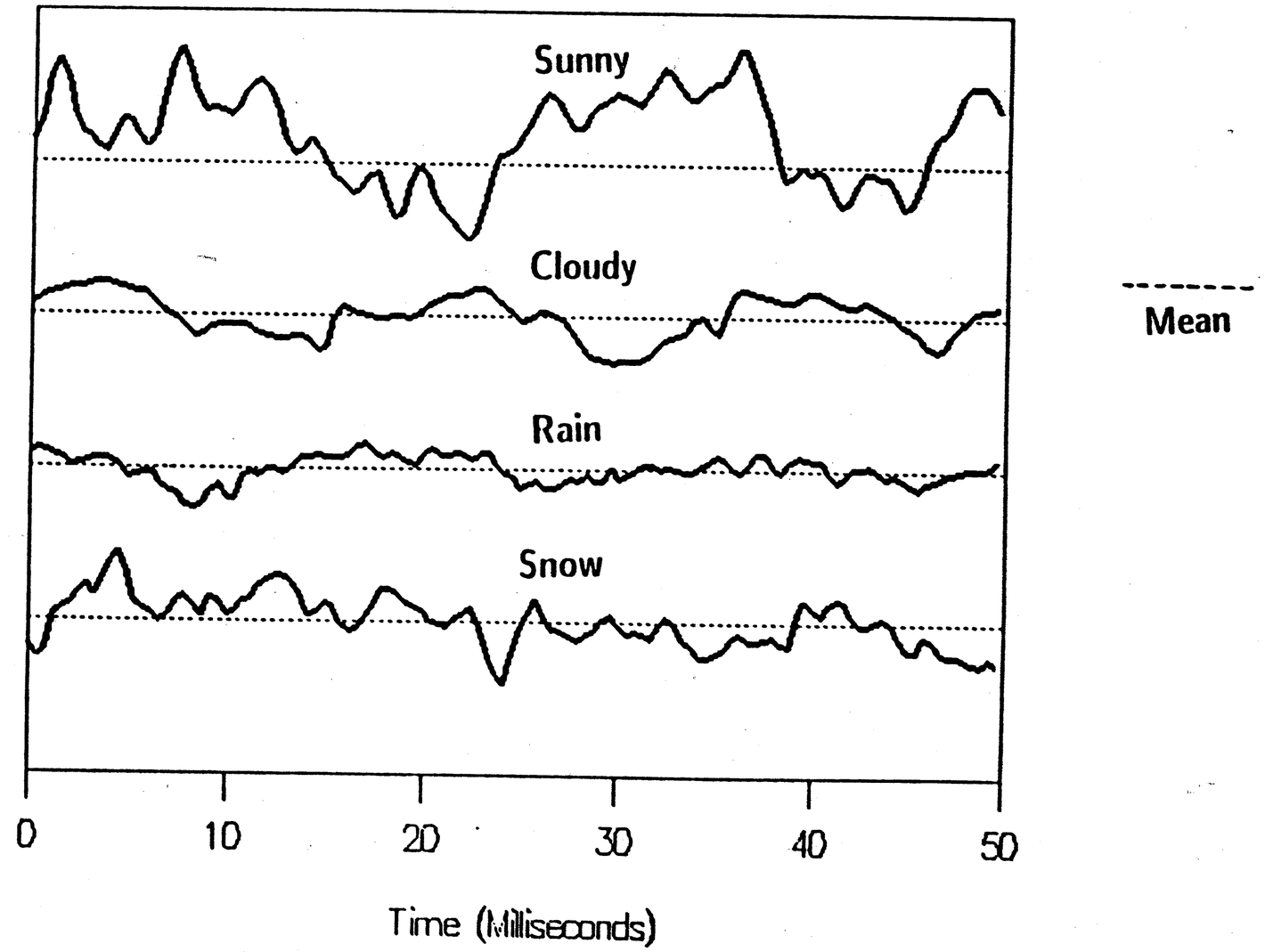


Figure 7-3

turbulence alone. Since rain is typically accompanied by clouds, there is less ground warming, and thus less atmospheric turbulence to cause fluctuations in the received signal. This makes fluctuations due to raindrops more pronounced. On long paths, the contribution to the variance of the signal due to raindrops is typically 10-20% (Borovoy et al). Snowflakes do not fall as fast as raindrops (see chapter 6), so one would expect fluctuations due to snowfall to vary slower than rain induced fluctuations; referring to figure 7-3, on can see that this is the case.

7.3.1.2 Long Term Amplitude Fluctuations

In the previous section, fluctuations were considered over a 50ms time frame. Over a longer time frame, the signal mean varies as well. Figure 7-4 shows a 5 minute record of the mean signal amplitude under clear conditions, recorded every second for both the 1km and 3km links. From the plot, one can see that even the long term average of the signal can vary by as much as ±1dB over a time scale of seconds. These changes are again due to slow changes in atmospheric turbulence, as well as changes in wind speed and direction. Furthermore, as predicted (see chapter 6), signal variance is proportional to path length. This is apparent in figure 7-4, since the 3km link has fluctuations approximately three times the amplitude of those on the 1km link. Note that the fluctuations are not exactly correlated, since the two received beams pass through two different paths through the atmosphere.

7.3.2 Probability Distribution of Amplitude Fluctuations

The distribution of the amplitude fluctuations about the mean theoretically varies as

Signal Mean Amplitude vs. Time for 1 and 3 Kilometer Links

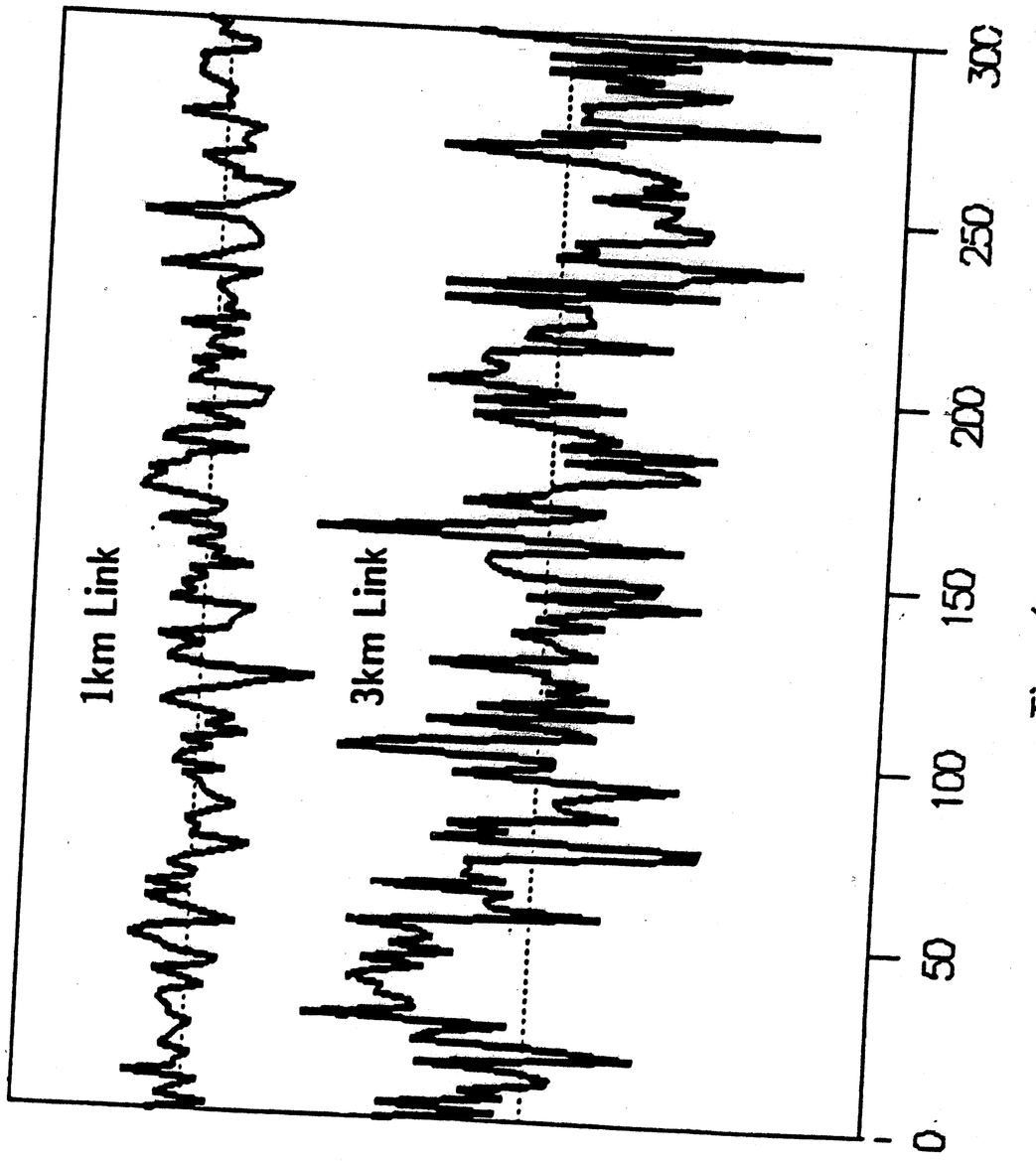


Figure 7-4

a function of turbulence strength, path length, and weather conditions. Measurements on the test links have shown a log-normal distribution, with standard deviations from 0.5dB to 3dB. Experimental data also shows the effects of weather on the fluctuation's amplitude.

7.3.2.1 Theoretical Distribution

To quantify the amplitude fluctuations, statistical tools, such as the probability density function, can be used. Recall from chapter 6 that the probability density function should be log-normal, and that the standard deviation $\sqrt{\sigma^2}$ of the distribution is given by

$$\sqrt{\sigma^2} = \sqrt{\left[\ln\frac{A}{A_0}\right]^2} = \sqrt{0.31\ C_n^{\,2}\ (2\pi/\lambda)^{7/6}\ L^{11/6}}$$

From this equation, one can see that log-amplitude fluctuations due to turbulence increase almost linearly with distance and the turbulence level, if the fluctuations have not saturated.

7.3.2.2 Calculation of Probability Density Function

In order to confirm these predictions, the probability density function of the signal is approximated by forming a histogram of values sampled form the signal. Here, the received signals were sampled at 4800 samples/sec, and 4×10^6 samples are used for a histogram computation. The histogram interval is 0.04dB, or 25 intervals/dB. The results are the histograms shown in figures 7-5 and 7-6, for the 3 and 1 kilometer

Histogram of Signal Fluctuations on 3 Kilometer Link

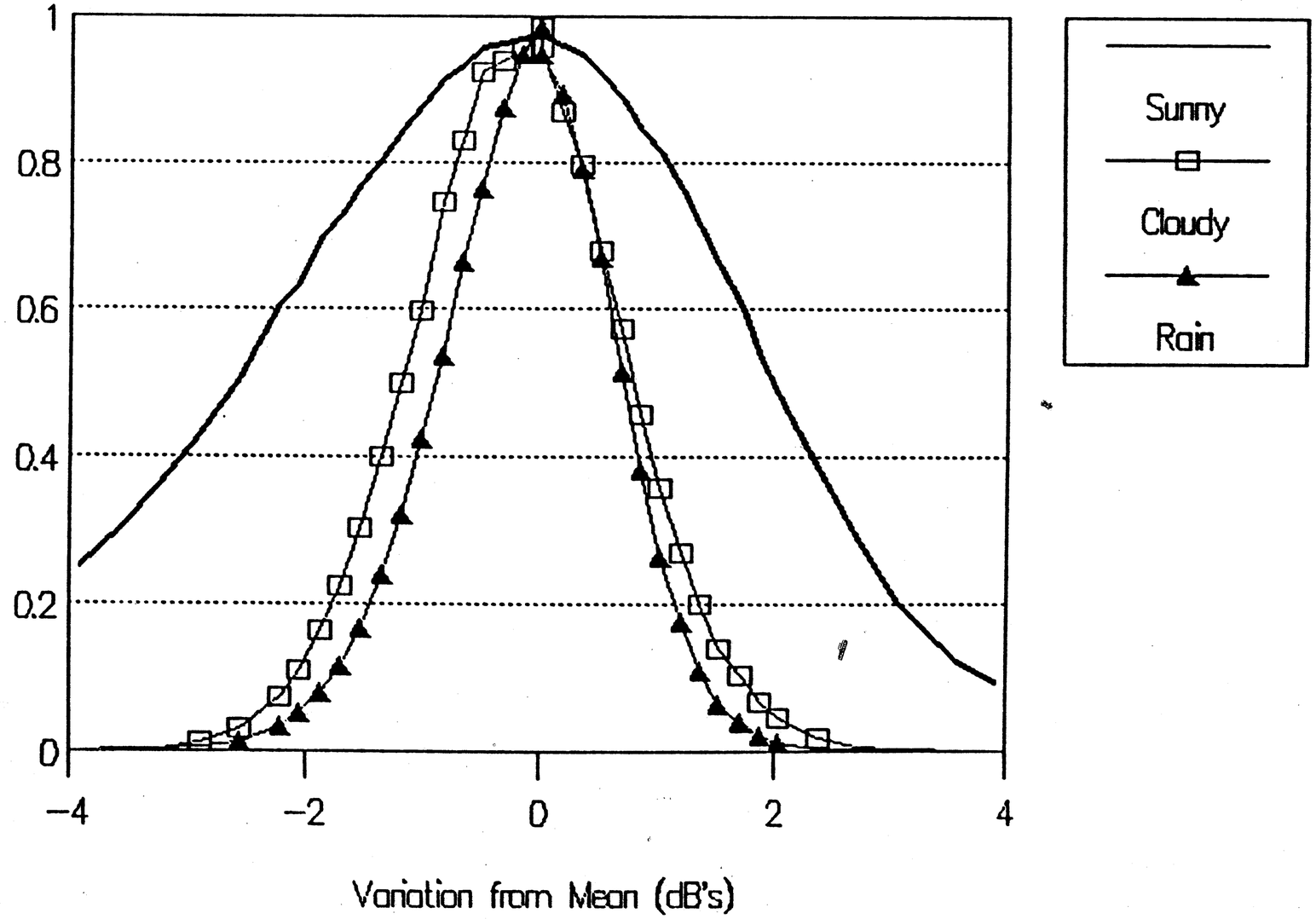


Figure 7-5

Histogram of Signal Fluctuations on 1 Kilometer Link

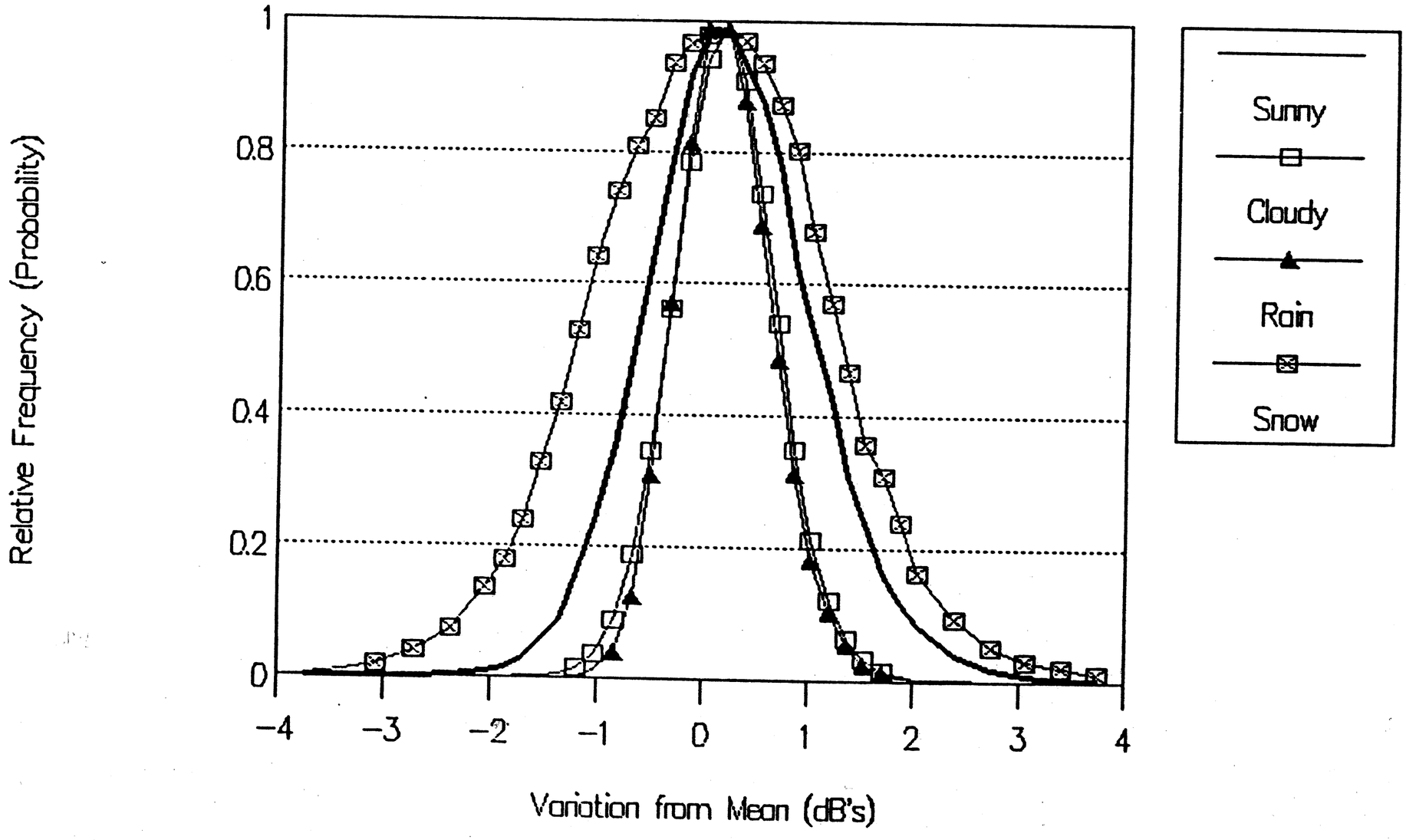


Figure 7-6

links, respectively. The histogram shows that the distribution is indeed log-normal under all weather conditions. Using this fact, the standard deviation of the amplitude fluctuations can be used as a measure of atmospheric turbulence under all weather conditions.

7.3.2.3 Standard Deviation vs. Weather

The standard deviation of the probability distribution varies according to the weather conditions, as can be seen in figures 7-5 and 7-6. Under sunny conditions, ground warming results in increased atmospheric turbulence (reflected in the C_n constant). This increases the amplitude of the signal fluctuations and the standard deviation of the histogram. Under cloudy conditions, turbulence levels are decreased, and the histogram is correspondingly narrower. Furthermore, since rain is often accompanied by clouds, and since raindrops do not cause more than 10-20% of the signal variance, the distributions with clouds and with rain are almost identical.

7.3.2.4 Standard Deviation vs. Time of Day

Even on clear days, the standard deviation is dependant on the turbulence level, which, in turn, is highly dependant on the presence of sunlight. At night, when no solar ground warming effects are present, there are decreased levels of turbulence. During daylight hours, the presence of sun and wind increases atmospheric turbulence, and thus increases the standard deviation of the signal fluctuations. This can be seen in figure 7-7, which is a plot of the standard deviation of the signal fluctuations over a 24 hour period with clear conditions. During nighttime hours, the

Standard Deviation of Fluctuations vs. Hour of Day

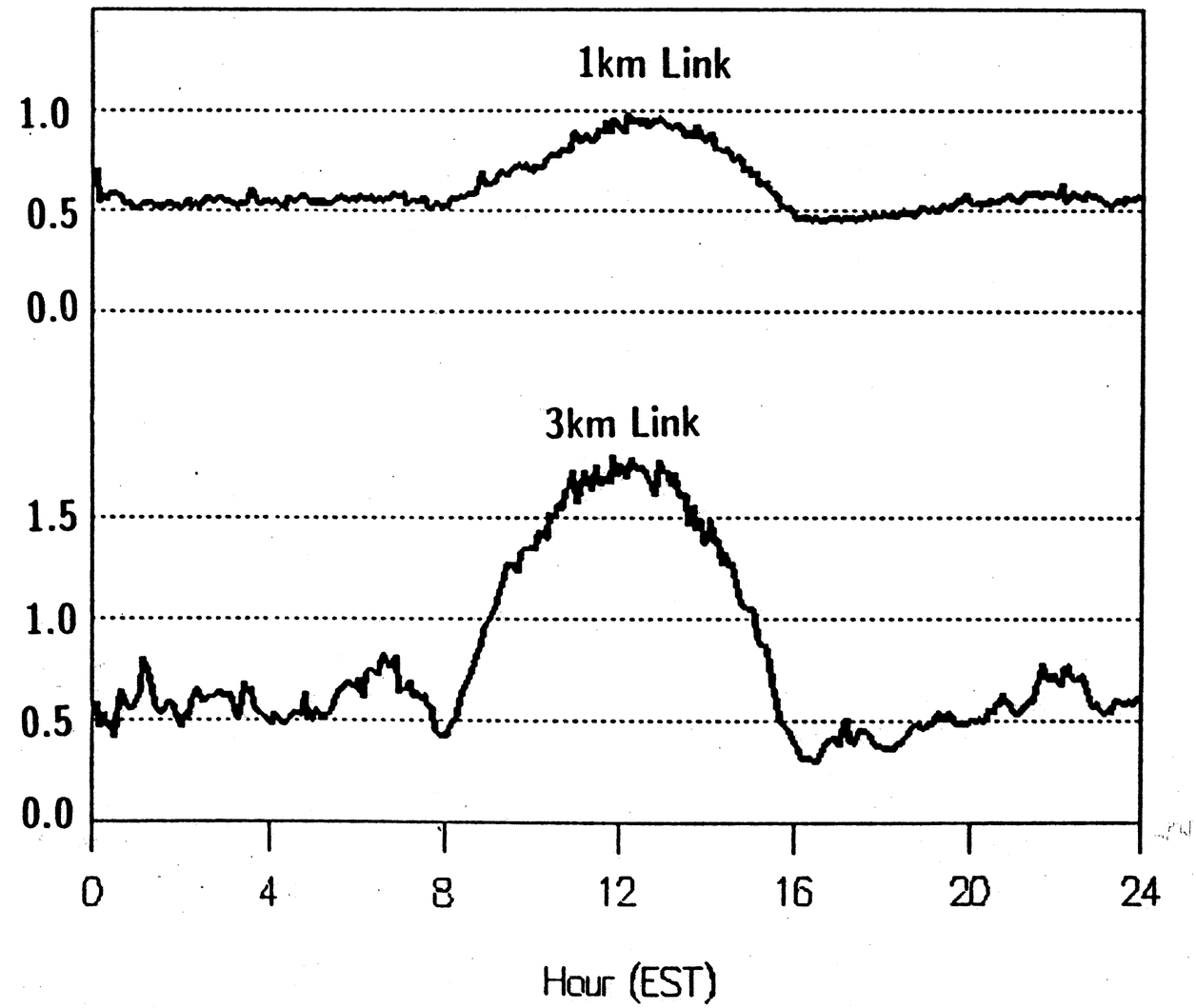


Figure 7-7

turbulence level on both links is 0.5dB. At noon, the standard deviations are 1.7dB on the 3km link, and 0.9dB on the 1km link.

7.3.2.5 Standard Deviation vs. Distance

The standard deviation of the probability distribution also varies according to distance. As can be seen from figures 7-5 and 7-6, the 3km path histograms have larger standard deviations than those on the 1km path under identical weather conditions. Theoretically (see chapter 6), fluctuations due to atmospheric turbulence should increase almost linearly with distance (if the fluctuations are not in the saturation region). Thus, the standard deviation on the 3km link should be triple that on the 1km link. This can be seen in the sunny conditions histograms, where the 3km link histogram has a standard deviation of 3.0dB, almost exactly three times the 1.1dB standard deviation at three times the 1km distance. Of the standard deviations recorded on the two links under all weather conditions, the standard deviations on the 3km link under sunny conditions are typically the largest, with values of approximately 3.0dB$\pm$0.3. This corresponds to the case of intermediate atmospheric turbulence.

Under cloudy and rainy conditions, the standard deviations have not tripled with distance, but have only doubled. This possibly indicates that the fluctuations have already saturated with distance at this lower turbulence level. While saturation phenomenon has yet to be adequately described theoretically, other experimental workers (Borovoy et al) (Ochs & Lawrence) have seen similar results in rain and clear conditions.

7.3.3 Spectrum of Amplitude Fluctuations

Spectral methods can also be used to describe the amplitude fluctuations. Measurements made on the test links confirm the spectrum models under various weather conditions.

7.3.3.1 Theoretical Model of Spectrum

Recall from chapter 6 that the spectrum of the log amplitude fluctuations due to turbulence theoretically increases up to a "breakpoint" frequency f_0, after which the power spectrum decreases as $f^{-8/3}$. The breakpoint frequency was predicted to be approximately 100Hz.

7.3.3.2 Spectra Calculations

To confirm this model, an FFT was used to find the spectrum of the sampled amplitude fluctuations. The FFT used 2048 points that were sampled at 4800 samples/second. The resulting FFT has a spectral resolution of 2.4Hz. Due to the statistical nature of the turbulence, ten spectra were averaged in order to find the mean spectrum over time. A typical spectrum of the amplitude fluctuations on the 3km link is shown in figure 7-8. As one can see from the graph, the signal has a broadband spectrum that decreases with increasing frequency. Most of the signal energy is below 800Hz. The appearance of the spectrum is not unlike that of a 1/f noise source or velocity in a turbulent flow, both of which are random processes with

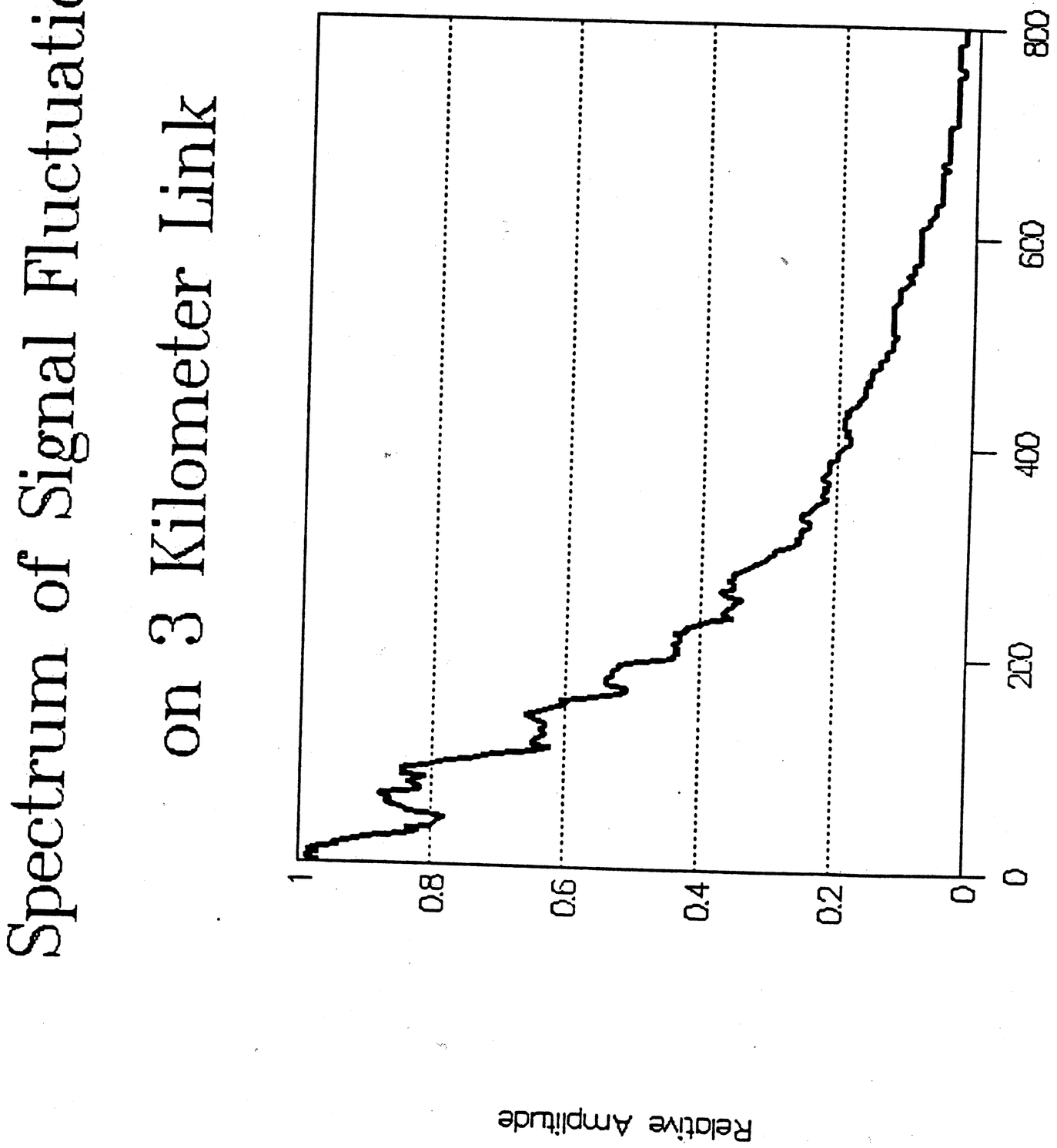

Spectrum of Signal Fluctuations
on 3 Kilometer Link
Relative Amplitude
1
0.8
0.6
0.4
0.2
0
0
200
400
600
800
Frequency (Hz)

Figure 7-8

broadband, decreasing spectra.

7.3.3.3 Spectra due to Turbulence

In order to further investigate the spectra due to turbulence, the results of the FFT were plotted on a log-log (Bode) plot for various weather conditions. These plots are shown in figure 7-9 and 7-10. The corner frequency, f_0, where the amplitude begins to decrease, is easily found from these log-log graphs. For sunny conditions on both links, $f_0 \simeq 200$Hz. Also, since the spectrum should decrease as $1/f^{8/3}$, the slope of the decreasing portion of the plot should be -8/3. A reference line with slope -8/3 has been placed next to the decreasing portion of the spectrum shown in figures 7-9 and 7-10. As one can see, the slopes of the spectra due to turbulence approximately match the theoretical slope. For $f \ll f_0$, the spectrum model predicts an increasing spectral amplitude with increasing frequency (see chapter 6). This, however, is not apparent in the plots. Other experimental work that measured the spectrum down to 2Hz (Subramanian & Collison, 1965) has shown a similar spectrum, without a decreasing amplitude at low frequencies. This may be due to the fact that the lowest observable frequency was only ~2Hz in both cases, which may not have been small enough to satisfy the condition $f \ll f_0$.

7.3.3.4 Spectra due to Rain & Snow

Since raindrops typically move faster than wind or turbulent eddies of rising warm air, rain causes a change in the spectrum. The bandwidth of fluctuations induced by turbulence is 500 to 1000Hz. With rainfall, the bandwidth can increase to 10KHz.

Spectra of Signal Fluctuations on 3 Kilometer Link

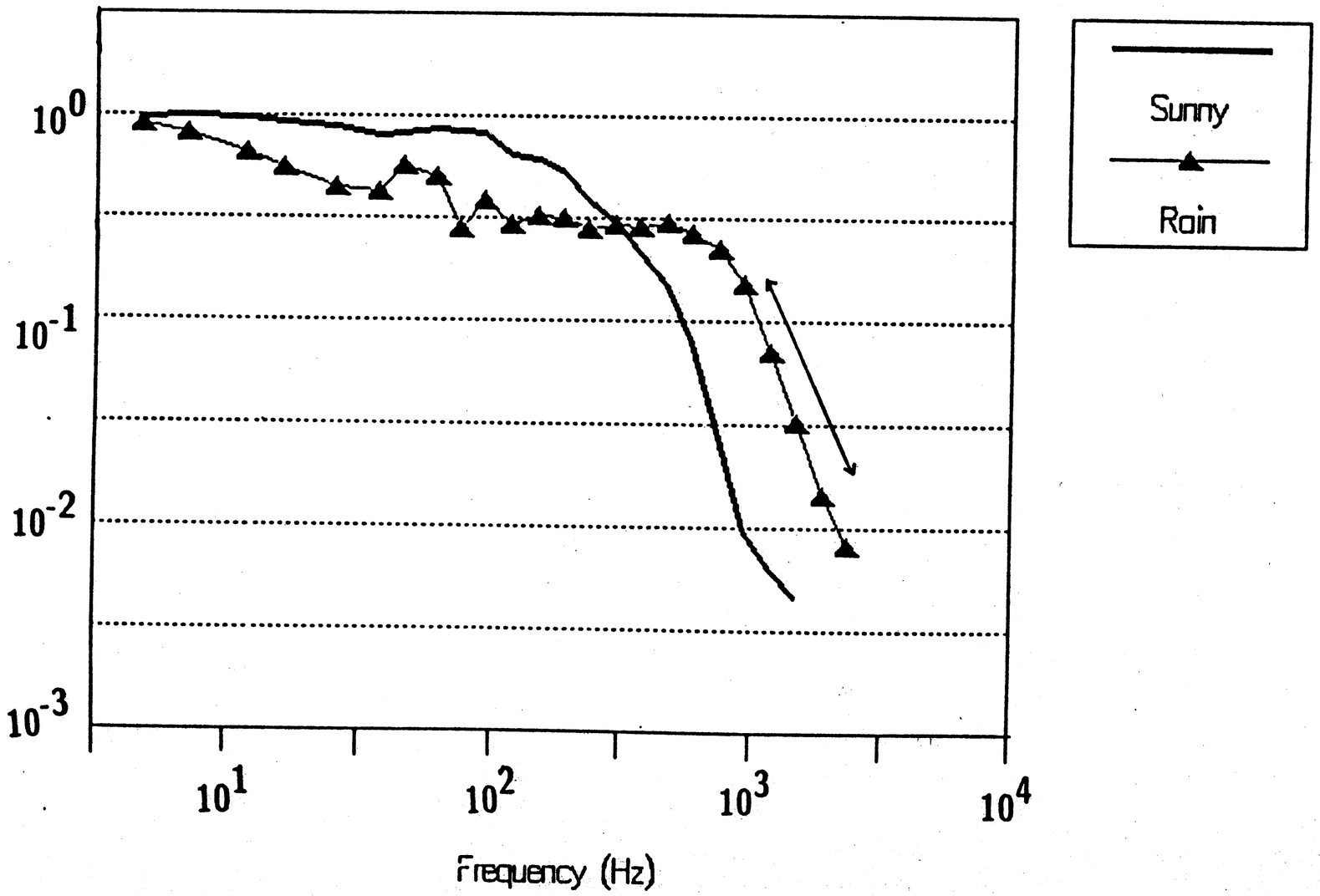


Figure 7-9

Spectra of Signal Fluctuations on 1 Kilometer Link

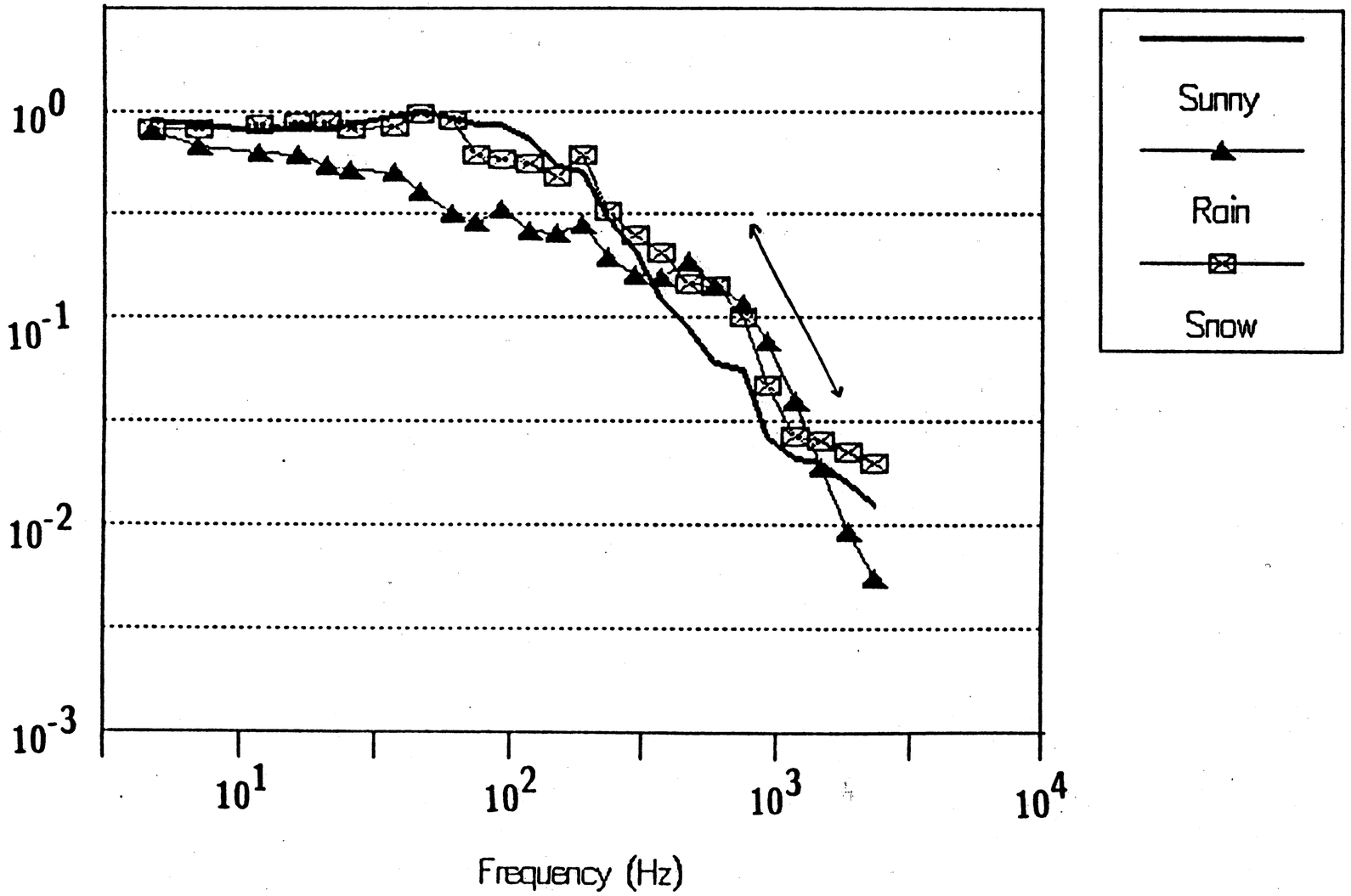


Figure 7-10

The added high frequency components are due to raindrops falling through the beam. The magnitude spectra due to rain has been modeled (Wang et al) as

$$W(f) = \frac{1}{(1 + 6.02\times10^{-8}\, f^2\, R^{0.21}\,)^{9/4}}$$

where R is the rainfall rate in mm/hr, and f is the frequency. A plot of this function is shown in figure 7-11 for R=1, 10, and 100 mm/hr. The modeled spectrum is relatively insensitive to the rainfall rate, R.

Actual spectra obtain during rainfall are shown in figure 7-9 and 7-10. On both the 3km and 1km links, the breakpoint of the rain spectrum is much higher than the breakpoint of the spectrum under sunny conditions. The increase in bandwidth is especially apparent on the 3km link, where the breakpoint frequency shifts from 200Hz under sunny conditions, to 1000Hz during rainfall. In actuality, the breakpoint frequency is most probably higher, but because of the A/D anti-aliasing filter breakpoint frequency at about 1KHz, the spectrum can only be considered accurate up to 1KHz. Despite this restriction, it is easy to see that the rainfall spectrum is much wider than the spectrum due to turbulence alone.

Since snowflakes do not fall as fast as raindrops (see chapter 6), one would expect the spectrum due to snowfall to be narrower than the spectrum due to rain. From figure 7-10, one can see that this is so, and that the snowfall spectrum closely matches the sunny conditions spectrum.

Model of Amplitude Fluctuation Spectrum during Rainfall

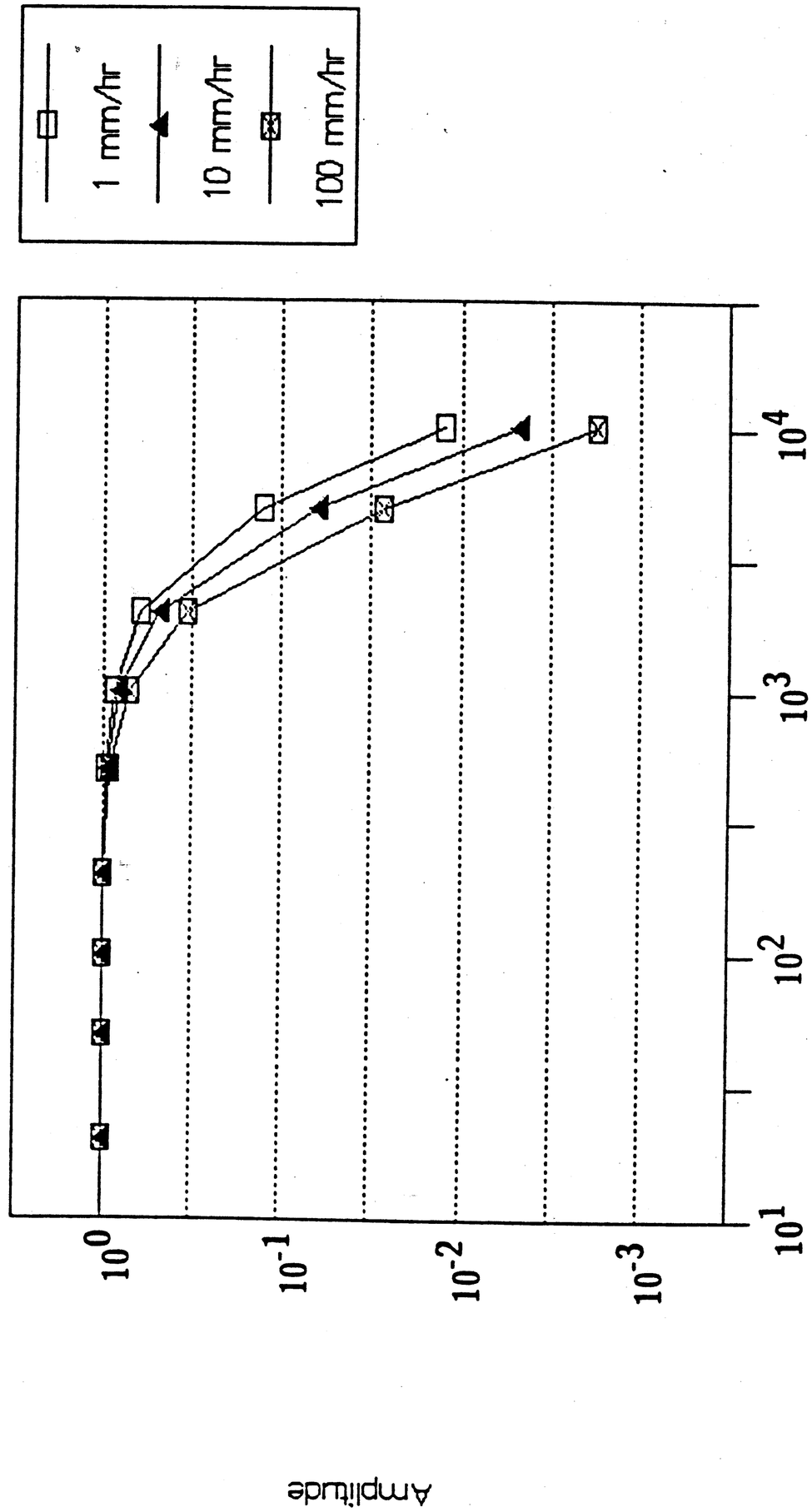


Figure 7-11

7.3.4 Fade Statistics

Due to the turbulence induced fluctuations in the received beam, the received signal level is often below an acceptable level. This undesired fade or dip in the mean signal level causes dropouts in an analog laser communication link, or bursts of bit errors in a digital link. In order to characterize the fades or dips, the duration and distribution of these dips was investigated, both by experiment and simulation.

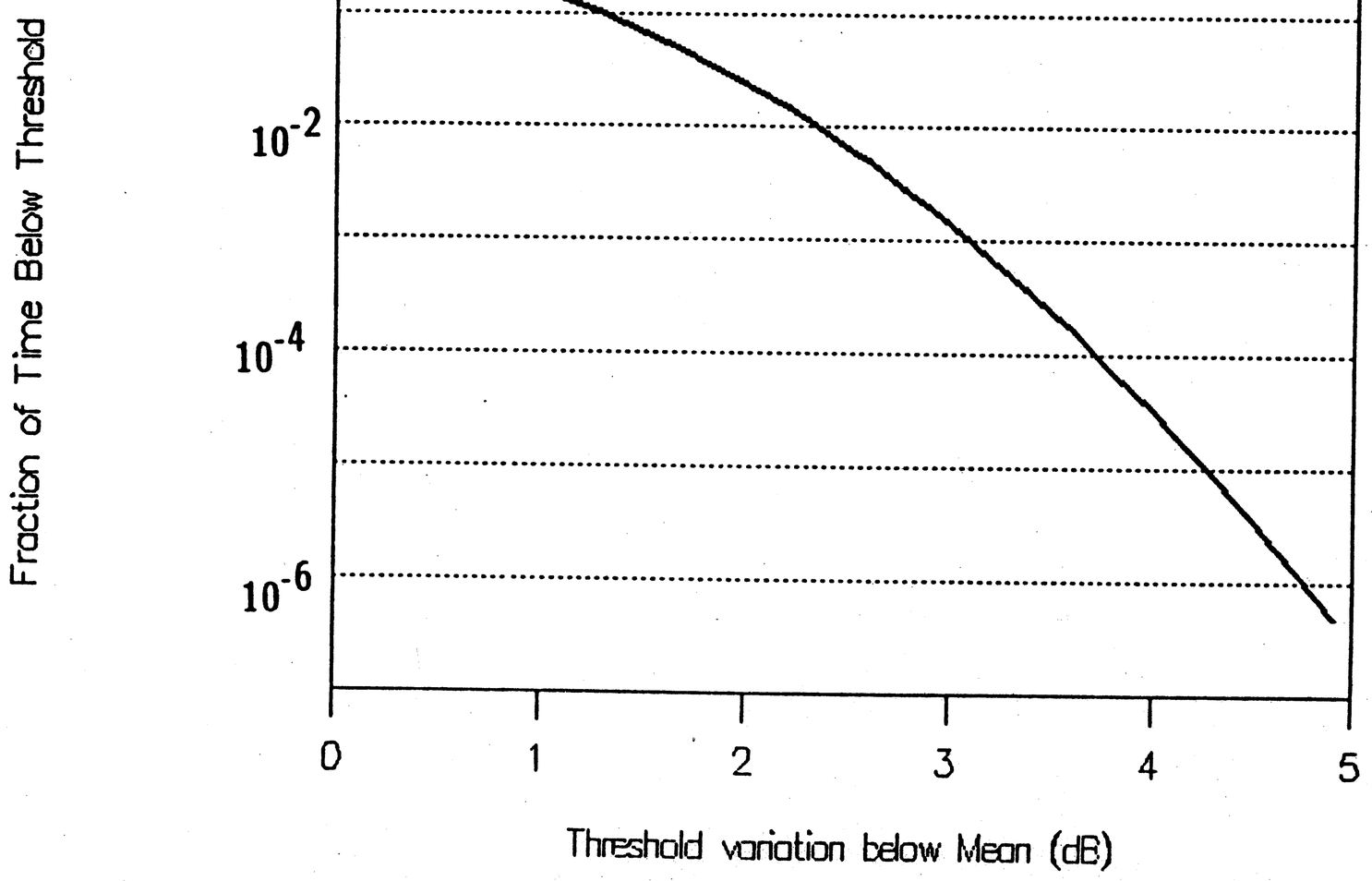
7.3.4.1 Time Fraction of Fades

Since the signal fluctuations are assumed to be a random process, where time averages are equal to ensemble averages, the fraction of time a signal X(t) spends below a certain level L can be found from the cumulative probability function of the signal (Yura & McKinley). Here, the cumulative probability function can be found by integrating over the log-normal distribution. Assume a threshold L decibels below the signal mean. The fraction of time that the signal spends below this threshold is

$$\text{Frac}(\ X(t) < L\ ) = \int_{-\infty}^{L} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-x^2/2\sigma^2}\, dx$$

where σ is the standard deviation of the log-normal distribution, in decibels. This integral cannot be calculated in closed form. Using numerical methods, this function was evaluated and is plotted in figure 7-12, with σ=1.0dB and L ranging from 0dB

Figure 7-12


Time Fraction below Threshold
with Standard Deviation = 1.0
Fraction of Time Below Threshold
10^0
10^{-2}
10^{-4}
10^{-6}
0
1
2
3
4
5
Threshold variation below Mean (dB)

to 5dB.

7.3.4.2 Experimental Fade Duration Histograms

To calculate the probability distribution of the fade duration, the signal level on the 1km link was first sampled. The duration of dips below thresholds of 0, 1, and 2 standard deviations below the mean were the obtained from the samples. Histograms of the durations below the thresholds were calculated. The results are presented in figure 7-13. From the histograms, it is apparent that the most probable dip time is approximately 1ms, and that the probability distribution is relatively insensitive to the threshold.

7.3.4.3 Simulated Fade Duration Histogram

To further investigate the dip duration statistics, the signal fluctuations were simulated. First, the magnitude spectrum of the atmospheric turbulence fluctuations was modeled as

$$W(f) = \frac{1}{1+(f/f_c)^{8/3}}$$

Using f_c=50Hz, the magnitude spectrum is plotted in figure 7-14. The phases of the phase spectrum were set at random in order to model the random motion of the turbulent eddies in the atmosphere. Using these models, 1024 points on the spectrum were calculated. An inverse FFT was taken in order to obtain a simulated

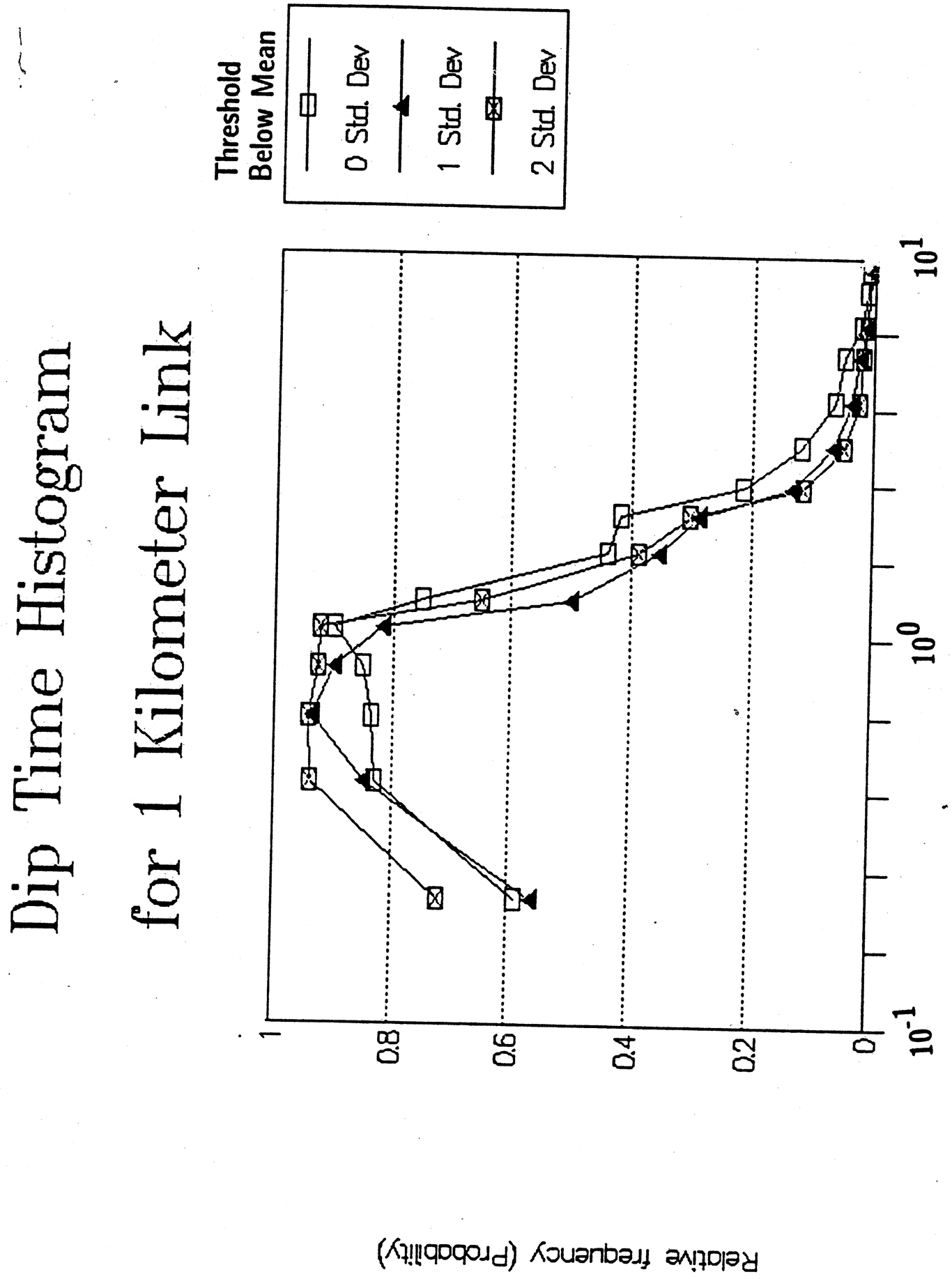

Dip Time Histogram
for 1 Kilometer Link
Threshold
Below Mean
0 Std. Dev
1 Std. Dev
2 Std. Dev
Time (milliseconds)
10^1
10^0
10^{-1}
1
0.8
0.6
0.4
0.2
0
Relative frequency (Probability)

Figure 7-13

Simulated Signal Magnitude Spectrum

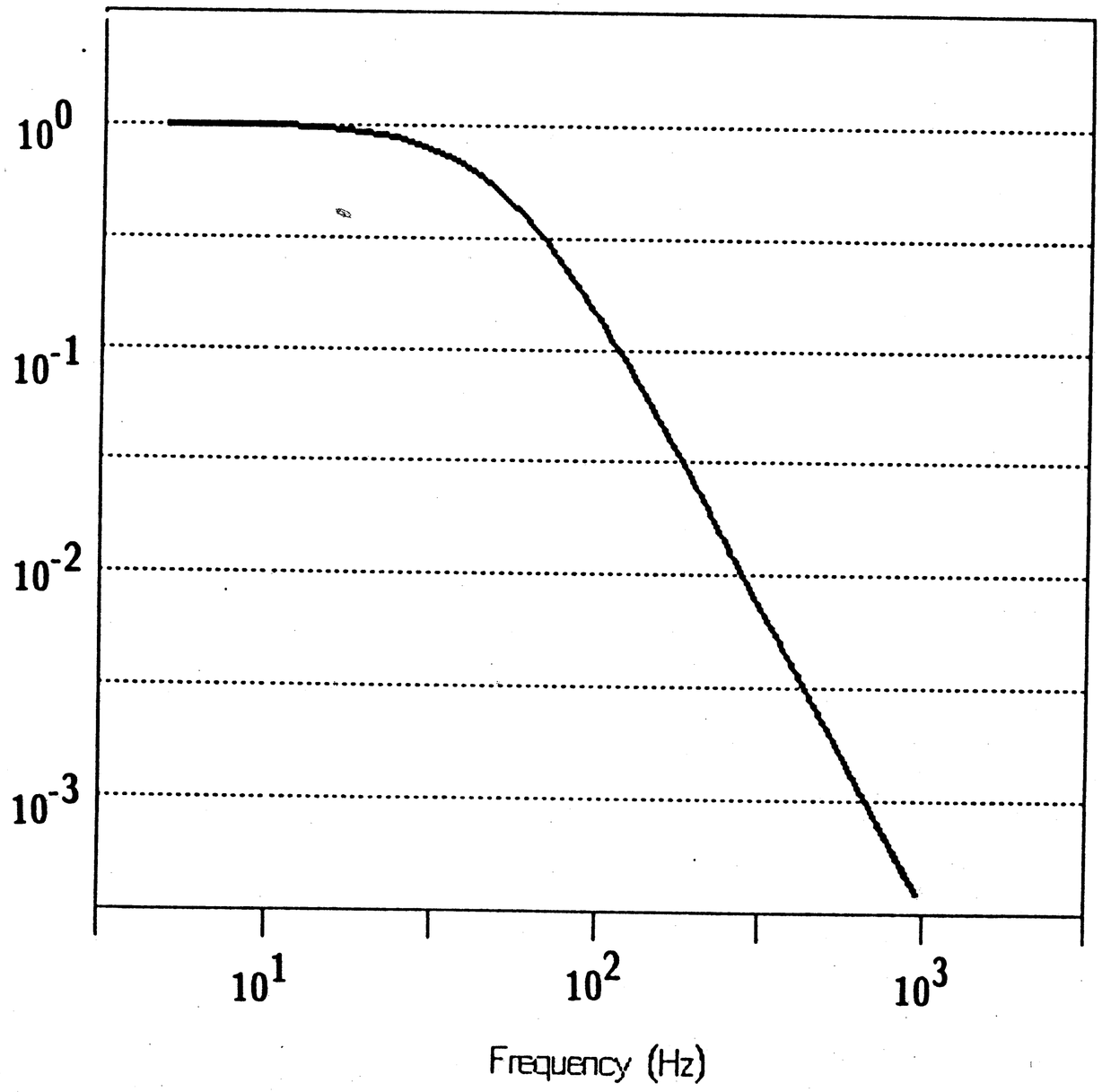


Figure 7-14

time waveform consisting of 2048 points. An example waveform is shown in figure 7-15, over a span of 100ms. In order to verify the accuracy of the probability distribution of the simulated signal, a histogram, shown in figure 7-16, was calculated. The probability distribution of the simulated signal has a standard deviation of 1.2dB, and closely matches the distribution of a real signal.

The dip duration histograms shown in figure 7-17 were calculated using multiple time waveforms. Although the histograms obtained via simulation are similar to the histograms calculated from actual data, there are some differences between the two. First, the most probable dip time varies between 2ms and 5ms in the simulation results. In the histograms created from experimental data, the most probable dip time was approximately 1ms. This discrepancy arises due to the choice of f_c for the simulation. The value of f_c was kept small to decrease the time and space requirements for the simulation; a more realistic value is f_c=200Hz. Second, the simulated histograms show a slight threshold dependance, but the experimental histograms do not. This discrepancy may be caused by the fact that the WXMON monitoring software does not dynamically readjust the threshold level with shifts in the signal mean or standard deviation. Since the dip duration histograms were collected over several hours, changing weather conditions may have caused the standard deviation or mean level of the signal to shift, and distort the resulting histograms.

7.3.5 Extinction Measurements

Extinction caused by rain, snow, and fog can render laser communication impossible.

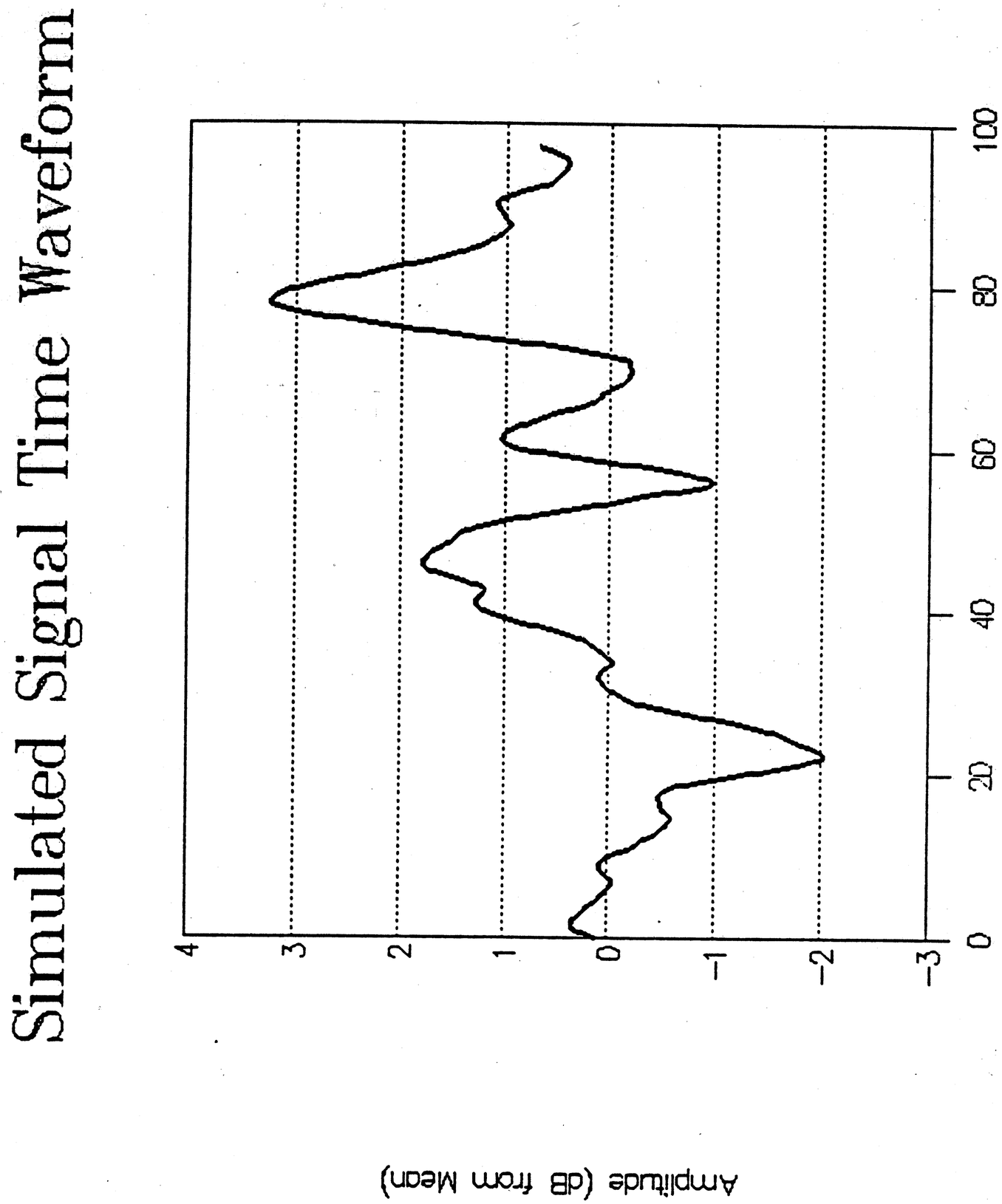

Simulated Signal Time Waveform
Amplitude (dB from Mean)
4
3
2
1
0
−1
−2
−3
0
20
40
60
80
100
Time (mS)

Figure 7-15

Histogram of Simulated Signal

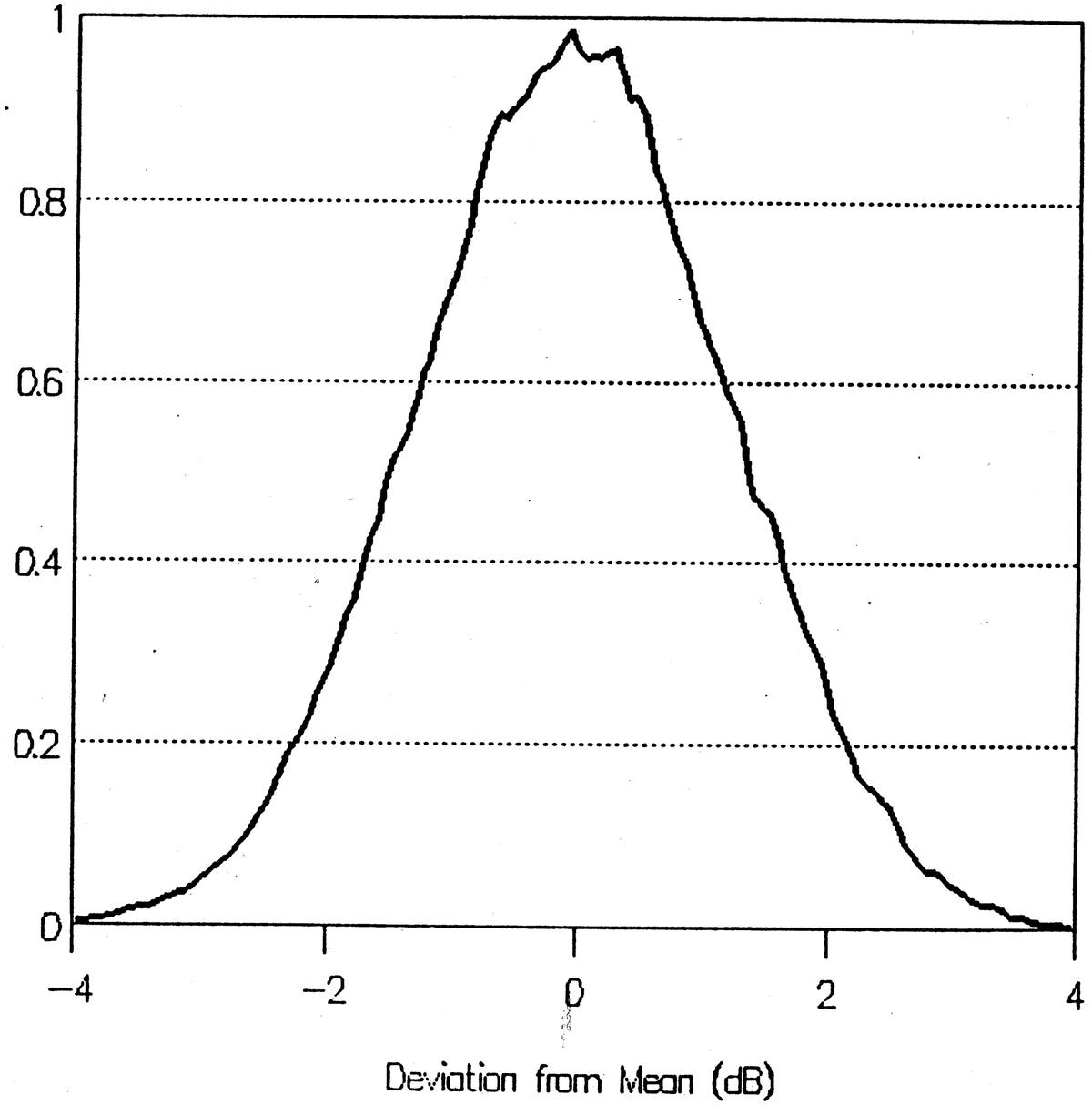


Figure 7-16

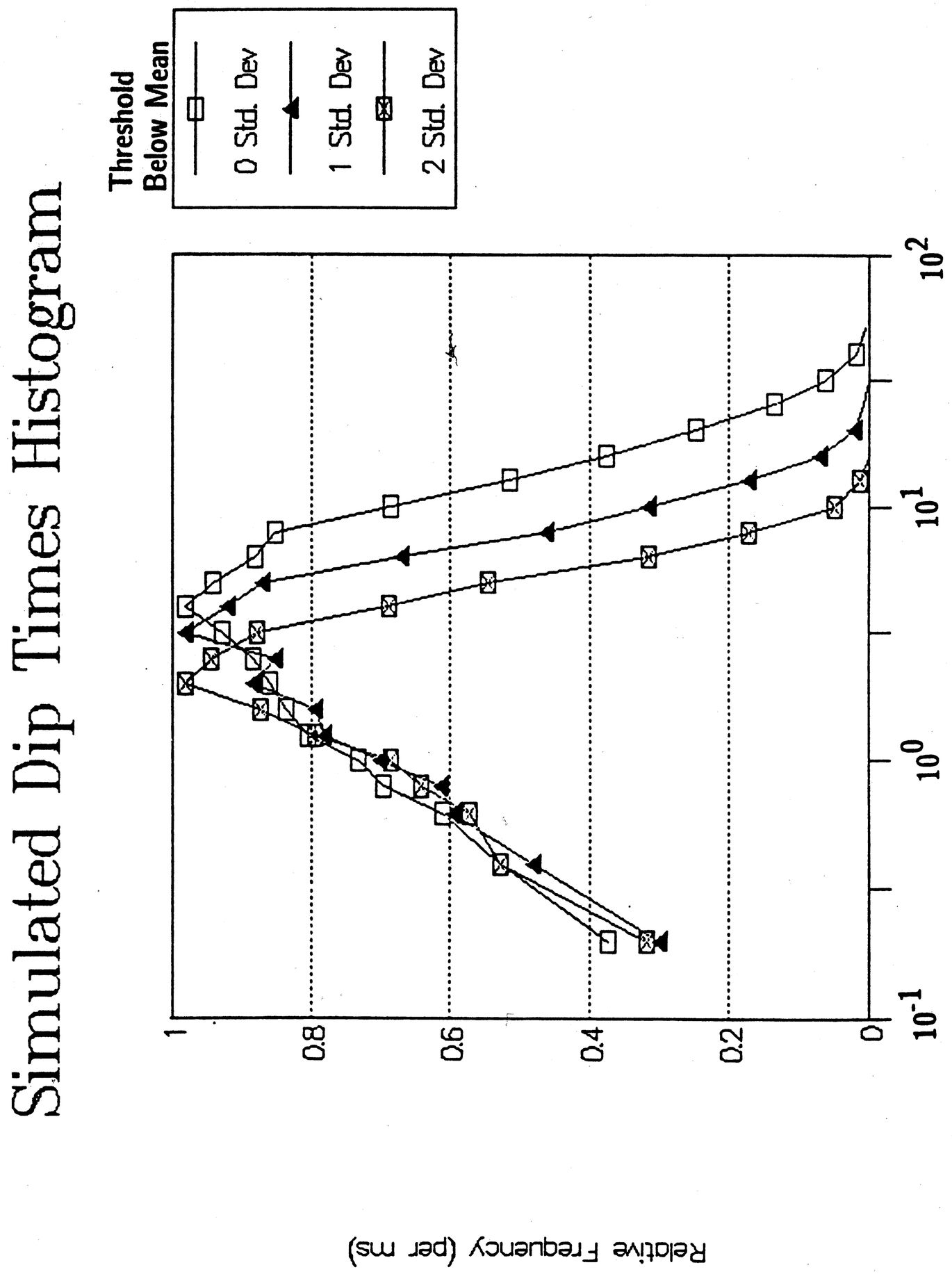

Simulated Dip Times Histogram
Threshold Below Mean
0 Std. Dev
1 Std. Dev
2 Std. Dev
Relative Frequency (per ms)
1
0.8
0.6
0.4
0.2
0
10^2
10^1
10^0
10^{-1}
Time (milliseconds)

Figure 7-17

Data was collected and analyzed in order to confirm the loss models for rain, snow, and fog that were presented in chapter 6. Measurements show that fog creates the highest losses, followed by snow, then rain.

7.3.5.1 Extinction due to Rain

In order to confirm the rain loss model presented in chapter 6, link loss data collected during rainstorms was correlated with rainfall data obtained from the National Weather Service. Using National Weather Service data introduces some uncertainty in the measurements since the rainfall rates are measured 6 km distant. Furthermore, the rainfall data has a time resolution of only 1 hour. Also, rainfalls are often accompanied with some amount of fog; this makes measurements of loss due to rain alone difficult. In order to overcome these difficulties, short term rainfalls (< 1 hr.) and rainfalls with significant amounts of fog were not used in the following analysis.

Fourteen hours of rainfall data was collected and analyzed. The result is shown in figure 7-18, which is a graph of measured path loss versus rainfall rate. Typical values for the rainfall rate are 0.25mm/hr (drizzle), 1mm/hr (light rain), 4mm/hr (moderate rain), and 16mm/hr (heavy rain). As one can see from the graph, most measured losses are under 2dB/km. Included in the graph are the losses predicted by the model presented in chapter 6,

$$\alpha_{dB} = 0.5862\ R^{0.727}\ \text{(dB/km Loss)}$$

Measured Losses due to Rain

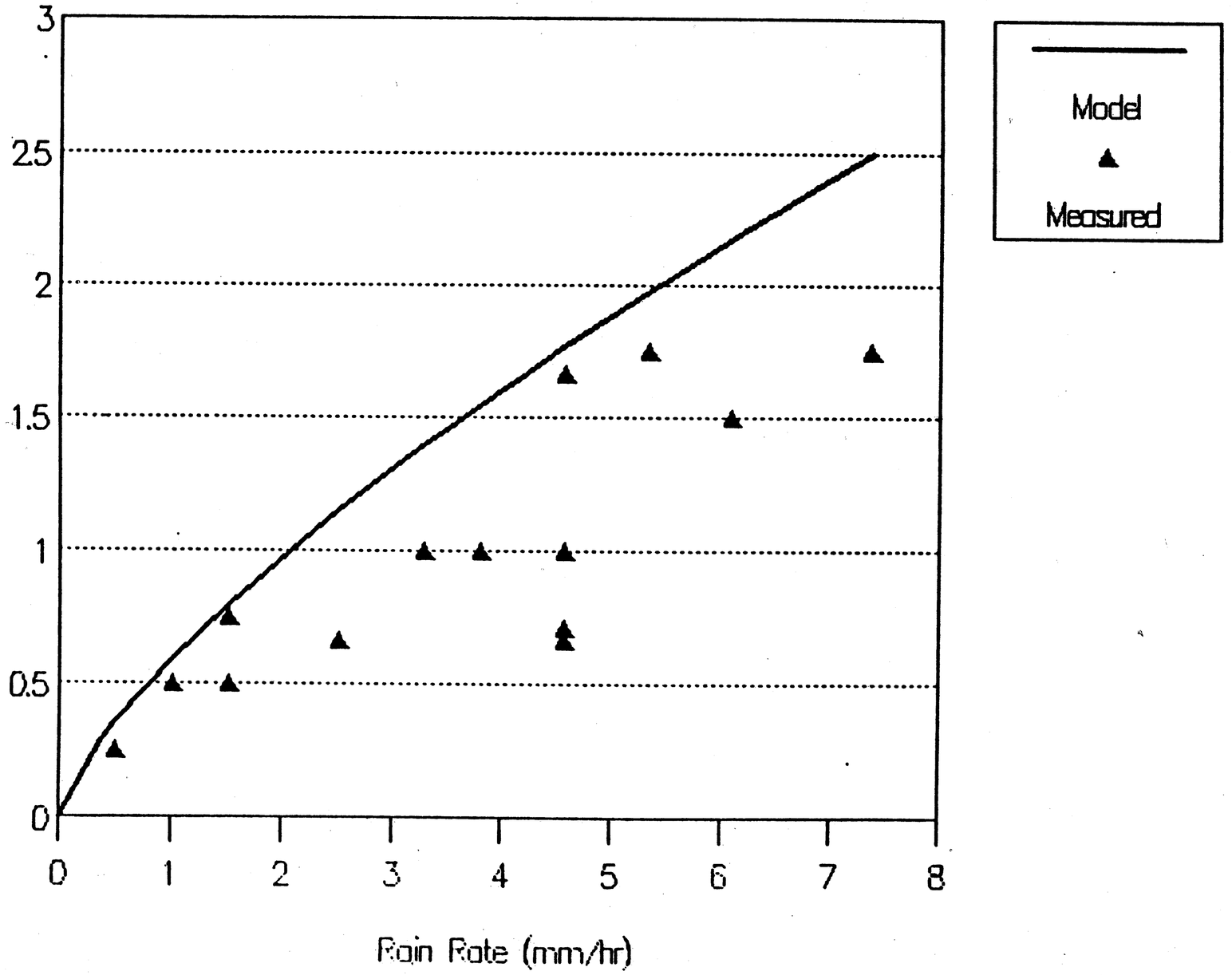


Figure 7-18

The model overestimates the losses somewhat, and a better model to fit this data set would be

$$\alpha_{dB} = 0.39\ R^{0.727}$$

However, in light of the uncertainties noted above, the model seems to approximately match the data.

The above measurements do not include rainfall rates of above 8mm/hr, since most rainfalls can not sustain such a rainfall rate for an entire hour. During a typical summer rainburst, however, the rainfall rate can be much higher for several minutes. The effect of a heavy summer rainfall is shown in figure 7-19, which is a graph of link losses measured every five minutes. As can be seen, the losses on the two links can reach as high as 4dB/km in heavy rain. The peak rainfall rate was estimated at 20mm/hr, which yields a predicted loss of 5dB/km.

7.3.5.2 Extinction due to Snow

Snowfall can introduce large losses in a communication link and render laser communication practically impossible. Snowfall has been observed to cause losses of over 20dB on both the 1km and the 3km link, essentially eliminating the beam.

Losses on the 1km link during a light snowstorm are plotted in figure 7-20, along with dry snowfall rates recorded by the National Weather Service. This data has a time resolution of 1 hour, and an accuracy of ±0.75mm. Using data similar to this gathered during ten snowfalls, figure 7-21 was constructed. This plot shows

Signal Levels during Rainstorm

7/27/89

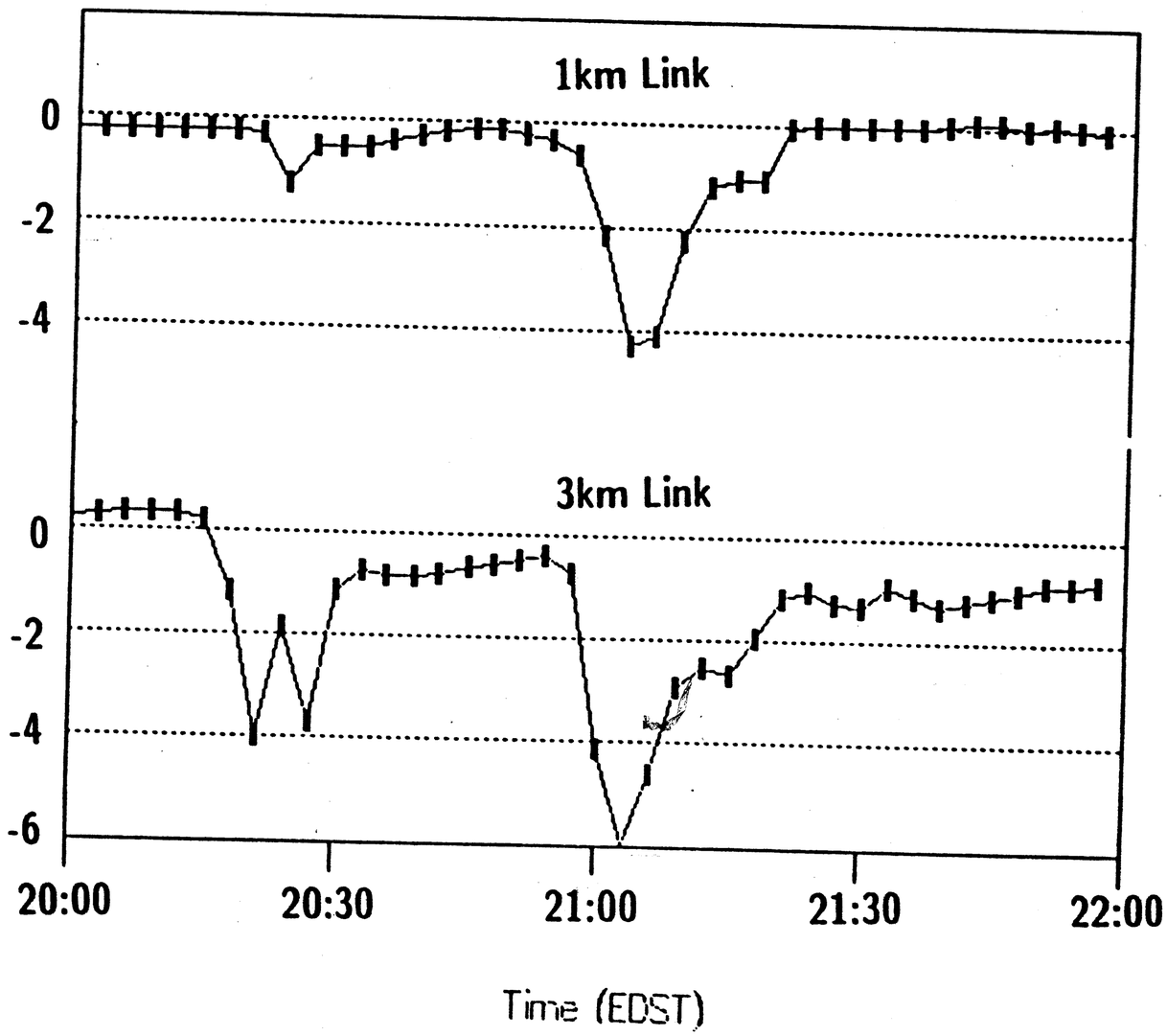


Figure 7-19

Signal Level on 1 Kilometer Link during 12/5/89 Snowstorm

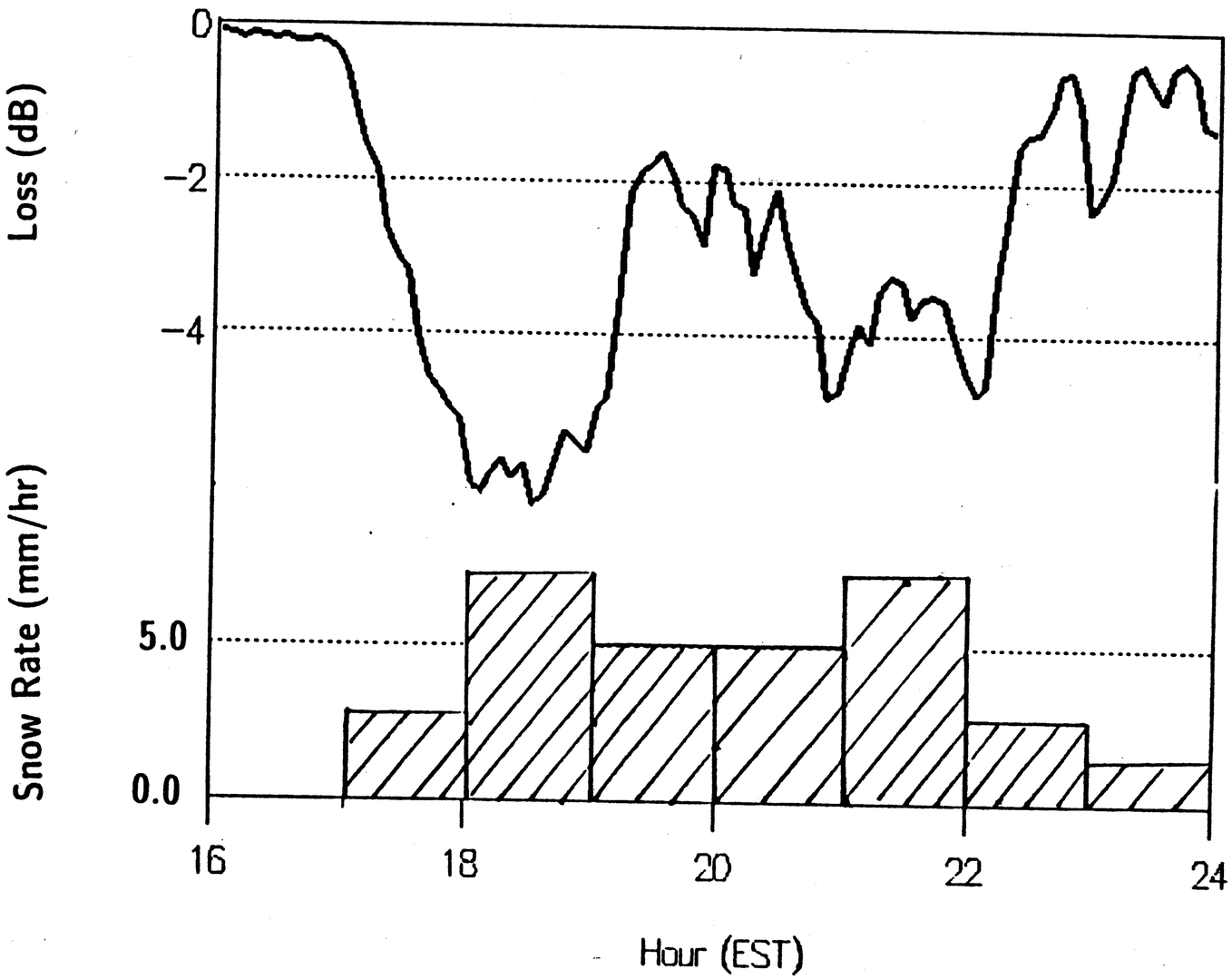


Figure 7-20

Measured Loss Due to Snow

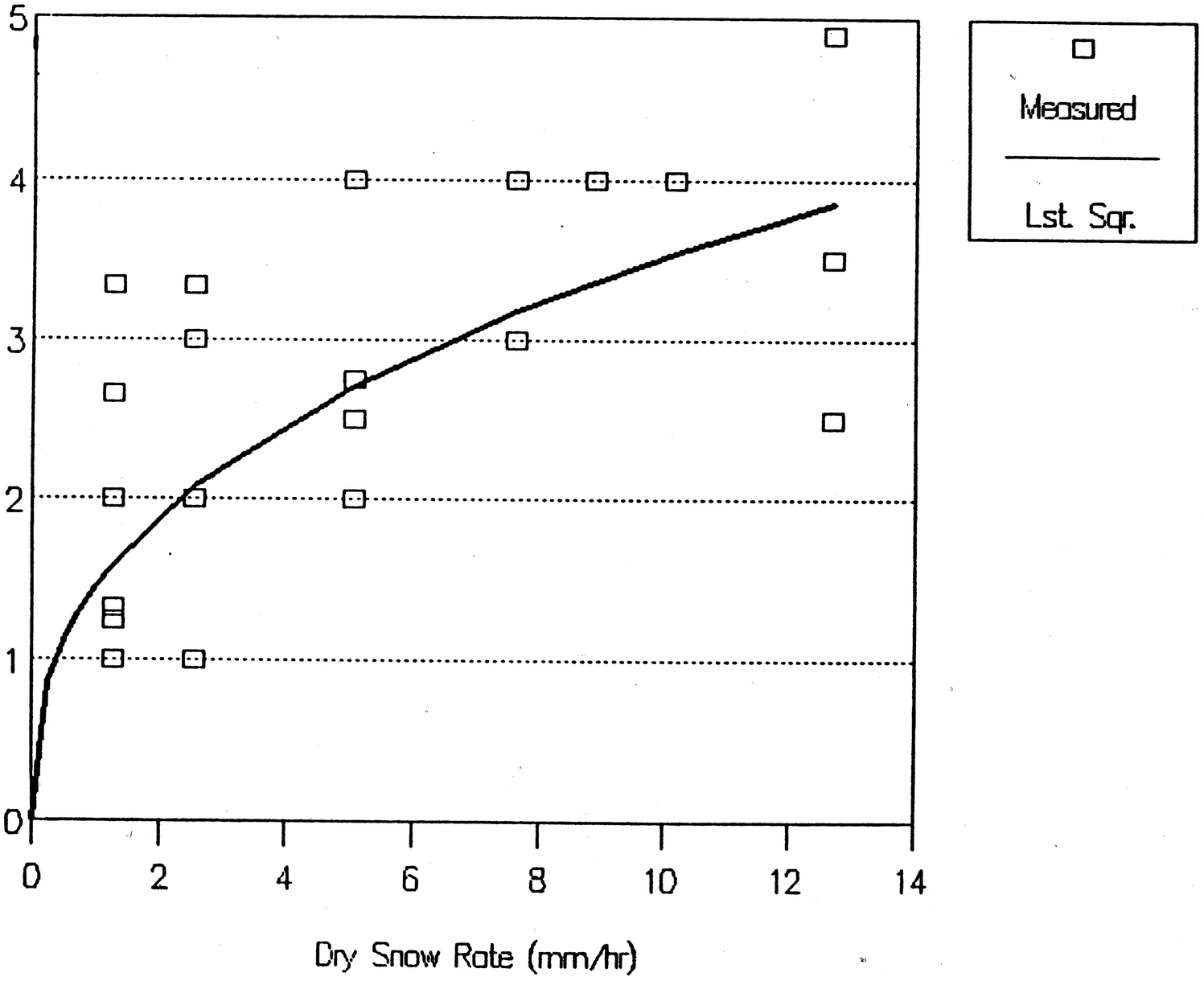


Figure 7-21

measured loss on the links versus dry snowfall rate. Using a nonlinear least squares fit, the best fit to this data is given by the model

$$\alpha_{dB} = 1.453\ R^{0.3844}$$

where α_{dB} is in dB/km, and R is the dry snowfall rate in mm/hr.

A similar model has been found by (Mellor), which is equivalent to

$$\alpha_{dB} = 1.81\ R^{0.421}$$

Empirical models presented in the literature, as shown in chapter 6, are quite varied. This is because the type and density of snow varies by region, and can even vary within the same snowstorm. The density of snow, which is the largest unknown factor in these measurements, can vary from $50kg/m^3$ to $200kg/m^3$. Here, a value of $100kg/m^3$ was assumed.

7.3.5.3 Extinction Due to Fog

The high losses due to fog can also render laser communication practically impossible. Measured losses due to fog have a range of 10dB/km and higher. The difficulty in modeling fog losses arises in finding a physical quantity of the fog to measure. Theoretical results shown in chapter 6 derived loss versus density of water in the air. No conclusive experiments were performed to confirm this model, since water density in air is difficult to estimate in the field.

Usually the easiest measurement to make is that of meteorological range, that is, the distance at which a black object is just noticeable against the sky. Recall from chapter 6 that the meteorological range, R_m (in kilometers) is given by

$$R_m = \frac{3.912}{\alpha}$$

Therefore, converting α to decibels, we can obtain

$$\alpha_{dB} = \frac{16.97}{R_m} \text{ (dB/km)}$$

While hourly visual range information was not taken, it was noted that light fogs can increase path losses by approximately 20dB when the 1km or 3km transmitting station was just visible. Using the above formula for visual ranges of 1km and 3km, the losses were predicted to be 16.97dB, close to the measured value of 20dB. Since the attenuation of a fog is wavelength independent at optical wavelengths (see chapter 6), this attenuation model is valid for the near-IR laser wavelength also.

8. CONCLUSION

8.1 Summary of Work

An atmospheric laser communication system for digital communication has been constructed and tested. The system is capable of asynchronous data rates of up to 20Mb/s over a path of 3km. The system uses a BPSK modulation scheme, since an analysis of modulation schemes has shown that binary phase shift keying provides optimum performance. The transmitter and receiver circuitry to implement BPSK was constructed and analyzed. The front-end of the receiver circuit was found to be a major factor influencing system performance. Optimization of the noise performance of the receiver front-end resulted in the selection of a transimpedance amplifier and APD gain of 100.

Transmitter and receiver optics were kept relatively simple in order to reduce complexity and costs. At the transmitter, beamwidth considerations compromise received power levels to pointing accuracy. The transmitted beamwidth of 1mR is large enough so that vibration and atmospheric turbulence effects do not move the beam off of the receiver. At the receiver, a large objective lens was selected with a focused spot size is smaller than the photodetector. This maximizes the angle of acceptance as well as the power delivered to the photodetector.

Studies of link performance under various weather conditions have shown that both extinction and atmospheric turbulence affect the received signal level. Atmospheric turbulence results in random amplitude fluctuations of the received beam. These

amplitude fluctuations are best modeled using a log-normal probability distribution, and a spectrum that decreases as $f^{-8/3}$. Extinction results from particles in the atmosphere, such as rain, snow and fog. Theoretical analysis of scattering by atmospheric particles has predicted that these losses may be up to many 10's of dB/km. Experimental data has confirmed that typical fog and snow conditions, which cause the most extinction, can make optical communication practically impossible at times.

8.2 Suggestions for Future Work

The performance of an atmospheric laser communication system can be improved considerably at the price of system complexity and cost. Promising methods to investigate include Manchester encoding, error correction codes, and automated beam tracking systems. Furthermore. since weather effects the atmospheric laser communication system, these systems can conceivably be used to make several atmospheric measurements.

8.2.1 Manchester Encoding

The system described here was designed to behave like a 'wire,' that is, it was bit rate independent and asynchronous. To obtain these qualities, the system uses BPSK, and which is essentially an analog modulation technique. If, in designing another laser communication system, the bit rate can be *fixed*, then Manchester encoding at the transmitter can be used. Manchester encoding exclusive-or's the incoming bit stream with a square wave running at twice the incoming bit rate.

This guarantees a transition every bit period. A Manchester encoded bit stream has no DC components, and little spectral energy at low frequencies. Thus, AC coupling can be used at the receiver, simplifying circuit design, and eliminating the effects of background light. At the receiver, a clock recovery circuit is used to recover the transmitter clock and decode the data.

Manchester encoding is very similar to BPSK modulation, since the incoming bit stream is being multiplied by a square-wave 'carrier' at twice the bit rate. In BPSK, the carrier is sinusoidal and not synchronized to the bit source; in Manchester encoding, the carrier is a square wave that is synchronized to the bit source. The advantage of Manchester encoding is the savings in bandwidth; for a 20MB/s source, the transmitted signal requires 40MHz. The BPSK system developed here requires 120MHz. Thus, using a Manchester encoding scheme, the same 120MHz bandwidth system can used to send data at a fixed bit rate of 60Mb/s.

8.2.2 Error Correction Codes

In a digital system with a fixed bit rate, bursts of bit errors due to deep and rapid fades in the received signal can be corrected by using error correction codes. Coding and decoding can be done by dedicated digital hardware, or microprocessors. Using simple codes and interleaving, (Davidson) has shown coding gains can range from 7dB to 11dB. Since these bursts occur in a time scale on the order of milliseconds, at high data rates (>1 Mb/s) dips in the signal cause errors spanning many thousands of data bits. Rather than resorting to the use of very long and complicated error correcting codes to correct for these long duration error bursts,

interleaving of simple codes is a more attractive alternative (Davidson). This requires that several milliseconds of the incoming bit stream be stored in memory in the transmitter.

8.2.3 Beam Tracking System

For communication over long distances, narrower beamwidths must be used in order to minimize optical power losses. Narrow beamwidths make aligning the transmitter and receiver difficult, increases the sensitivity of the system to vibration, and makes beam wandering due to atmospheric turbulence more pronounced. Clearly, a system employing feedback of receiver signal strength to the transmitter could be used to help automatically aim the transmitting beam for best reception at the receiver. Successful experiments by (King et al), which used an automatic aiming system for a 0.1mR beam over a 23 mile path, demonstrated the results possible when using such a system.

8.2.4 Weather Measurements

Since the performance of an atmospheric laser communication system is highly dependant on weather conditions, a slight modification of a communication system could be used to measure weather statistics. For example, using the amplitude fluctuation spectrum, (Wang et al) has demonstrated the use of a laser system to measure rainfall rates. As rainfall increases, the spectrum of the fluctuations broadens, and the width of the spectrum can give an absolute measurement of the rainfall rate. Another potential application is in measurements of visibility.

Visibility could be quantified by measuring backscattered intensity or received intensity over a measured path. Also, wind speeds can be measured by correlating the amplitude fluctuation signals from two receivers several meters apart. Lastly, the strength of the amplitude fluctuations due to turbulence can be used to estimate the atmospheric turbulence level, which is an important consideration at airports.

References

Arnulf, A., and Bricard, J., "Transmission by Haze and Fog in the Spectral Region 0.35 to 10 Microns" Journal of the Optical Society of America, Vol. 47 (1957), pp. 491-498.

Best, Roland E. Phase-Locked Loops : Theory, Design, and Applications New York: McGraw-Hill Book Co., 1984.

Borovoy, A. G., Patrushev, G. Ya., and Petrov, A. I., "Laser beam propagation through the turbulent atmosphere with precipitation" Applied Optics, Vol. 27 (1988), pp. 3704-3714.

Brookner, Eli, "Atmosphere Propagation and Communication Channel Model for Laser Wavelengths" IEEE Transactions on Communication Technology, Vol. Com-18 (1970), pp. 396-416.

Chu, T.S., and Hogg, D.C., "Effects of Precipitation on Propagation at 0.63, 3.5, and 10.6 Microns" Bell System Technical Journal, Volume 47 (1968), pp. 723-759.

Clarke, Keneth K., and Hess, Donald T. Communication Circuits : Analysis and Design Reading, Mass. : Addison Wesley Publishing Co., 1971.

Davenport, W.B. Jr., "Signal-to-Noise Ratios in Band-Pass Limiters" Journal of Applied Physics, Vol. 24 (1953), p. 720-727.

Davidson, Frederic M., and Koh, Yutai T., "Interleaved Convolutional Coding for the Turbulent Atmospheric Optical Communication Channel" IEEE Transactions on Communications, Vol. 36 (1988), pp. 993-1003.

Davis, J. I., "Consideration of Atmospheric Turbulence in Laser Systems Design" Applied Optics, Vol. 5 (1966), pp. 139-146.

deWolf, D.A., "Saturation of irradiance fluctuations due to turbulent atmosphere" Journal of the Optical Society of America, Vol. 58 (1968), pp. 461-466.

Fried, D.L., "Aperture Averaging of Scintillation" Journal of the Optical Society of America, Vol. 57 (1967), pp. 169-175.

Gray, D.M., and Male, D.H., ed. Handbook of Snow: principles, processes, management & use New York: Pergamon Press, 1981.

Hohn, D. H., "Effects of Atmospheric Turbulence on the Transmission of a Laser Beam at 6328A. I-Distribution of Intensity" Applied Optics, Vol. 5 (1966), pp. 1427-1431.

Houghton, David D., ed. Handbook of Applied Meteorology New York: Wiley, 1985.

Hufnagel, R.E., and Stanley, N.R., "Modulation Transfer Function Associated with Image Transmission through Turbulent Media" Journal of the Optical Society of America, Vol. 54 (1961) p. 52-61.

Hullett, J.L. and Muoi, T.V., "A Feedback Receive Amplifier for Optical Transmission Systems" IEEE Transactions on Communications, (1976), pp. 1180-1185.

Jones, Kenneth A. Introduction to Optical Electronics, New York: Harper & Row Publishers, 1987.

Jones, J. J., "Hard Limiting of Two Signals in Random Noise" IEEE Transactions on Information Theory, (1963), p. 34 - 42.

Kasper, B.L., Campbell, J.C., Talman, J.R., Gnauck, A.H., Bowers, J.E., and Holden. W.S., "An APD/FET Optical Receiver Operating at 8Gbit/s" Journal of Lightwave Technology, Vol. LT-5 (1987), pp. 344-347.

Katzman, Morris Laser Satellite Communications Englewood Cliffs, NJ : Prentice-Hall, 1987.

King, B. G., Fitzgerald, P. J., and Stein, H. A., "An Experimental Study of Atmospheric Optical Transmission" Bell System Technical Journal, Vol. 62 (1983), pp. 607-629.

Koh, Y.T, and Davidson, F., "Interleaved Concatenated Coding for the Turbulent Atmospheric Direct Detection Optical Communication Channel" IEEE Transactions on Communications, Vol. 37 (1989), pp. 648-651.

Lawrence, Robert S., and Strohbehn, John W., "A Survey of Clean-Air Propatation Effects Relevant to Optical Communications" Proceedings of the IEEE, Vol. 58 (1970), pp. 1523-1545.

Laws, J.O., and Parsons, D.A., "The relation of Rain Drop Size to Intensity" Trans. Amer. Geophys. Union, Vol. 24 (1943), pp. 452-460.

Lillesaeter, O., "Parallel-Beam Attenuation of Light, Particularly by Falling Snow" Journal of Applied Meteorology, Vol. 4 (1965), pp. 607-613.

McCartney, Earl J. Optics of the Atmosphere: Scattering by Molecules and Particles New York: Wiley, 1976.

Mellor, M., "Light scattering and particle aggregation in snow" J. Glaciol., Vol. 6 (1966), pp. 237-248.

Ochs, G. R., Bergman, R. R., and Snyder, J. R., "Laser-Beam Scintillation over Horizontal Paths from 5.5 to 145 Kilometers" Journal of the Optical Society of America, Vol. 59 (1969) pp. 231-234.

Ochs, G.R., and Lawrence, R.S., "Saturation of Laser-Beam Scintillation under Conditions of Strong Atmospheric Turbulence" Applied Optics, Vol. 59 (1969), pp. 226-227.

O'Shea, Donald C. Elements of Modern Optical Design New York: Wiley, 1985.

Personick, S.D., "Receiver Design for Digital Fiber Optic Communication Systems, I & II" Bell System Technical Journal, Vol. 52 (1973), pp. 843-886.

Pratt, William K. Laser Communication Systems New York: John Wiley & Sons, 1969.

Rensch, D.B., and Long, R.K., "Comparative Studies of Extinction and Backscattering by Aerosols, Fog, and Rain at 10.6μ and 0.63μ" Applied Optics, Vol. 9 (1970), pp. 1563-1573.

Roden, Martin S. Digital Communication System Design Englewood Cliffs, NJ: Prentice-Hall, 1988.

Seagraves, Mary Ann, "Precipitation Rate and Extinction in Falling Snow" Journal of the Atmospheric Sciences, Vol. 41 (1984), pp. 1827-1835.

Smith, R.G., Brackett, C.A., and Reinbold, H.W., "Atlanta Fiber System Experiment : Optical Detector Package" The Bell System Technical Journal, Vol. 57 (1978), pp. 1809-1822.

Smith, Jack Modern Communication Circuits New York: McGraw-Hill, 1986.

Subramanian, M., and Collinson, J. A., "Modulation of Laser Beams by Atmospheric Turbulence" The Bell System Technical Journal, Vol. 44 (1965), pp. 543-546.

Subramanian, M., and Collinson, J. A., "Modulation of Laser Beams by Atmospheric Turbulence - Depth of Modulation" The Bell System Technical Journal, Vol. 46 (1967), pp. 623-648.

Tatarski, V. I. Wave Propagation in a Turbulent Medium New York: McGraw-Hill, 1961.

Van de Hulst, H.C. Light Scattering by Small Particles New York: John Wiley and Sons, 1957.

Van Muoi, Tran, "Receiver Design for High-Speed Optical-Fiber Systems" IEEE Journal of Lightwave Technology, Vol. LT-2 (1984), pp. 243-267.

Unwin, R.R., "A High-Speed Optical Receiver" Optical and Quantum Electronics, Vol. 14 (1982), pp. 61-66

Wang, Ting-i, Kumar, P.N., and Fang, D.J., "Laser rain gauge: near-field effect" Applied Optics, Vol. 22 (1983), pp. 4008-4012.

Yura, H. T., and McKinley, W. G., "Optical scintillation statistics for IR ground-to-space laser communication systems" Applied Optics, Vol. 22 (1983), pp. 3353-3358.

Ziemer, R.E., and Tranter, W.H. Principles of Communications : Systems, Modulation, and Noise. 2nd Edition Lawrenceville, NJ: Houghton Mifflin Co., 1985.

Vita

Christopher J. Hefele was born in Yonkers, New York on January 27, 1966. He graduated with honors from Lehigh University in 1988 with a Bachelor of Science degree in Electrical Engineering, with a minor in philosophy. Currently, he is working towards his Master of Science degree in Electrical Engineering at Lehigh as a research assistant. He is a member of IEEE, as well as Tau Beta Pi and Eta Kappa Nu honoraries. After receiving his MSEE degree in June, 1990, Mr. Hefele plans to join AT&T in Holmdel, NJ.