THE DEVELOPMENT OF HETERODYNE-BASED OPTICAL INTERSATELLITE COMMUNICATIONS «FINAL REPORT»

MPB

MPB TECHNOLOGIES INC.

1725 North Service Road | Trans-Canada Highway
Dorval, Québec H9P 1J1

MPB TECHNOLOGIES INC.
1725 North Service Road | Trans-Canada Highway
Dorval, Québec H9P 1J1
(514) 683-1490 Telex 05-823509

DOC-CR-SP-86-046

MPB

THE DEVELOPMENT OF HETERODYNE-BASED OPTICAL INTERSATELLITE COMMUNICATIONS «FINAL REPORT»

J. Gazdewich, A. Waksberg, C. Drèze

Prepared for:

Department of Communications
Communications Research Centre
P.O. Box 11490, Station H
Shirley Bay
Ottawa, Ontario
K2H 8S2

DSS Contract No. 1ST84-00177

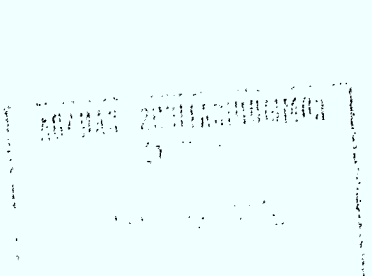
Prepared by:

MPB Technologies Inc.
1725 North Service Road
Trans Canada Highway
Dorval, Quebec
H9P 1J1

Dr. M.P. Bachynski
President

July 1986

S.O. 239



COMMUNICATIONS RESEARCH CENTRE

RELEASABLE

DOC-CR-SP-86-046

ORIGINAL DOCUMENT:REVIEW AND PUBLICATION RECORD

SECTOR	BRANCH	DATE
DGSTA	DSS	May 25, 1987

PURPOSE This form is for use during review of the DOC-CR contractor reports.
It is designed to: record decisions for classification,
record reasons for classification and cautionary marki
provide for indexing requirements.

INSTRUCTIONS * 1 copy of the completed form must accompany the contractor report package submitted to the CRC Library.

* Complete the following items as applicable.

1. DOC-CR NO.	2. DSS CONTRACT NO.	
DOC-CR-SP-86-046	36001-4-1050	
3. TITLE: The development of heterodyne-based optical intersatellite communications : final report		4. DATE July 1986
5. CONTRACTOR MPB Technologies Inc.		
6. SCIENTIFIC AUTHORITY B. Felstead	7. LOCATION DSS/CRC	8. TEL. NO. 998-2373

9. CONTRACTOR REPORT CLASSIFICATION:

RELEASABLE [XXX] CONDITIONALLY RELEASABLE [ ] NON-RELEASABLE [ ]

* REASONS FOR CLASSIFICATION:

10. NO. OF COPIES SUBMITTED TO LIBRARY:

EXECUTIVE SUMMARY [ ] FINAL REPORT [XXX] 5 copies

....

Scientific Authority's Signature Date

This form is not official therefore it is not signed.

TABLE OF CONTENTS

TABLE OF CONTENTS - Cont'd.

TABLE OF CONTENTS

ABSTRACT

A study was made to develop key technologies in intersatellite links (ISL) based on the laser diode. Theoretical work was done including determinations of the receiver sensitivity as a function of a number of parameters such as bit rate, type of detectors, gain and detection techniques, both direct and heterodyne. Investigations were also made on modulation formats, laser diodes, detectors, noise and noise reduction techniques as well as acquisition and tracking.

A communication testbed was built to study the problems associated with a laser diode based system. The transmitter used a GaAlAs laser diode operating at 830 nm and was modulated up to 1 Gbit/s. Both direct and heterodyne detection were employed using Silicon avalanche photodetectors. Laser and detector noises were measured and found to be within their expected ranges. The advantages in sensitivity of heterodyne detection were found to depend on the success of removing excess noise from the local oscillator.

The results of our study show that Canada has the capability to participate fully in the ISL high technology worldwide activity. A Canadian Industrial Strategy to the future ISL market is proposed.

1.0 INTRODUCTION

This is the final report for DSS Contract No. 1ST84-00177 entitled "The Development of Heterodyne-Based Optical Intersatellite Communications".

The contents of this report will describe work that was performed at MPB Technologies Inc., both theoretical and experimental, to develop key technologies in optical intersatellite links (ISL) based on the laser diode.

The work included a survey of the state-of-the-art as well as original system studies to fill in gaps not found in the literature, and experimental work on a testbed to identify the potentials and problems of this new technology. Many aspects of ISL have been considered and discussed in this report. These include the transmitter and receiver subsystems, modulation formats, sensitivity of the receiver, noise, aquisition and tracking, etc.

The breadboard that was built in the lab provided us with invaluable first hand experience on the idiosyncrasies of this technology. The transmitter used a Ga AlAs laser diode emitting at 0.83 μm and was modulated up to 1 Gbit/s. The same type of laser was used as the local oscillator. Direct and heterodyne detection were studied using silicon avalanche photodetectors. FSK modulation was successfully investigated up to 1 Gbit. This work permitted us to evaluate realistic expectations of ISL using laser diodes. This report therefore includes a list of problem areas that should be tackled to make this technology a reality.

In view of Canada's present involvement, recommendations are also made that Canada participates fully in the ISL high

technology field and Canadian industrial strategy to the future ISL market is proposed.

2.0 STATE-OF-THE-ART

Research into the use of heterodyne techniques in optical communication systems is primarily motivated by a significant gain in receiver sensitivity. It has been estimated that the use of heterodyne detection could add a potential 15-20 dB[2.1] to the system power budget. Such a gain could increase repeater spacings in long haul fiber optic communication systems by as much as 100 km. In inter orbit satellite-to-satellite communications such a gain in receiver sensitivities would allow the use of low power lasers and/or a reduction in antenna size. Such considerations are a very important aspect of satellite optical communication system design.

The bulk of R&D activity is currently found in the field of fiber optic communications. Although all optical heterodyne communication systems are fundamentally the same, there are some significant differences in optical fiber heterodyne systems, and those appropriate for space communication. These differences rest primarily with the type of semiconductor laser most suitable for each. Notwithstanding these differences, much of the research into fiber systems is certainly applicable to satellite systems.

Section 2.1 will present a brief discussion of various considerations related to the space environment, outlining both operational and environmental issues important to system design. Section 2.2 then proceeds with a review of the fundamental elements of an optical heterodyne communication system. Section 2.3 discusses some of the inherent problems in such a system. It is on these problems that current research and development effort

is focussed. Finally Section 2.4 will survey the state-of-the-art and the progress made in developing solutions to some of these problems.

2.1 Space Requirements

Operation of optical communication systems in a space environment places unique boundary conditions on system design. These conditions affect both the designs of system operation and environmental protection. A brief discussion of these space-related factors follows.

2.1.1 Operational Factors

Communications in a space environment are unimpeded by beam attenuation or dispersion in the beam transport medium as is experienced in fiber optic or atmospheric systems. However, because of beam spreading due to diffraction effects, the captured optical power at the receiver is only a fraction of the total signal power, for large baselengths L. This beam spreading is a function of wavelength, exit aperture size and spacecraft separation. The intensity, on the optical axis, at the receiver is given by:

$$I_o(L) = \frac{2}{\pi} \frac{P_L}{D^2(1+4L^2/k^2D^4)} \; ; \; k = \frac{2\pi}{\lambda} \qquad (2.1.1)$$

which for large baselengths L, reduces to:

$$I_o(L) = 2\pi \frac{D^2 \; P_L}{\lambda^2 \; L^2} \qquad (2.1.2)$$

where D = exit aperture diameter

λ = laser wavelength

L = baselength (i.e. T_x - R_x distance)

P_L = launch power.

Therefore, it is obvious that in order to maximize received optical power the system designer should tend towards short wavelength sources whose power is maximized. This is perhaps the single greatest advantage of optical systems when compared with microwave systems, since optical wavelengths are typically four orders of magnitude shorter than microwave, resulting in greater optical antenna gains.

Ensuring adequate received power levels for high speed, high performance systems is a major concern, considering the potentially large baselengths possible in ISL links. Inter orbit baselengths may be as large as 55,000 km. Assuming a 10 cm diameter transmitting telescope and a 20 cm diameter receiving telescope a power loss of approximately 60 dB may result for a short wavelength semiconductor laser. Although the telescope (antenna) sizes are very much a fluid design parameter, these figures suggest that practical optical sources should operated at power levels at least on the order of 10 dBm - 20 dBm.

An ISL system must also account for the fact that the baselength is constantly changing. This situation is accentuated for communication between non-geostationary satellites. The condition of variable baselength presents two problems to an ISL communication system. The first is a non-constant average received power level since received power has been shown to vary as the inverse square of the baselength. Thus some sort of automatic power control would likely be required to counteract

this effect. Secondly a problem with dynamic carrier frequencies caused by doppler shifting may result for significant satellite relative velocities. This could pose a problem for systems requiring a high degree of signal spectral purity. Solutions to the problem may involve tracking out the doppler shift or utilizing wide optical filters such that the signal is prevented from drifting outside the optical window of the receiver.

The use of a large optical window may be precluded by receiver noise limitations due to high levels of background radiation. Sources of such optical background include the sun, moon, airglow and general stellar background. Of these sources, the greatest amount of care must be taken to protect against solar noise. Hence relatively narrow optical filters may be a requirement. Another aid in guarding against background noise is to restrict the receiver field of view, taking care that intense point sources such as the sun or moon are not, or at least are rarely seen by the detector. However, the narrower the system field-of-view, the more difficult become pointing and tracking problems.

Because of the large distances between spacecraft, the problems of signal pointing, acquisition and tracking may indeed be quite formidable. Although the beam spot size at the receiver may be on the order of km's in diameter, corresponding baselengths on the order of tens of thousands of kilometers result in pointing demands better than fractions of an arc second. Such accuracy is needed for signal acquisition. However the power density of the beam is not constant over the entire radius of the beam spot, therefore, although acquisition may be

obtained on the perimeter of the spot, system operation must occur near the center where the power density is at a maximum. This further increases the severity of the pointing requirements. Once signal has been acquired, a tracking function must be performed. This tracking capability is necessary for two reasons. First the spacecrafts are never completely stationary with respect to each other, changing not only in separation, but in position as well. Secondly, there is always present a certain amount of spacecraft jitter. This problem would be most severe at the transmitting end where there is enormous leverage. This jitter must be tracked out in the system pointing optics. There may also be a need for point-ahead-optics for tracking spacecraft with high relative velocities such as may occur in communication between satellites in non-geostationary orbits.

2.1.2 Environmental Factors

One of the most noteable characteristics of space systems is a lack of easy access for the purposes of maintenance, repair or control. Because of this fact, space systems must be designed and manufactured with a high degree of reliability. Satellite systems are expected to demonstrate a lifetime of typically seven years. Any redundancy built into the system to help meet these reliability requirements must be automatically controlled.

Since the system is in a remote and hostile environment, it must be responsible for the generation of its own supply power, unlike terrestrial systems where there is relatively easy access to power supply. Because of this situation, power on board a satellite is at a premium. Therefore it is important that all systems are as power efficient as possible. In

addition, the availability of certain supply voltage and current levels may be restricted.

Again, because of the remote location of space systems, and because of the difficulty and cost of transport to these locations, it is important to minimize system volume and weight as much as possible. Minimizing system volume results in either a smaller satellite, or in allowing the provision of greater capabilities in a satellite of given size. Minimizing system weight results in either reduced launch costs since launch vehicle fuel requirements are governed by vehicle weight, or in easing weight requirements for other systems while still maintaining a set weight budget for the spacecraft as a whole.

Therefore, in short, the remote location of space systems presents a requirement for small, light, rugged, reliable and power efficient subsystems. This requirement implies subsystems with as low a complexity level as is feasible in order to get the job done.

Any ISL communication system must be designed and constructed in a manner such that it will survive the rigors of launch. This demands a system resistance to severe mechanical shock, vibration and stress.

The space environment has associated with it many unique hazards that any ISL system must withstand. Among these hazards is a severe thermal cycling that may occur if the spacecraft passes in and out of the sun's shadow. Indeed, high thermal gradients will exist within the satellite itself, especially if it is despun. The ISL system must be able to withstand bombardment by various sources of radiation including cosmic

rays, Van Allen belt-trapped radiation and magnetospheric boundary accelerated solar particles, nuclear events producing hard X-rays, γ-rays and neutron/proton radiation. Other potential hazards include contamination from rocket outgassing and micro-meteorite bombardment.

This section has discussed in a qualitative sense some of the special considerations that must be accounted for in the design of a satellite ISL optical communication system. Keeping these in mind, a study regarding the preliminary design and feasibility of a high capacity optical heterodyne communication system will be presented in the balance of this report.

2.2 Optical Heterodyne System Elements

The general theory of heterodyne communications is well known from radio and microwave applications. However, the details of actual implementation can differ substantially between various systems. Certainly this is true with respect to optical and microwave systems. The fundamental operational elements of a laser based optical heterodyne system are shown in Figure 2.2.1. The individual components are described in detail in Chapter 6.

The most critical components in such a system are the lasers (the local oscillator indicated in Figure 2.2.1 is also a laser similar to that used in the transmitter). It is the laser that governs the critical system parameters of signal power, AM and FM noise and mixing efficiency. The lasers currently under consideration for use in fiber optic systems and now for intersatellite systems are single mode semiconductor injection lasers. Most of the necessary development work required to bring these systems into practice revolve around these lasers.

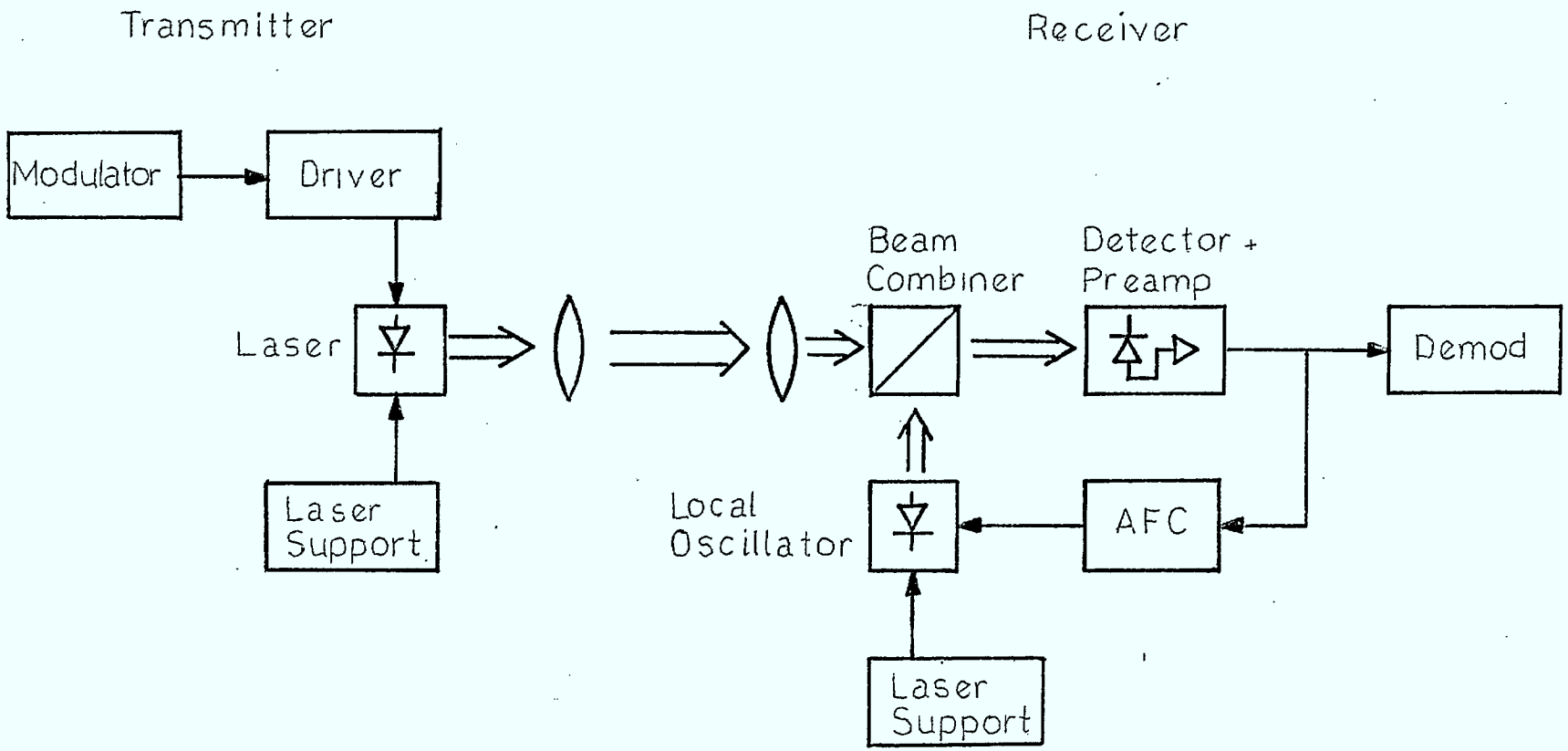


FIGURE 2.2.1: Fundamental Optical Heterodyne System

The signal and local oscillator beams are mixed at the receiver in an optical beam combiner. The beam combiner may simply be a conventional beam splitter. The mixed beam is then detected by a photodiode which recovers the beat frequencies between the two beams. This electrical signal is amplified and fed into the system electronics.

The local oscillator is frequency stabilized by an adjustment of its bias current. The adjustment is made by the error signal derived from a frequency discriminator used in the automatic frequency control (AFC) loop. This loop senses the stability of a specific frequency component in the electrical intermediate frequency (IF) spectrum. In many respects the local oscillator laser can be thought of as a VCO.

The laser drive electronics in the transmitter represent circuitry unique to a semiconductor laser heterodyne communication system. These electronics are used to directly modulate the transmitter laser according to a specified modulation format. The operation of the modulation and demodulation electronics should be similar to those generally used in other systems. However, actual implementation may again present some new problems because of the speeds capable of these lasers and because of the level of integration demanded by satellite systems.

Modulation formats that have received consideration for use in optical heterodyne systems are basically the shift keyed formats commonly used in heterodyne communication systems in general. These include amplitude shift keying (ASK), frequency shift keying (FSK) and phase shift keying (PSK) and variations on

these formats such as differential phase shift keying (DPSK), minimum frequency shift keying (MFSK), etc. M-ary techniques can also be employed in these systems. These modulation formats will be discussed in greater detail in Chapter 4. However, it can be said that the FSK and DPSK formats have been found to be the most attractive for use in optical heterodyne systems.

2.3 Areas of Technical Concern

Many properties of semiconductor laser diodes are ideally suited for space application. There are however, many properties that pose severe technical problems for ISL optical heterodyne communications. The two major of these are listed below.

1. Semiconductor lasers are restricted to relatively low output powers. These limits are caused by thermal damage due to high injection currents and to optical damage at the laser facets caused by a high electric field. Currently available single mode GaAlAs lasers are rated at 30 mW in cw operation. Lasers rated to 50 mW are expected in the near future. The prime motivator for high power short wavelength lasers has been the optical recording industry. Long wavelength InGaAsP lasers, made for the fiber optics industry, have their powers limited typically between 3 mW and 5 mW. There is no incentive for power levels to be increased beyond this point because of nonlinear optical effects which begin to occur in the fiber. Multimode lasers are commercially available with cw output powers as high as 200 mW.

2. Semiconductor lasers tend to exhibit a very dynamic spectrum. Unlike gas laser systems where optical transitions occur in discrete energy states, semiconductor lasers can experience optical transistions over a band of energy states. This results in a broad gain spectrum extending over many cavity modes. This spectrum is also dynamic, being a function of carrier density (injection current) and temperature. A drift in the gain spectrum can result in mode hopping behaviour and partition noise, both of which are detrimental to heterodyne applications. Single mode lasing has been established in those lasers with the development of index guiding structures. Single mode behaviour is mandatory in heterodyne systems. The single modedness or spectral purity of the laser is also proportional to the lasing power. The mode's linewidth is inversely proportional to the laser power. Hence, the higher the operating power of the laser, the greater is the spectral purity of the laser. The actual lasing or dominant mode is determined by the peak gain wavelength.

In addition to possessing a dynamic gain spectrum the laser also has a dynamic cavity mode structure. The laser cavity is in essence a Fabry Perot cavity where the mode structure is a strong function of cavity length and group refractive index. The refractive index is influenced to varying degrees

by the Al content in the active layer, the doping, temperature, absolute carrier concentration, and gain. Of these effects the first two result in laser to laser differences, the third usually results in long term drifts while the last two can result in short term instabilities and operating point dependence. The last three effects allow the laser to be temperature and current tuned while the gain dependence allows it to be FM modulated.

The existence of a dynamic gain and cavity mode structure make frequency matching and stabilization very difficult. The matching requirements (i.e. frequency offset or IF, linewidth, and long and short term stability) are set by the particular modulation format in use and the magnitude of the transmission rate.

When communications are attempted at high frequencies approaching the Gbit/s region, additional technical problems are encountered. An immediate problem is in ensuring sufficient bandwidth at the receiver. This involves the use of a high speed yet sensitive detector. A special detector has been developed for fiberoptic applications. This is the SAGM (separate absorption graded multiplication) Avalanche Photodiode Detector (APD). Unfortunately, this device is made from InGaAs/InP and so is not well matched to GaAlAs lasers. Si-APD's

which have peak efficiency around the GaAlAs laser wavelength are commercially available with bandwidths in excess of 2 GHz and Si-p-i-n photodiodes are commercially available with bandwidths in excess of 5 GHz.

In addition to the high speed electro-optic components, high speed modulation and demodulation electronics will be needed if Gbit/s transmission rates are desired. Such electronics are now becoming available with the development of GaAs electronic technology, especially integrated circuit technology.

2.4 Technology Survey

There has been a tremendous research and design effort in the area of optical heterodyne technology since 1980, when an initial optical heterodyne detection experiment was conducted using directly frequency modulated semiconductor lasers. Evidence of this activity is given by the considerable amount of papers being published in the literature and by the devotion of technical sessions to the subject at major scientific conferences such as the Conference on Optical Fiber Communication and CLEO, among others. Substantial progress has been made in developing this complex technology into a commercial reality.

The bulk of heterodyne, or coherent, research has been taking place in the field of fiber optic, long haul communications. The objective is to expand repeater spacings and increase the data rate by taking advantage of the increase in receiver sensitivities offered by coherent transmission technology. Because transmission is through fiber, the lasers are being developed for operation at $\lambda = 1.55$ μm where there is a minimum fiber loss window of $\sim$ 0.2 dB/km. The current status

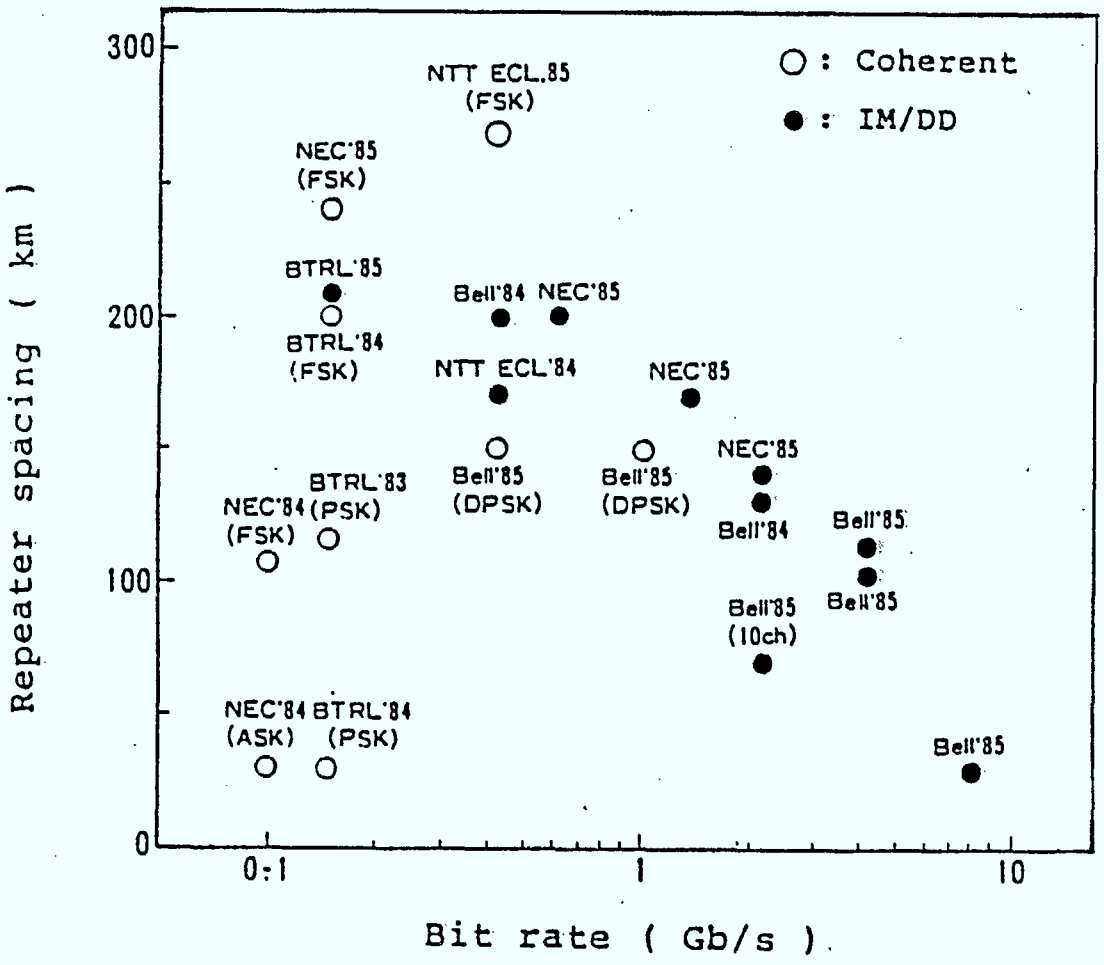


FIGURE 2.4.1: Current Status of Coherent Optical Transmission Systems

of this research is shown in Figure 2.4.1. A common system performance measure in fiber optic system analysis is bit rate distance product. Link loss can be estimated by using a fiber loss of approximately 0.22 dB/km for 1.55 μm systems and 0.42 dB/km for 1.3 μm systems. Recent research has concentrated on 1.55 μm systems using InGaAsP lasers. Figure 2.4.1[2.2] shows that most systems to date have been designed for operation between 100 - 500 Mbit/s with FSK proving to be a popular choice of modulation format. ASK was a common modulation format in early systems, but was soon replaced by FSK and DPSK formats. These formats offer an improved performance/ease of implementation tradeoff. High bit rate systems are still dominated by intensity modulated, direct detection systems, however it must be realized that conventional direct detection technology is much more mature than coherent technology. Progress in the two technologies, using this performance measure, are compared in Figure 2.4.2.[2.3] The rapid development of coherent systems is clearly indicated. Receiver sensitivities in recent coherent experiments have shown improvements over equivalent direct detection systems of 5.4 dB in a 280 Mbit/s, 204 km, FSK, single filter detection system,[2.4] and 1.7 dB and 7.5 dB in a 400 Mbit/s and 1 Gbit/s, DPSK system respectively.[2.5]

A laboratory demonstration of an end-to-end link operating at 100 Mbit/s designed for space application has recently been made.[2.6] An experimental package using heterodyne technology and consisting of the transmitting end of an optical communication system capable of a 220 Mbit/s data rate has been

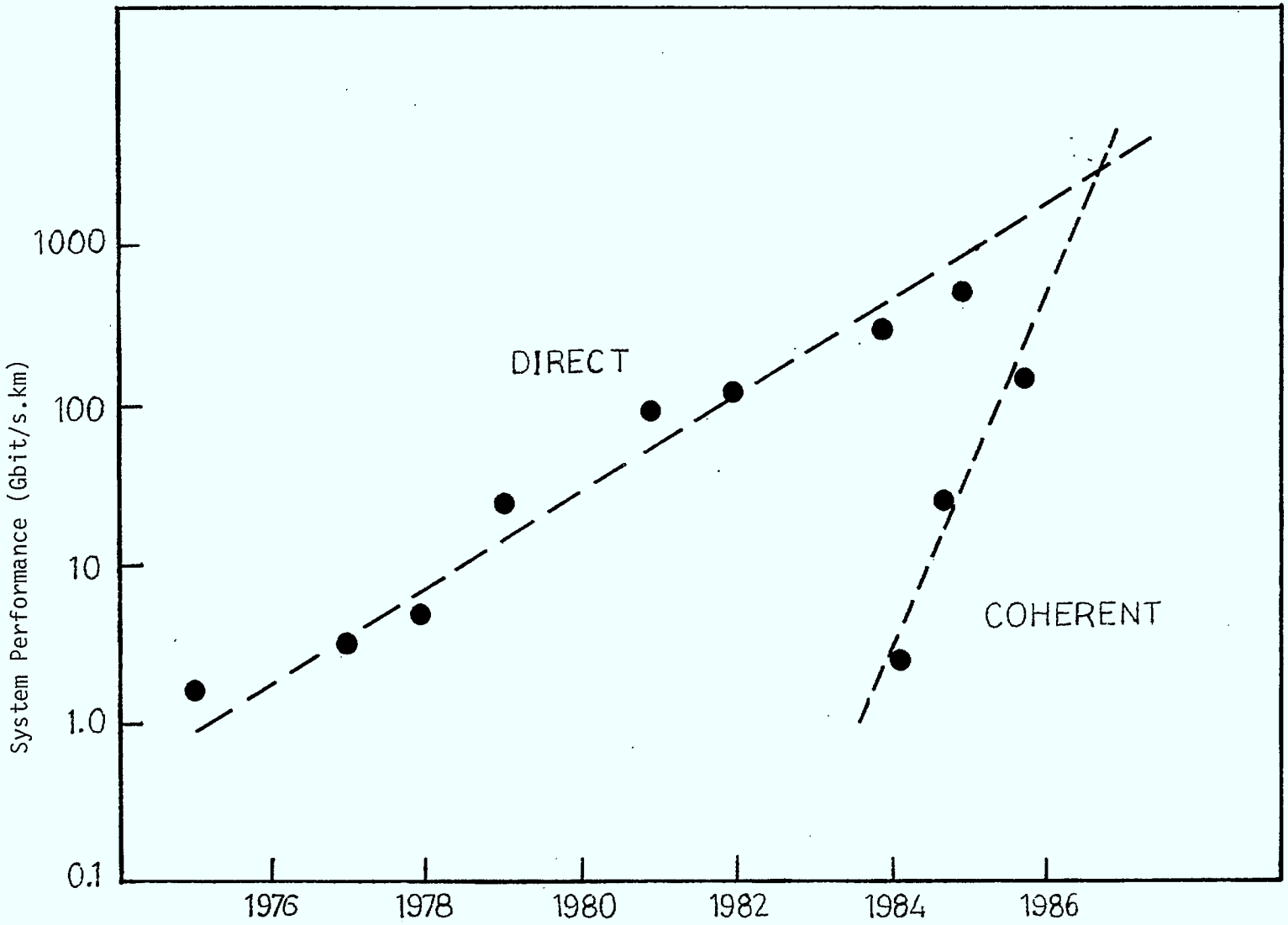


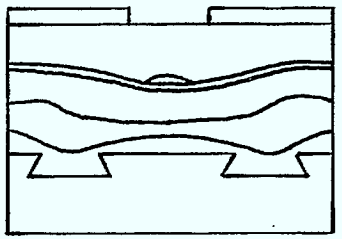
FIGURE 2.4.2: Progress in Coherent System Performance

proposed for space qualification on the NASA Advanced Communications Technology Satellite (ACTS), scheduled for launch in late 1989. A system capable of full duplex heterodyne operation is being planned for space deployment around 1991.

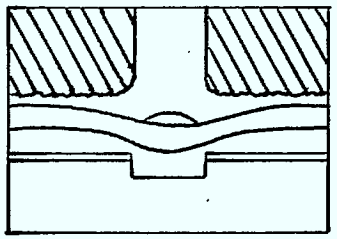
Communication systems designed for satellite application are not laser power restricted by the non-linear optical effects displayed by optical fibers. Currently available single mode, GaAlAs lasers are capable of 10 dB higher output power levels than those currently used in fiber systems. It is anticipated that these levels will increase still further in the near future.

A large amount of device research in GaAlAs lasers has been directed towards high power single mode lasers. Three major device structures currently exist for high power AlGaAs lasers. These are the large-optical-cavity structures, the thin-active-layer structures and nonabsorbing mirror structures. Lasers designed with LOC structures and TAL structures are shown in Figure 2.4.3. The best published results of light-current characteristics for GaAlAs devices of these structures are shown in Figure 2.4.4.[2.7] The Matsushita diode has emitted a record single mode cw power at 100mW and a maximum cw power of 200mW. RCA's CDH-LOC laser with a regrown nonabsorbing mirror achieved a peak pulse power of 1.5W with pulses 100 ns wide which is the highest peak power yet achieved from a single mode laser. The bulk of the laser was catastrophically destroyed at this power level. Although several types of these lasers demonstrated cw powers in excess of 100 mW, their reliable cw power levels are presently limited to between 15 mW - 30 mW. These levels are expected to increase to between 30 mW and 60 mW in the near

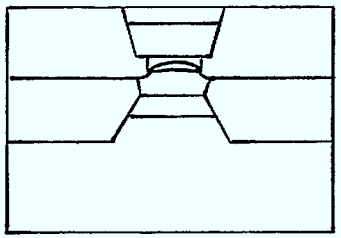
LARGE-OPTICAL-CAVITY STRUCTURES



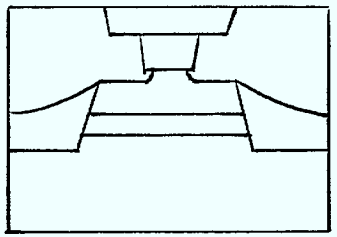
Constricted Double Heterostructure (CDH-LOC)



Channeled Narrow Stripe (CNS-LOC)

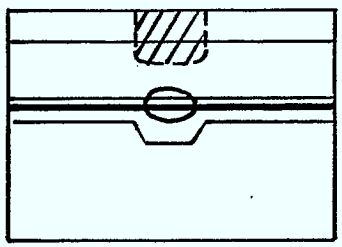


Buried Coarctate Mesa (BCM)

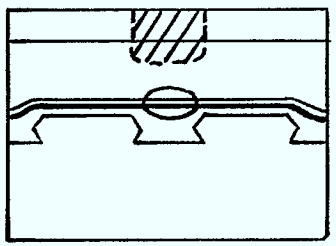


Self-aligned Strip Buried Heterostructure (SSBH)

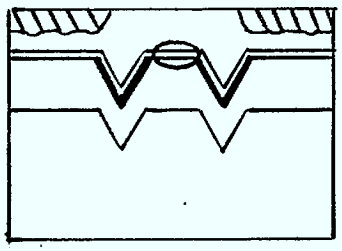
THIN-ACTIVE-LAYER STRUCTURES



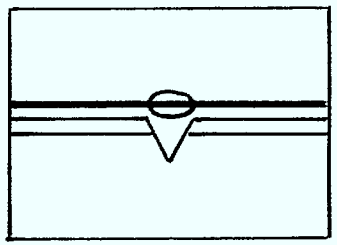
Channeled-Substrate Planer (CSP



Twin-Ridge Substrate (TRS)



Twin-Channel Substrate Mesa (TCSM)



V-Groove-Substrate Inner Stripe (VSIS)

FIGURE 2.4.3: Current High Power Laser Structures

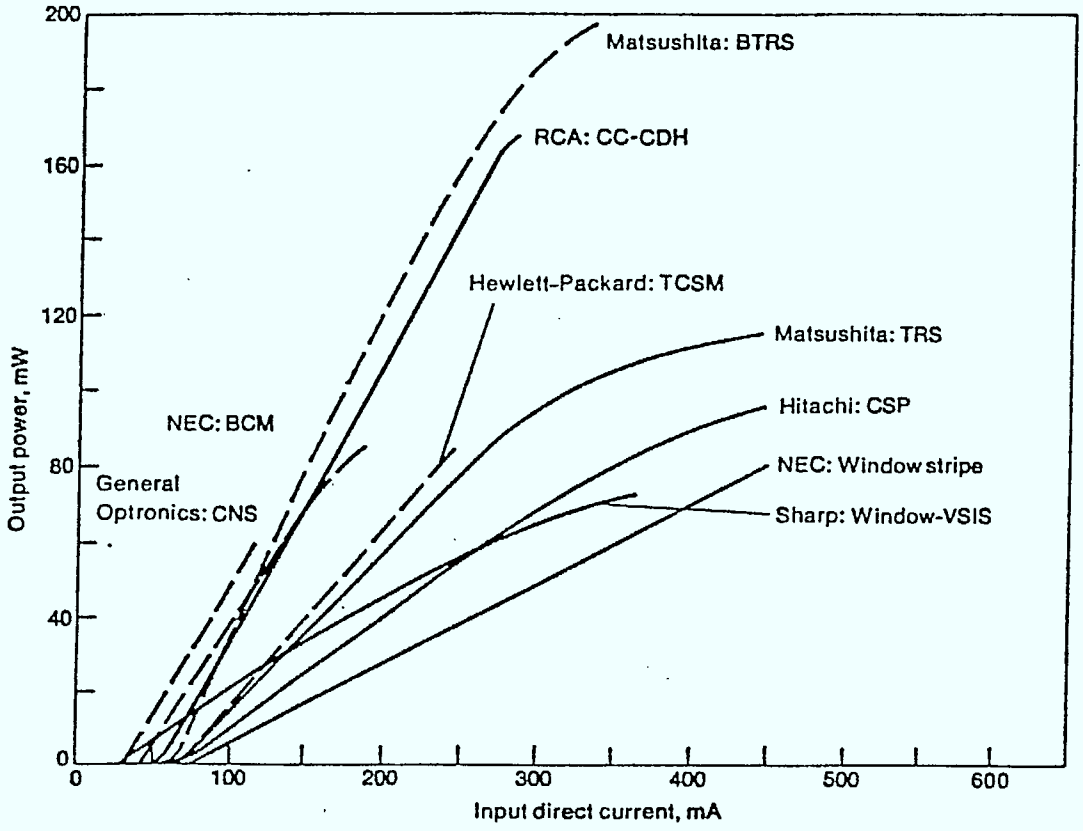


FIGURE 2.4.4: Best Published Results of Light-Current Characteristics for GaAlAs Devices of the Structures Shown in Figure 2.4.3.

future.

Several structures have also been developed in an effort to improve spectral purity. A number of these structures are shown in Figure 2.4.5. The structures that have received the greatest amount of attention are the distributed feedback (DFB), cleaved coupled-cavity (C^3), and external cavity. External cavities have been constructed using curved mirrors, gratings, and silvered grin-rod lenses referred to as grin rod external coupled cavity (GRECC) structures.

Narrow laser linewidths are important in reducing noise in the heterodyne process and in facilitating locking of the AFC loop. When the IF linewidth becomes on the order of the modulation bandwidth then significant sideband distortion or crosstalk in FSK systems can occur. Acceptable limits are determined by the transmission rate and by the particular modulation format and detection method chosen. It should be possible to tolerate larger linewidths in high speed envelope detecting systems. Should the linewidths become large relative to the bandwidth of the frequency discriminator in the AFC loop, then tracking and stabilization of the center frequency can be seriously degraded.

The natural linewidths of semiconductor lasers are typically in the range of 10 MHz - 100 MHz, full width at half max (FWHM). Long term laser drifts may well extend into the GHz region. Doppler shifts of up to 10 GHz at a rate of 10 MHz can be expected in satellite systems. It is therefore absolutely necessary that efficient frequency tracking loops be included in semiconductor laser heterodyne systems. However, even once

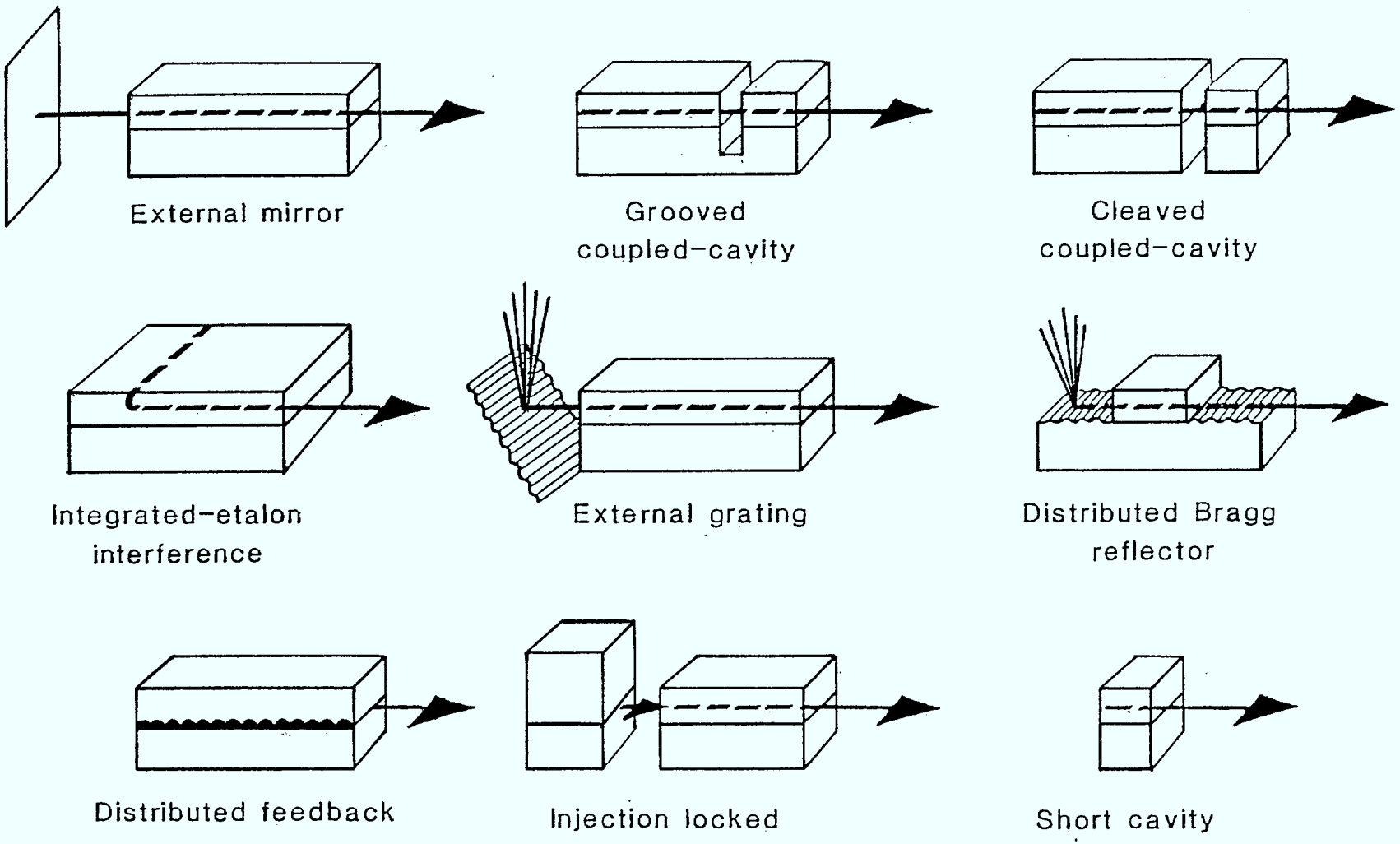


FIGURE 2.4.5: Laser Structures Used for Improving Laser Spectral Purity

frequency stabilization has been obtained, linewidths on the order of tens of megahertz are generally unacceptable. This provided the incentive for development of the structures shown in Figure 2.4.5.

Distributed feedback lasers have been receiving the largest amount of attention in fiber optic research. These lasers are fabricated from InGaAsP and lase at 1.55 μm. Linewidths of 1 MHz and below have been obtained, however these lasers are currently being used in conjunction with external cavity configurations to provide further enhanced linewidth narrowing down to hundreds of kilohertz. Unfortunately, such lasers are not common at short wavelengths because of the difficulty in constructing the necessary short period gratings. External cavity configurations are much more common at short wavelengths. Linewidths down to tens of kilohertz have been reported. Although longer cavities result in narrower linewidths, they also degrade FSK modulation efficiency. In addition long cavities are more difficult to stabilize. Therefore tradeoffs must be made to reach an optimum cavity length. Grating feedback is generally employed as it allows greater tuning capability. Also, adjustment of cavity length by piezoelectric elements increases frequency stabilization capability.

Laser linewidths and frequency stability have not yet reached the stage where homodyne and PSK techniques requiring automatic phase control can be effectively used.

Optics specially designed for laser diodes are becoming

commercially available. Diffraction limited AR coated lenses are being designed for laser collimation, beam expansion and receiving. Anamorphic prism pairs are being made for beam circularization. Optical isolators are used in some systems to prevent reflected light from feeding back into the laser cavity and in doing so, degrading laser dynamic performance. The isolators are usually based upon the Faraday Rotator. Optical mixing is usually performed by conventional beam splitters. In fiber optic systems fiber couplers are used. Waveguide mixers may eventually find use in heterodyne systems. Miniature optical components (i.e. lenses, grating, prisms, etc.) may become very important to these systems, especially if stable, compact external cavities prove necessary. Grin-rod (graded refractive index) lenses have already found common use.

InGaAs APD's and InGaAs PIN: GaAs FET detectors are those commonly employed in fiber optic heterodyne systems. An SAGM quarternary APD has recently been developed by J.C. Campbell et al at AT&T Bell Laboratories.(2.8) This device is reported to exhibit a gain-bandwidth product of 60 GHz. Unfortunately such devices are not well matched to GaAlAs lasers at .85 m. These short wavelength systems use Si-APD's or Si-PIN's, APD's being more common. At these wavelengths the low excess noise of silicon APD's has allowed < 300 photons/bit for 10^{-9} error rate.(2.9) High speed GaAlAs/GaAs PIN photodiodes are also available at these wavelengths. Both Si based and III-V based detectors have been produced with bandwidths approaching 10 GHz. Custom devices are capable of still higher speeds. Continued research and development should increase these bandwidths still

further in the near future.

Some of the best direct detection receiver results at a number of wavelengths and bit rates, and a variety of detector/amplifier combinations are shown in Figure 2.4.6.[2.9] The decrease in receiver sensitivity with increasing bit rate is clearly evident from this figure.

As modulation rates exceed 565 Mbit/s and approach the Gbit/s region, the availability of adequate electronics becomes a major problem. Laser drivers will require the use of GaAs FET components. IF frequencies in excess of 1 GHz will require the use of microwave components. Electronics involving logic functions will again most likely use GaAs IC technology although Si-bipolar ECL may be used below 1 Gbit/s if custom designed. Manufacturers of high speed GaAs logic circuits have been marketing devices for over a year now, and are rapidly bringing new devices on line.[2.10] The level of integration is reaching LSI levels with some manufacturers offering semi-custom gate arrays.[2.11]

In addition, appropriate measurement and test equipment is only now becoming available for operation at Gbit/s plus levels. Bit error rate transmission test equipment in particular is currently being offered by at least three manufacturers.

Another problem that is encountered when working above GHz frequencies involves the intrinsic noise behaviour of the laser. Both the AM and FM noises of the laser peak at the resonance frequency of the laser. The resonance peak typically occurs between 2 - 5 GHz at lasing and moves to higher frequencies with an increase in power. Therefore operation at these frequencies

	N	B	λ	Det	Amp	Auth	Year
1	225	250	0.83	SA	SB-TZ	Muoi	1982
2	264	50	0.85	SA	SF-HZ	Runge	1976
3	591	34	1.52	GA	GF-TZ	Walker	1984
4	640	6.3	0.9	SA	SF-HZ	Goell	1973
5	642	140	1.52	GA	GF-TZ	Walker	1984
6	847	2000	1.54	IA	GF-HZ	Kasper	1984
7	930	420	1.55	IA	GF-HZ	Campbell	1983
8	1046	140	1.3	IP	GF-HZ	Smith	1984
9	1077	45	1.3	IA	GF-TZ	Forrest	1981
10	1186	446	1.52	GA	GF-HZ	Toba	1984
11	1436	1200	1.53	IP	GF-HZ	Brain	1984
12	1440	4000	1.51	IA	GF-HZ	Kasper	1985
13	1635	565	1.3	IP	GF-HZ	Smith	1982
14	1872	800	1.3	GA	SB-LZ	Yamada	1982
15	2053	45	1.3	IP	GF-TZ	Williams	1982

N = Photons/Bit

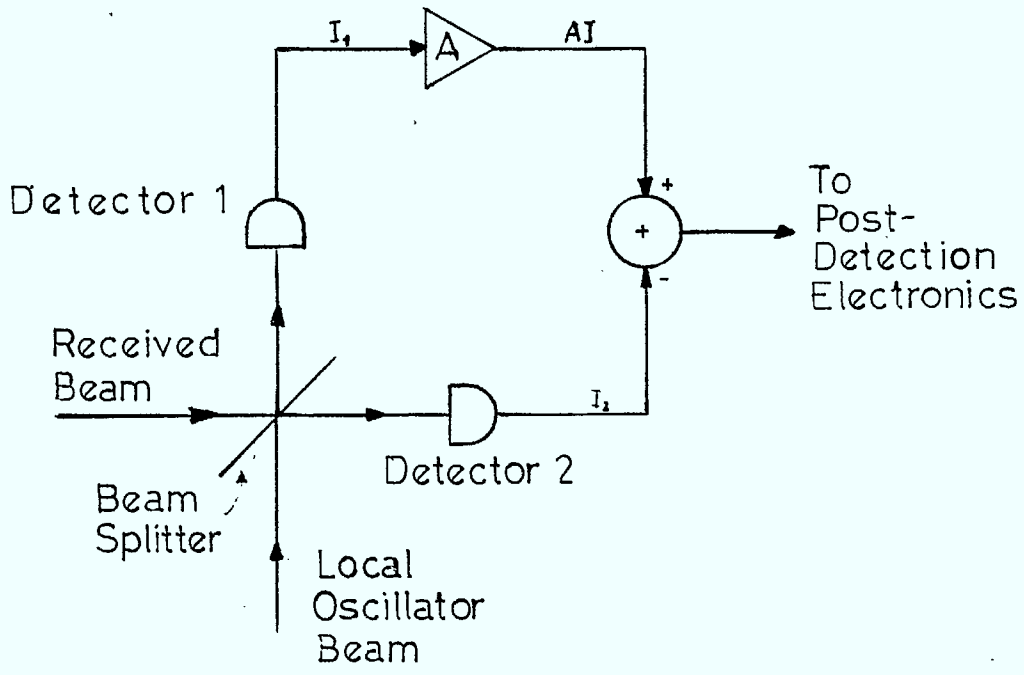
B = Bit Rate in Mbit/s

FIGURE 2.4.6: Comparison of Direct Detection Receiver Sensitivities

can cause a degradation in the system signal-to-noise ratio. This problem has been alleviated to a large extent by the development of a dual balanced detector receiver. A schematic of such a receiver and its measured noise performance is shown in Figure 2.4.7.[2.12] Such a receiver was used in the 1 Gbit/s DPSK coherent experiment mentioned earlier.[2.13] The experimental configuration is shown in Figure 2.4.8. The AM excess-noise currents from the two photodetectors are in phase with each other while the signal components are out of phase. Therefore, by differentially combining the two outputs, the AM excess noise is cancelled while the signal is completely recovered. The FM noise cannot be removed by this technique.

Finally, one recent development worth noting is that of a packaged frequency-stable tunable, 20 KHz linewidth, 1.5 μm InGaAsP external cavity laser.[2.14] The packaged grating-tuned external cavity InGaAsP/InP laser was developed for use in coherent transmission systems operating at a wavelength of 1.5μm. The lasing frequency is selectable with a 600 A range during manufacture and can subsequently be tuned over a 5 A or 63 GHz range. Long term relative frequency stability of better than 150 MHz was demonstrated. A schematic diagram of the device is shown in Figure 2.4.9.

Experiments have also been performed to demonstrate the ability of optical heterodyne systems to perform frequency division multiplexing.[2.15] The concept of optical FDM is similar to that commonly used in microwave and radio communications. Frequency channels are defined by the lasing frequency difference between the transmitter and local oscillator



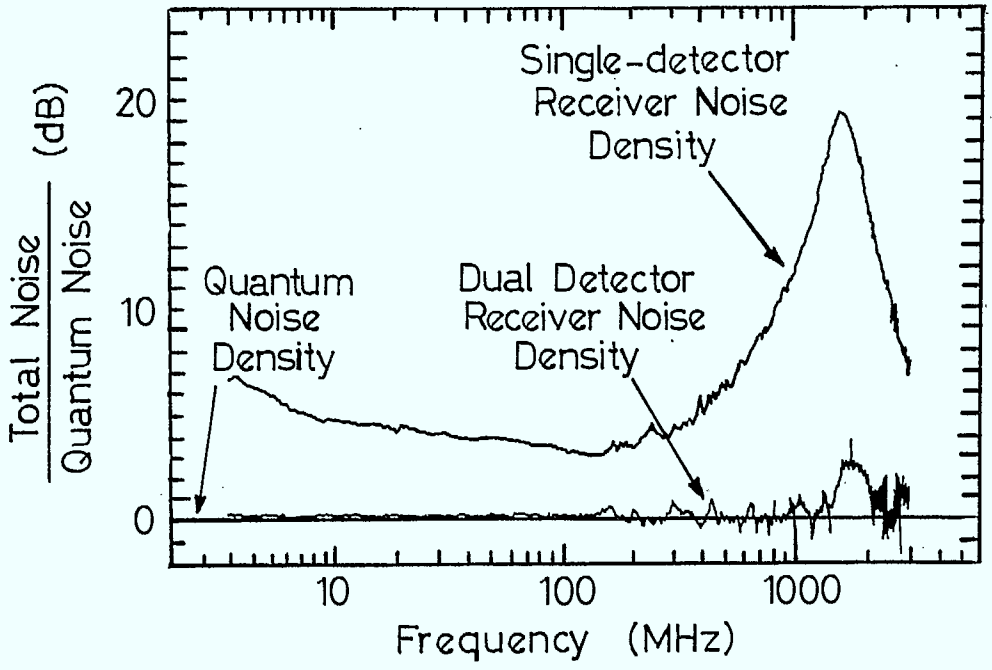


FIGURE 2.4.7: Schematic of a Dual Balanced Detector Receiver and Measured Noise Performance

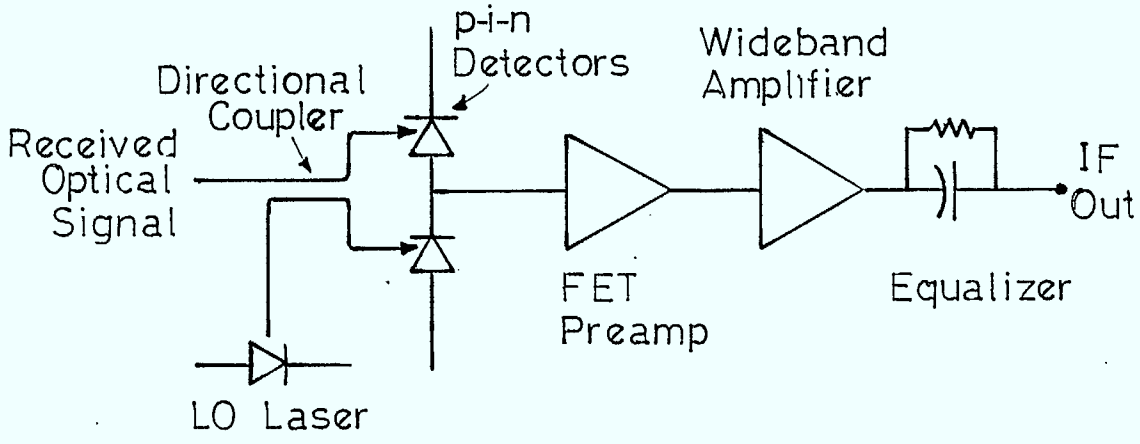


FIGURE 2.4.8: Dual Balanced Detector Configuration Used in the 1 Gbit/s DPSK Coherent Experiment

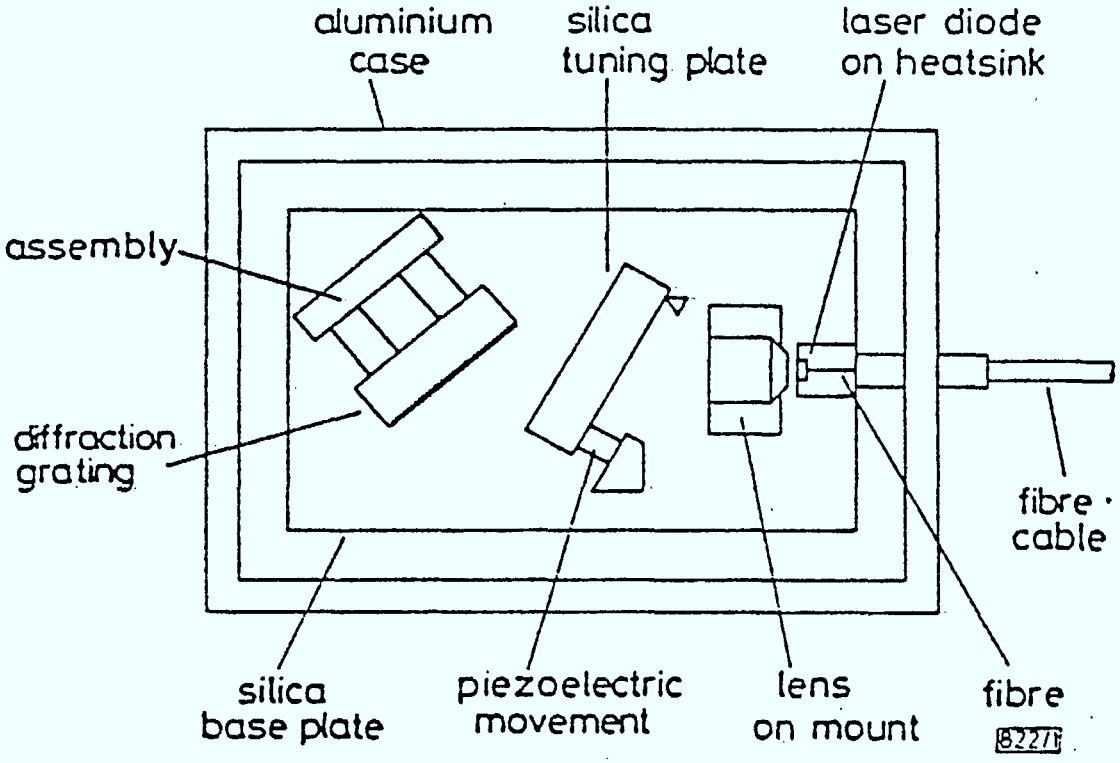


FIGURE 2.4.9: Schematic Diagram of the Grating-Tuned External Cavity Laser

lasers. Therefore channel selection can be performed by tuning either the transmitting laser or the local oscillator laser. FDM would provide access to a potential 22 THz of useful bandwidth in the .83 μm window.

Such developments clearly demonstrate the progress made in realizing the vast potential offered by optical heterodyne techniques in practical communication systems.

3.0 RECEIVER SYSTEM STUDY

In order that a receiver might be optimized as to its sensitivity, a great deal has to be known about its various components and the expected signal to be received. For example, the detector time constant, quantum efficiency, noise sources, the detector area, the pulse width, information format are all important parameters that will contribute to the sensitivity of the receiver.

It is beyond the scope of this report to give a comprehensive analysis of optical receivers. The reader is referred for instance to R.G. Smith and S.D. Personick[3.1] for a good general discussion of such receivers.

The derivation in that paper is appealing as it attempts to optimize the complete system in a general way. On the other hand, the optimization requires the design of a mathematically derived equalizer which might or might not be physically realizable. Also, the derivations of this paper are rather abstract and do not give a good insight into the contribution of each element to the total performance of the receiver. It is therefore difficult to optimize piece-meal, as is usually required in a real experiment when one is stuck with non optimum components such as the best detector which is available, etc. We shall therefore derive as much theory as required to be able to follow the signal retrieval process through the components of interest. We will then compare our results with Smith et al to see how close to optimum our system approaches.

3.1 Dectector Study

The heart of the receiver system is usually the detector itself. We shall therefore dwell most of our emphasis on this component of the receiver.

Two types of silicon avalanche photodiode detectors will be examined. One will be the RCA type C30950G which contains its own preamplifier as part of the chip module, the other is the ANTEL type which is strictly the detector chip and bias resistor . The advantage of the RCA detector is that the capacitance can be minimized and impedance mismatched between detector and transmission line can be removed by the buffer pre-amp. Unfortunately, the time constant so far is limited to about 1 nsec or so. Because their behaviour is now well characterized, these detectors merit our attention. On the other hand the ANTEL type can be very fast but their performance is inferior by necessity.

We shall derive the theory required to examine the behaviour of a pulse as it passes through the detectors and the accompanying noise.

3.1.1 Pulse Evolution Through the Detector RCA C30950G

Assume that the input to the detector is a rectangular laser pulse of duration τ_p. This pulse is transformed by the detector into an electrical pulse by the detector proper, which is transformed in turn by the module preamp into an output pulse.

The preamp which is part of the detector module has a frequency and impulse response given approximately by (3.2)

$$f(\omega) \approx \frac{1}{1 + \omega^2\tau^2_d}, \tag{3.1.1}$$

$$F(t) \approx \frac{t}{\tau^2_d} \exp(-t/\tau_d) \tag{3.1.2}$$

where τ_d is the detector time constant.

The normalized time development of the detector output signal pulse is thus given by

$$F_{sn}(t) = \int_0^t P_R(t').R_{di}\, F(t - t')dt', \qquad (3.1.3)$$

where $P_R(t)$ is the time dependence of the received laser pulse power (watt) and R_{di} is the current responsivity of the detector assembly. (Amps/watt)

By substituting the impulse function (3.1.2) into (3.1.3) and performing the integration, we obtain for a rectangular pulse P of duration τ_p.

$$F_1(t) = P\left[1 - \exp(-\tau/\tau_d)\left(\frac{t}{\tau_d} + 1\right)\right] \qquad (3.1.4a)$$

for $t \leq \tau_p$ and

$$F_2(t) = P\,\frac{\exp(-t/\tau_d)}{\tau_d}\{t(e^X-1)-\tau_d[e^X(x-1) + 1]\} \qquad (3.1.4b)$$

for $t \geq \tau_p$ with $x = \tau_p/\tau_d$

Figure 3.1.1 shows the time evolution of a rectangular pulse of duration τ_p as it passes through the detector. The output is shown normalized to the output value obtained for $\tau_p \to \infty$ and for P = 1. The output is plotted as a function of $y = t/\tau_d$ for a number of values of $x = \tau_p/\tau_d$. It is seen that if a detector is first compared to τ_p (i.e., $x >> 1$), the curves reach 90% of their maximum attainable values, when $t \gtrsim 4\tau_d$. On the other hand, when $x \lesssim 1$, the output never makes it to the asymptote and reaches its maximum at about a time equal to the time constant of the detector and, therefore, after the initiating pulse is terminated. The maximum can be found by differentiating Eq. (3.1.4) to obtain

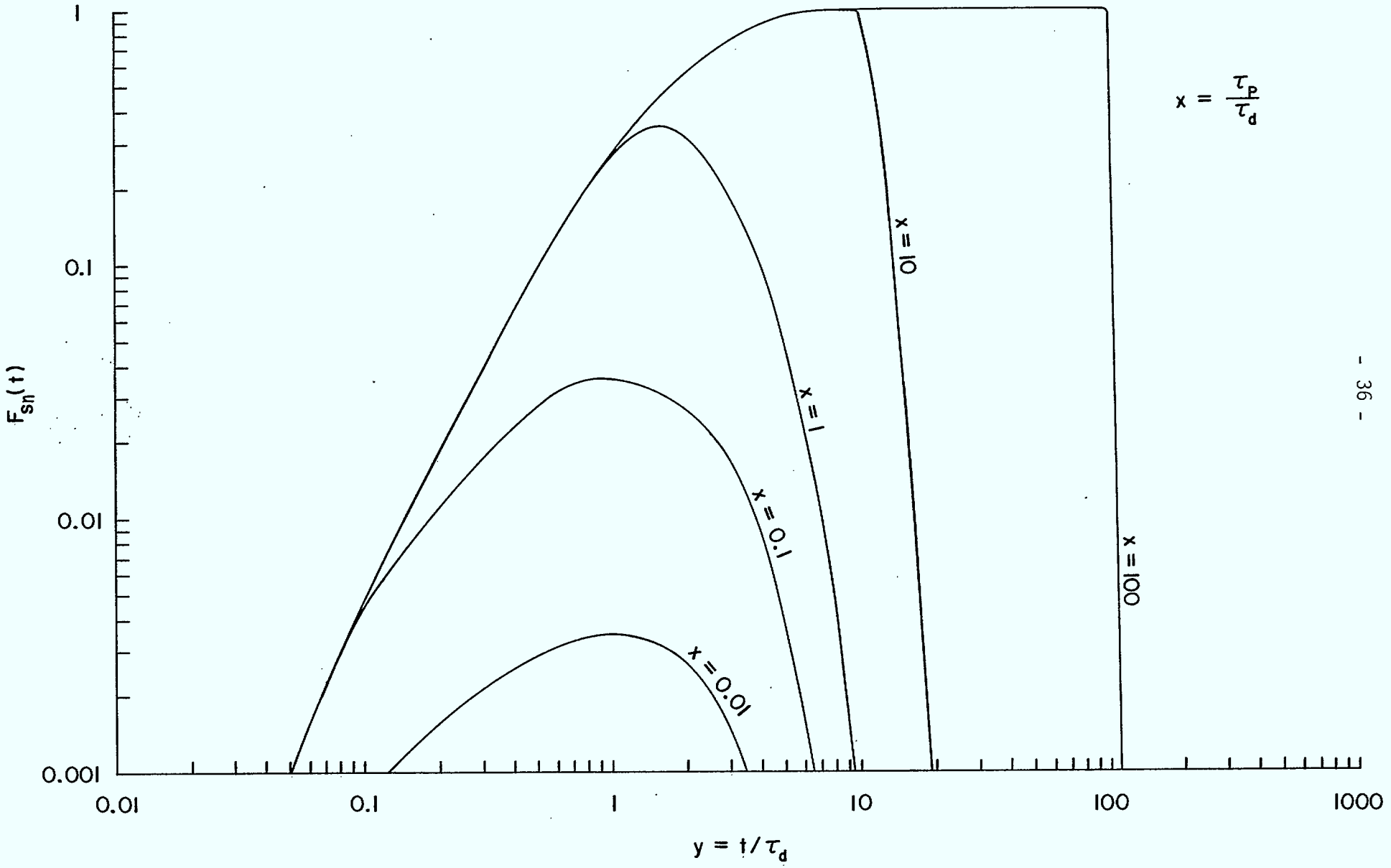


FIGURE 3.1.1: Time Evolution of a Rectangualr Input Pulse

$$F(t_m) = P[\exp(x) - 1] \exp\left(\frac{-x}{1 - \exp(-x)}\right) \tag{3.1.5}$$

This occurs at the optimum time

$$t_m = \frac{x\tau_d}{1 - \exp(-x)} \tag{3.1.6}$$

Figure 3.1.2 is a normalized plot of these maxima as a function of x. It is seen that the maximum available signal can be obtained by making $x \gtrsim 4$, i.e. keeping $\tau_d \lesssim 0.25\ \tau_p$.

To bring out the features of the pulse output as it pertains to pulse detection, we shall replot the time evolution on a linear basis for 3 values of x. This is shown in Figure 3.1.3. The absisca is divided in time slots of duration τ_p. If it is assumed that the sampling is done at the end of the time slot then it is seen that for x = 10, there is essentially no intermodulation at $t = 2\tau_p$. On the other hand, for x = 1, the signal from the previous pulse is larger than that generated at the second time slot (in the ratio of .33 to .27). Also, the actual maximum reached is only 35% of that when the response of the detector assembly is faster (x larger). For the case of x = .1 this effect is even worse. It is evident that strictly from the point of view of the signal alone, it is advantageous to have a detector with a time constant as fast as possible (x as large as possible). This is good both for the amplitude and intermodulation point of view. However, as we shall see, the larger x, the larger the bandwidth of the detector and hence the larger the noise admitted into the system. Evidently there should be a compromised detector time constant that would give the optimum error bit rate performance. In order to find it, we must obtain a functional relationship for noise. This will now be discussed.

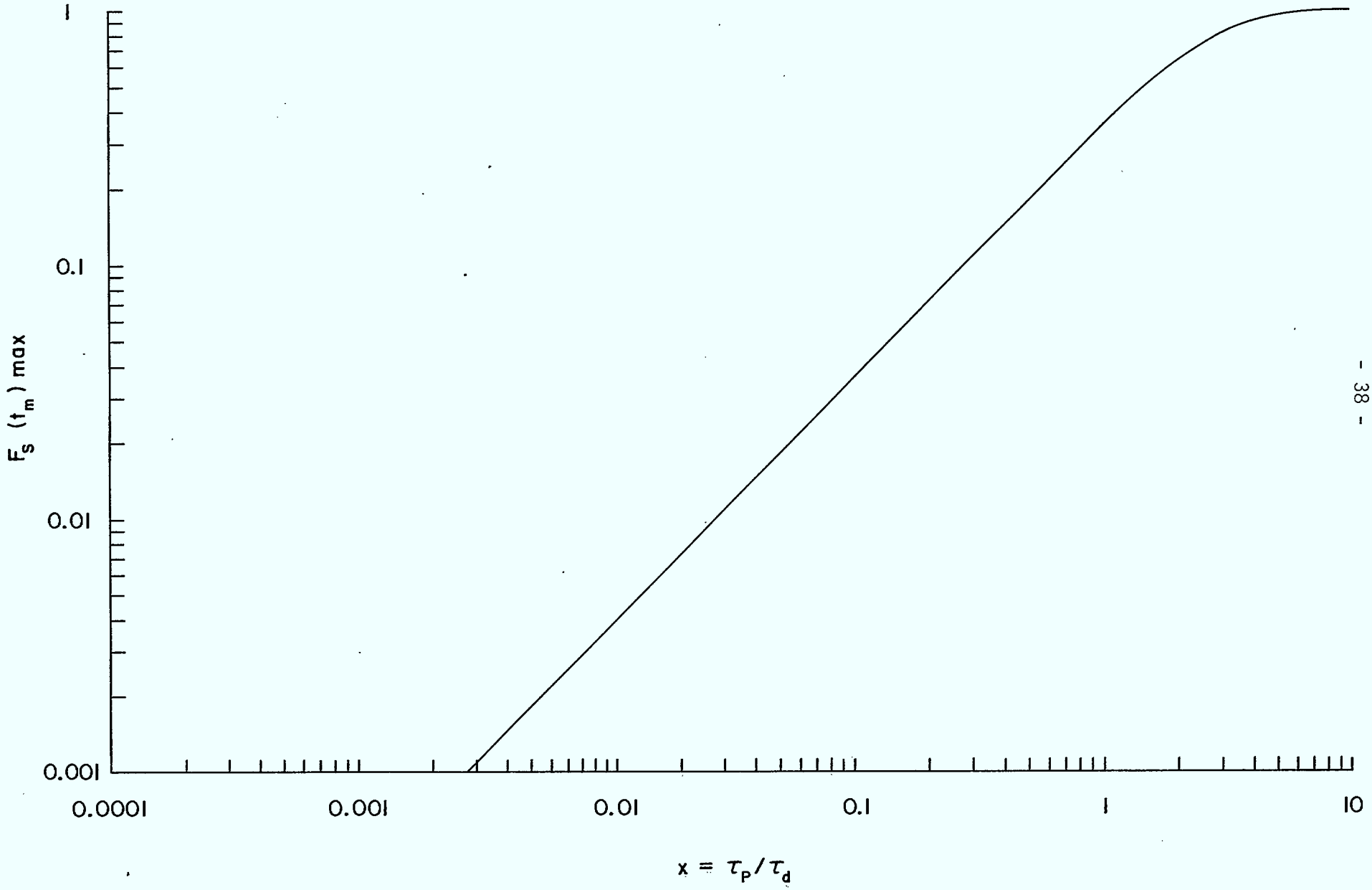


FIGURE 3.1.2: Maximum Output as a Function of x

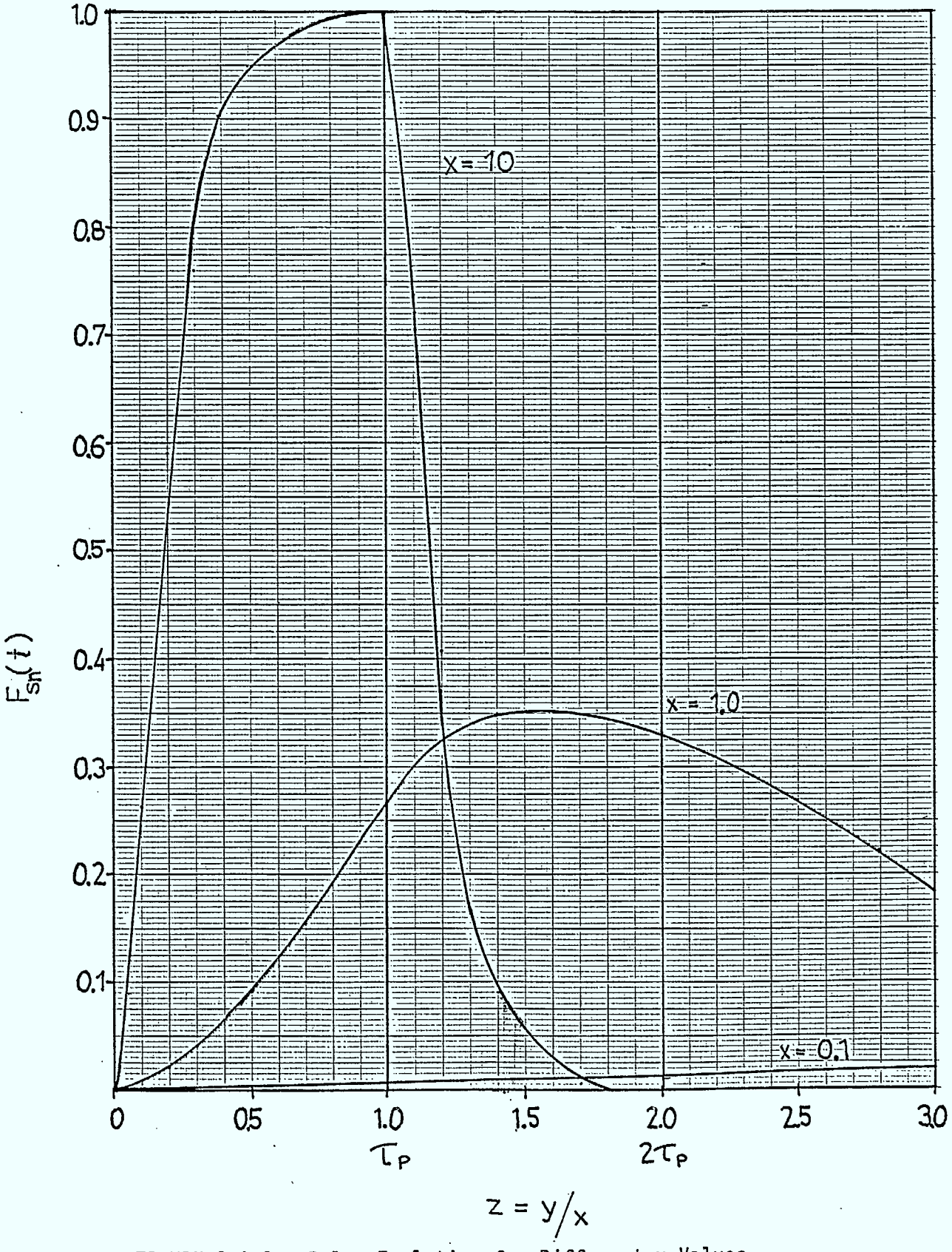


FIGURE 3.1.3: Pulse Evolution for Different x Values

3.2 Noise Study

When an optical detector is used to detect optical pulses, the noise output can be split into two major categories. One category is time independent and is produced by the internal mechanism of the detector. This includes dark current noise, gain noise, preamplifier noise, etc. The other category is time dependent: the main contribution coming from shot noise from the signal itself. It is not clear from Smith et al theory whether that noise contribution comes mostly from the pulse being sampled or whether a substantial amount is contributed by the past pulses. This is because in the theory formalism, all the past pulse terms appear. We shall therefore derive the noise contribution for one pulse for a realistic case and determine its influence on subsequent pulses.

3.2.1 Time dependence of Shot Noise

The mean square of the noise current at the output of the detector module is given by[3.2]

$$\dot{I}_n^2(t) = qM^2E\left(\frac{R_d}{R_L}\right) \int_{-\infty}^{\infty} [R_o.p(t' - t_o)].F^2(t - t')dt' \qquad (3.2.1)$$

where q is the electronic charge

M the mean current gain

$E = k_e M + (1 - k_e)(2 - \frac{1}{M})$ is the excess noise factor for an avalanche diode

k_e = a constant

R_o = the unity gain responsivity (Amp/watt)

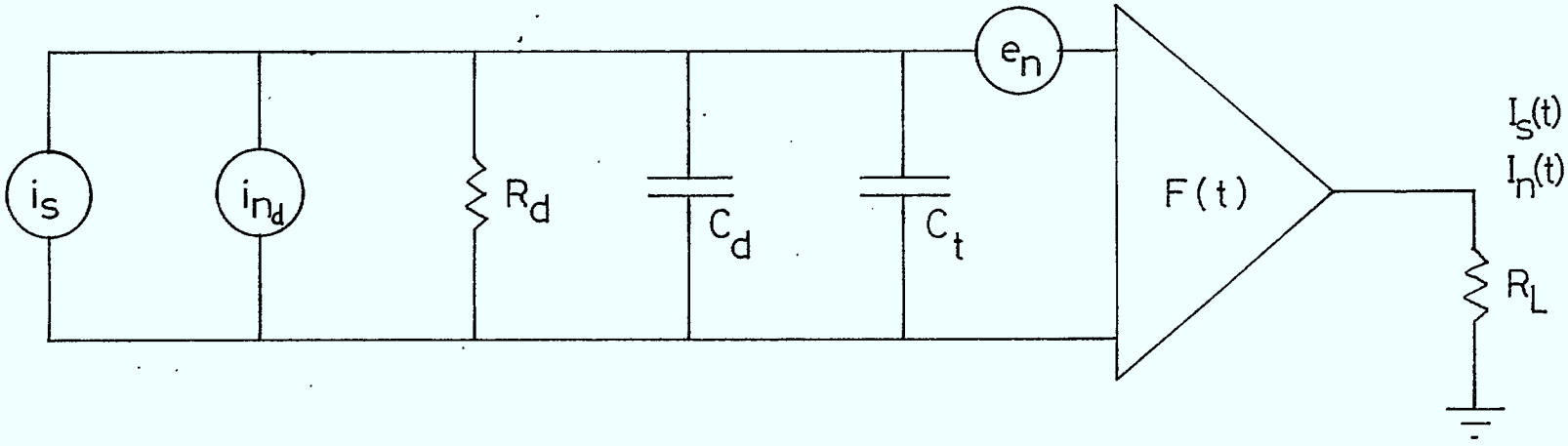
p(t) = the time dependence of the received input pulse (watt)

F(t) = impulse response of the detector module

R_d is the load resistance of the detector (see Figure 3.1.4)

R_L is the load resistance of the output circuit (see Figure 3.1.4)

We can substitute into 3.2.1 the pulse of duration τ_p starting at $t_o = 0$, and the impulse function given in 3.2.1 and doing the required integration.



where i_{n_d} = detector noise before the preamplifier

FIGURE 3.1.4: Detector Equivalent Circuit

This gives

a) for $t > \tau_p$

$$I_n^2(t) = \frac{qM^2E\,R_oP}{2\tau_d}\left(\frac{R_d}{R_L}\right)^2 e^{-2y}\{y^2(e^{2x}-1)-y[e^{2x}(2x-1)+1] + \frac{1}{2}[e^{2x}(2x^2-2x+1)-1]\} \quad (3.2.2)$$

b) for $t < \tau_p$

$$I_n^2(t) = \frac{qM^2E\,R_oP}{2\tau_d}\left(\frac{R_d}{R_L}\right)^2 \{\frac{1}{2} - e^{-2y}[\frac{1}{2}+y^2+y]\} \quad (3.2.3)$$

where P is the magnitude of the optical pulse (Watts)

$y = t/\tau_d$

$x = \tau_p/\tau_d$

3.2.2 Total Noise of Detector Module

In addition to the time dependent noise term, the detector will also generate a steady state noise term given by[3.3]

$$I_n^2 = \frac{1}{2\pi}\left(\frac{R_d}{R_L}\right)^2 \int_0^\infty (S_0 + \omega^2 S_2)\cdot f^2(\omega)\,d\omega \quad (3.2.4)$$

where[3.3] the noise power spectrum (Watts/Hz) of the detector is

$$S_0 = 2q[I_{ds} + I_t + (i_{db} + i_b)M^2E] + \frac{4kT}{R_d}, \quad (3.2.5)$$

with q = electronic charge (Coulomb)

I_{ds} = unmultiplied portion of the detector dark current (Amp)

I_t = effective gate leakage current of the preamp transistor (Amp)

M = multiplication gain

$E = k_e M + (1 - k_e)(2 - 1/M)$ is the excess noise factor

k_e = a constant

R_d = load resistance of detector

R_L = output load of the circuit

i_b = average detector current due to background radiation power

i_{db} = multiplied portion of dark current

k, t have the usual significance and $f(\omega)$ was defined in Eq. (3.2.1)

The other term contributed by the transistor preamp is

$$S_2 = e_n^2(C_d + C_t)^2, \tag{3.2.6}$$

where e_n = equivalent transistor noise voltage

C_d = detector capacitance, and

C_t = effective transistor input capacitance.

The integration of Eq. (3.2.4) gives

$$I_n^2 = \frac{1}{8\tau^3_d}(S_0\tau^2_d + S_2)(R_d/R_L)^2 \tag{3.2.7}$$

which is the average current noise power.

3.2.3 Sample Calculations

In order to have a feel for how the detector parameter will affect the system performance we shall examine a specific case based on the RCA C30950G detector module.

The performance of a communication system is usually given in terms of the bit-error-rate. We start by noting that the probability of

error for a two level system is given approximately by

$$P_T = \frac{1}{4}\left(\operatorname{erfc}\left(\frac{D - A}{\sqrt{2}\sigma_A}\right) + \operatorname{erfc}\left(\frac{B - D}{\sqrt{2}\sigma_B}\right)\right) \tag{3.2.8}$$

where erfc = 1 - erf

erf = the error function

A and B are the two signal levels (B > A) at the sampling time

σ_A, σ_B are the r.m.s. noise occurring at level A and B respectively

D is the optimum decision level.

It is assumed here that both signal A and B are equally likely and that the signal-to-noise ratios are not too small. For simplicity in this discussion, we assume that $\sigma_A \approx \sigma_B$ which implies that the system is not signal shot noise limited. Under this simplification, the probability of error is simply

$$P_T = \frac{1}{2}\operatorname{erfc}(Q/\sqrt{2}) \tag{3.2.9}$$

where $Q = \dfrac{D - F_{s_{min}}}{\sigma} = \dfrac{F_{s_{max}} - D}{\sigma}$

and A and B have been replaced by the two possible signal levels $F_{s_{min}}$ and $F_{s_{max}}$ occuring at the sampling time τ_p and $2\tau_p$ obtained from Eqn. 3.1.4a and b.

Solving for Q by eliminating D, we get

$$Q^2 = \frac{\left(F_{s_{max}} - F_{s_{min}}\right)^2}{4\sigma^2} = \frac{1}{4}\left(\frac{S_{pp}}{N}\right) \tag{3.2.10}$$

where S_{pp}/N is the peak-to-peak signal to noise power ratio.

Determining the performance of the system requires therefore the evaluation of S_{pp}/N which we shall proceed to do.

Let us try to determine the detector time response in a semi-quantitative fashion. Suppose we wish the system to have a BER of 1 part in 10^9 if there were no intersymbol interference. Then the question that can be asked is what is the largest intersymbol interference we could tolerate that would at most change the BER from 1 to 2 parts in 10^9? By solving this problem, it can be shown that this requires that

$$F_s(2\tau_p) \lesssim .01$$

Solving from Eqn. (3.1.4b) we get

$$x \gtrsim 6.6 = \frac{\tau_p}{\tau_d}$$

or the detector should be about 6 times faster than the pulse width.

A plot of two such pulses inserted into two adjacent slots, is shown in Fig. 3.2.1 where the normalized signal is plotted against units of τ_p. It is seen that such a signal would be close to ideal since at $z = 2$, the sampling time, the signal would rise close to its maximum value while the past signal is decreased close to zero at that point.

Next, it would be interesting to calculate the time development of the shot noise for such a detector and to determine its relation with respect to the steady state noise. By using Eqns. (3.1.4a,b, 3.2.2, 3.2.3, 3.2.7) and using the following typical values for the constants

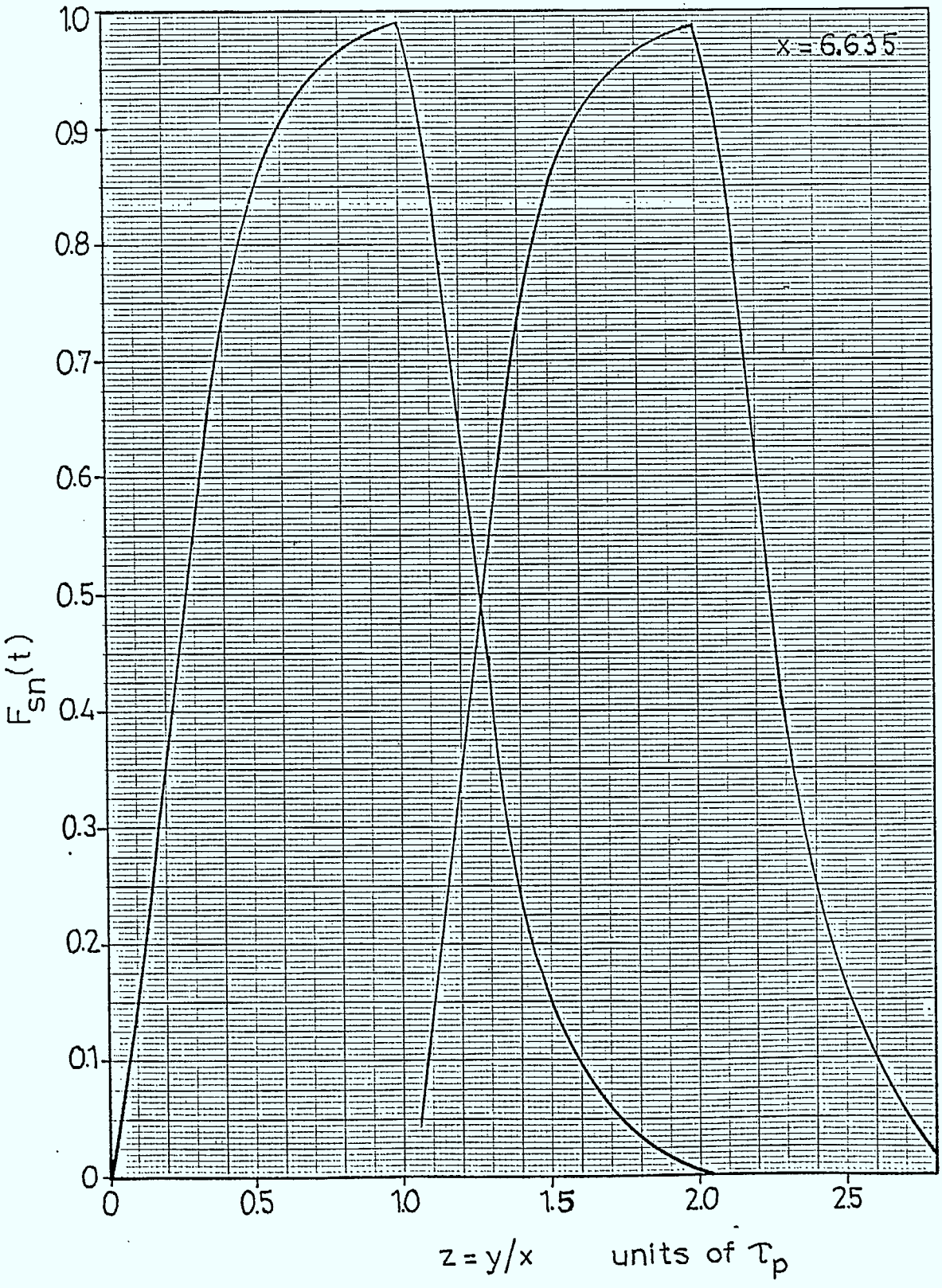


FIGURE 3.2.1: Signal Time Evolution

$$
\begin{aligned}
I_{ds} &= 300 \text{ nA} \\
I_{db} &= .4 \text{ nA} \\
I_t &= 3 \ \mu\text{A} \\
e_n &= 2 \text{ nV/Hz}^{1/2} \\
C_d &= 10 \text{ pF} \\
C_t &= 2 \text{ pF} \\
T &= 293 \text{ K} \\
k_e &= .02 \\
R_o &= .4 \text{ A/W} \\
M &= 60 \\
\tau_d &= 10 \text{ ns} \\
\tau_p &= 66.35 \text{ ns}
\end{aligned}
\qquad (3.2.11)
$$

we can calculate the time evolution of the signal noise $I_n^2(t)$ and steady state noise I_n^2 and S/N. The result is plotted in Fig. 3.2.2. It is seen that the time dependent noise term does push through above the steady state noise level at about y = 2, i.e. at a time equal to twice the detector time constant or 20 nsec. It stays above this level during the sampling time ($t = \tau_p$). In fact, the loss in S/N due to the time dependent noise term is about 4 dB. In this particular case however, because the S/N is high ($\sim$ 22 dB), the BER is less than 1 part in 10^{20} and this loss is therefore inconsequential. By the time the second sampling time ($t = 2\tau_p$) is reached, that noise term is completely inconsequential. At that time only the steady state noise is important.

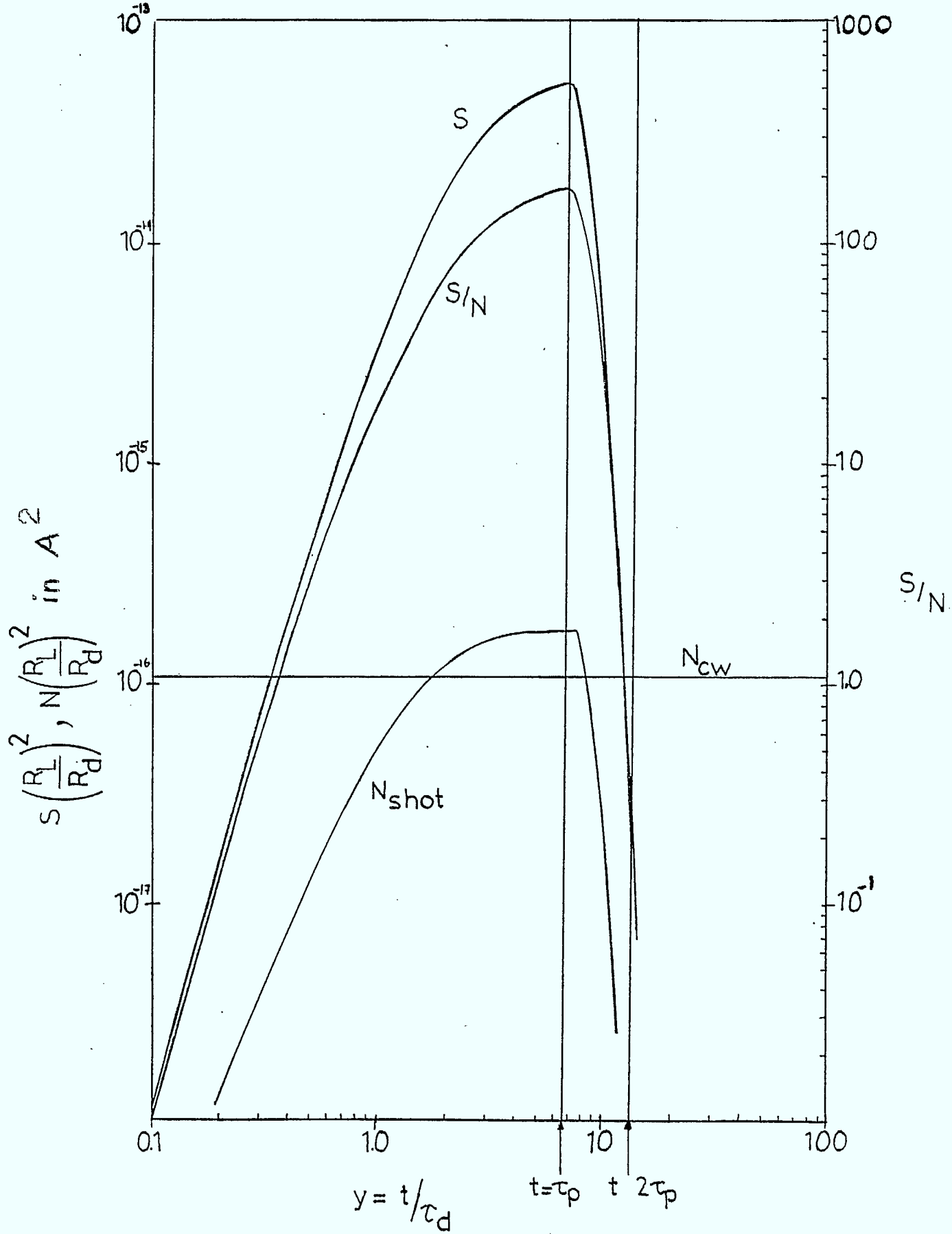


FIGURE 3.2.2: Signal, Noise and S/N Time Evolution

3.3 Computer Optimization

The previous calculation is a useful guide in showing the interaction between a fast detector that provides quick rise and fall of pulse with little intersymbol interference and time dependent noise which does rise above the other noise sources and thus limits the ultimate S/N. The calculation was carried out for a gain of M = 60. Different gain values would produce different proportions of noise as it affects both the steady state and time dependent noise in different ways. It would be desirable also, to carry out the analysis without simplification such as assuming that the noise of the high and low signals are the same and to make use of the time development of noise we have derived. The only practical way to optimize the detector module is therefore to use the computer to maximize the S_{pp}/N by determing the optimized value of τ_d and M at the sampling time.

3.3.1 Case 1 - Optimization of RCA Preamp Type Detector - Assumption 1

Assumption 1 is that at the sampling time, the previous signal was on if the present signal is off and vice versa, i.e. the previous signal was off when the present signal is on.

It can be shown that under these conditions and for a S/N which is not too low and for Gaussian noise the decision level D can be found by solving

$$\frac{i_h - D}{\sigma_h} = Q \tag{3.3.1}$$

$$\frac{D - i_\ell}{\sigma_\ell} = Q \tag{3.3.2}$$

where i_h, i_ℓ are the signal current for on and off state respectively and σ_h, σ_ℓ are the r.m.s. noise corresponding to the on and off state respectively.

With Assumption 1, this entails

$$i_h = i(\tau_p) = R_o MP(\frac{R_d}{R_L})F_1(\tau_p) \tag{3.3.4}$$

$$i_\ell = i(2\tau_p) = R_o MP(\frac{R_d}{R_L})F_2(2\tau_p) \tag{3.3.5}$$

$$\sigma_h^2 = i_n^2 + I_n^2(\tau_p) \tag{3.3.6}$$

$$\sigma_\ell^2 = i_n^2 + I_n^2(2\tau_p) \tag{3.3.7}$$

where F_1, F_2, i_n and I_n have been defined in Eqn. (3.1.4a and b, 3.2.4, 3.2.1).

Eliminating D and solving for Q^2 we get

$$Q^2 = \frac{(i_h - i_\ell)^2}{(\sigma_\ell + \sigma_h)^2} = \frac{S_{pp}}{N_T} \tag{3.3.8}$$

From this definition, S_{pp}/N_T is thus the peak-to-peak signal power to effective noise power σ_T^2

where $\sigma_T = \sigma_\ell + \sigma_h$.

A computer algorithm is then set up to maximize the S_{pp}/N_T by optimizing the detector time constant and the gain. The plotted results are shown in Figs. 3.3.1 to 3.3.4. Figure 3.3.1 shows the optimum S_{pp}/N_T as a function of P_R the received pulse power at the

detector where $N_T = \sigma_T^2$. This is done for a number of pulse widths τ_p (or bit rate). The detector parameter values are the same as those selected in 3.2.11 and R_d = 15 KΩ and R_L = 50 Ω. We observe that for larger τ_p, the S_{pp}/N_T is greater. This is to be expected since a narrower bandwidth can be used in the receiver with a reduction in noise. A line is drawn to show the level at which the bit-error-rate is 1 part in 10^9. This graph therefore reveals the amount of received power P_R which is required at the bit-error-rate for various bit rates (pulse widths). Figure 3.3.2 shows the optimum gain required as a function of P_R for each pulse width (bit rate). It is seen that in general, the lower the received power and/or the higher the bit rate the higher is the gain.

Figure 3.3.3 gives the optimized detector time constant as a function of P_R for various bit rates. It is found that in general, the higher the bit rate the lower τ_d is expected. Also, except for the very high bit rate (10 Gbit), the lower the received power the higher the required detector time constant, although the change is relatively small.

To see the relation of optimum τ_d to τ_p, we have plotted these parameters in Fig. 3.3.4. It is seen that at low P_R, optimum τ_d is essentially proportional to τ_p. In fact $\tau_{d_{opt}} \approx 1/4\tau_p$.

3.3.2 Case 2 - Optimization of RCA Preamp Type Detector - Assumption 2

Assumption 2 is that it is now assumed that at sampling time, previous signal was always on whether present signal is on or off. This entails now that the intersymbol interference is always present, so is the noise from the previous pulse. This is therefore slightly more pessimistic than the previous case. We shall, nevertheless, use this assumption from here on so as to be able to compare our results with those of Smith & Personick who made that assumption.

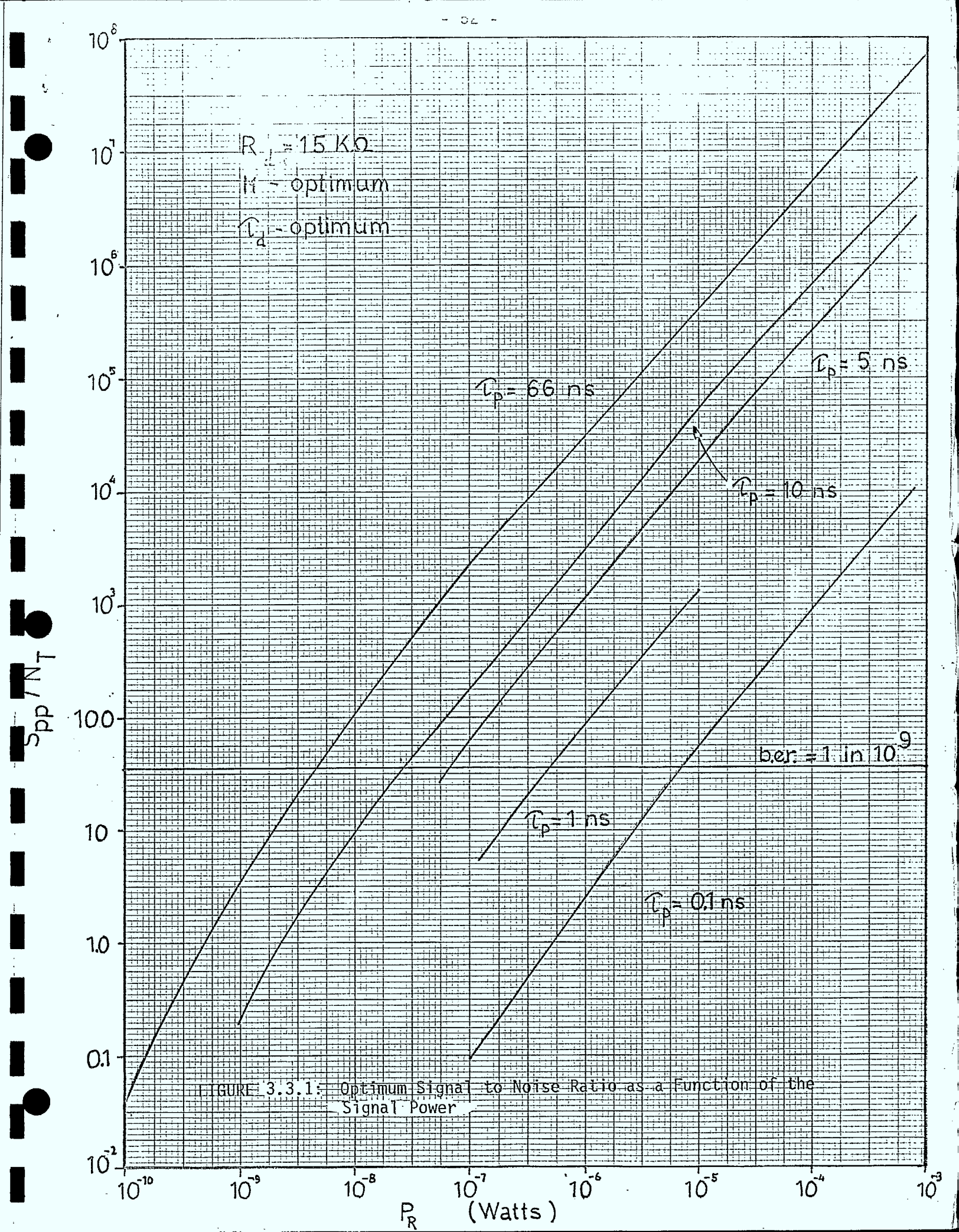


FIGURE 3.3.1: Optimum Signal to Noise Ratio as a Function of the Signal Power

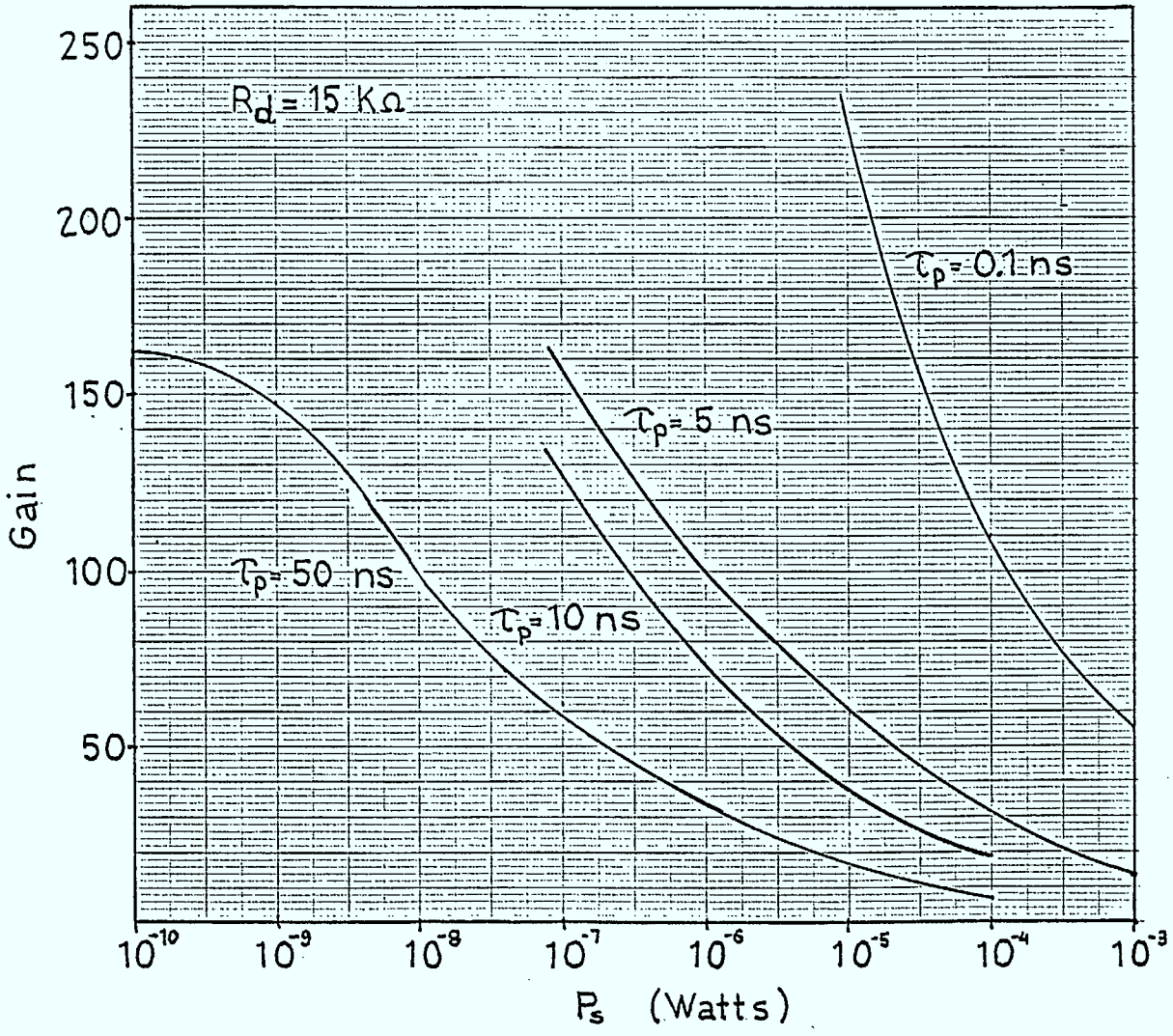


FIGURE 3.3.2: Optimum Gain as a Function of the Signal Power

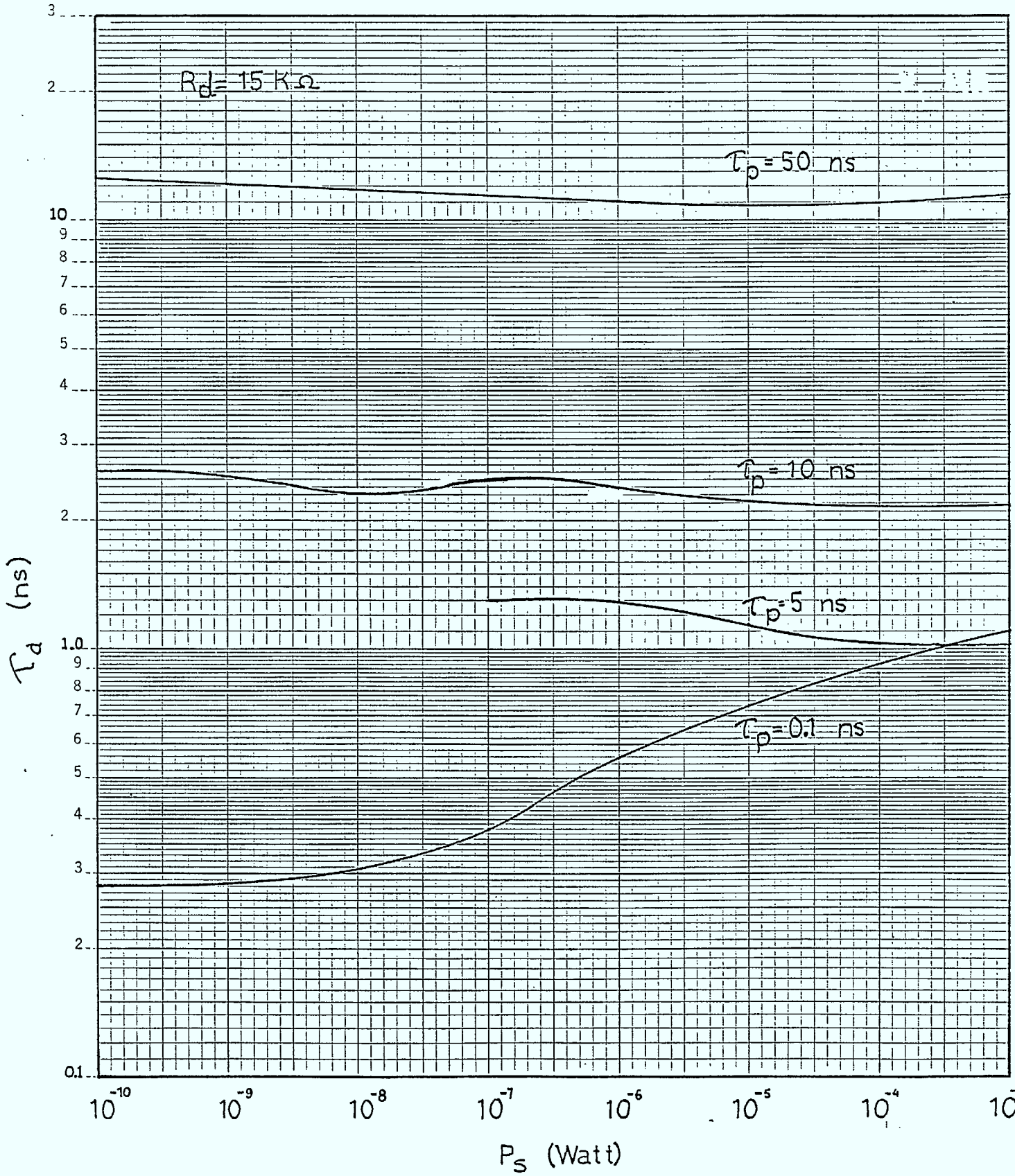


FIGURE 3.3.3: Optimum Detector Time Constant as a Function of the Signal Power

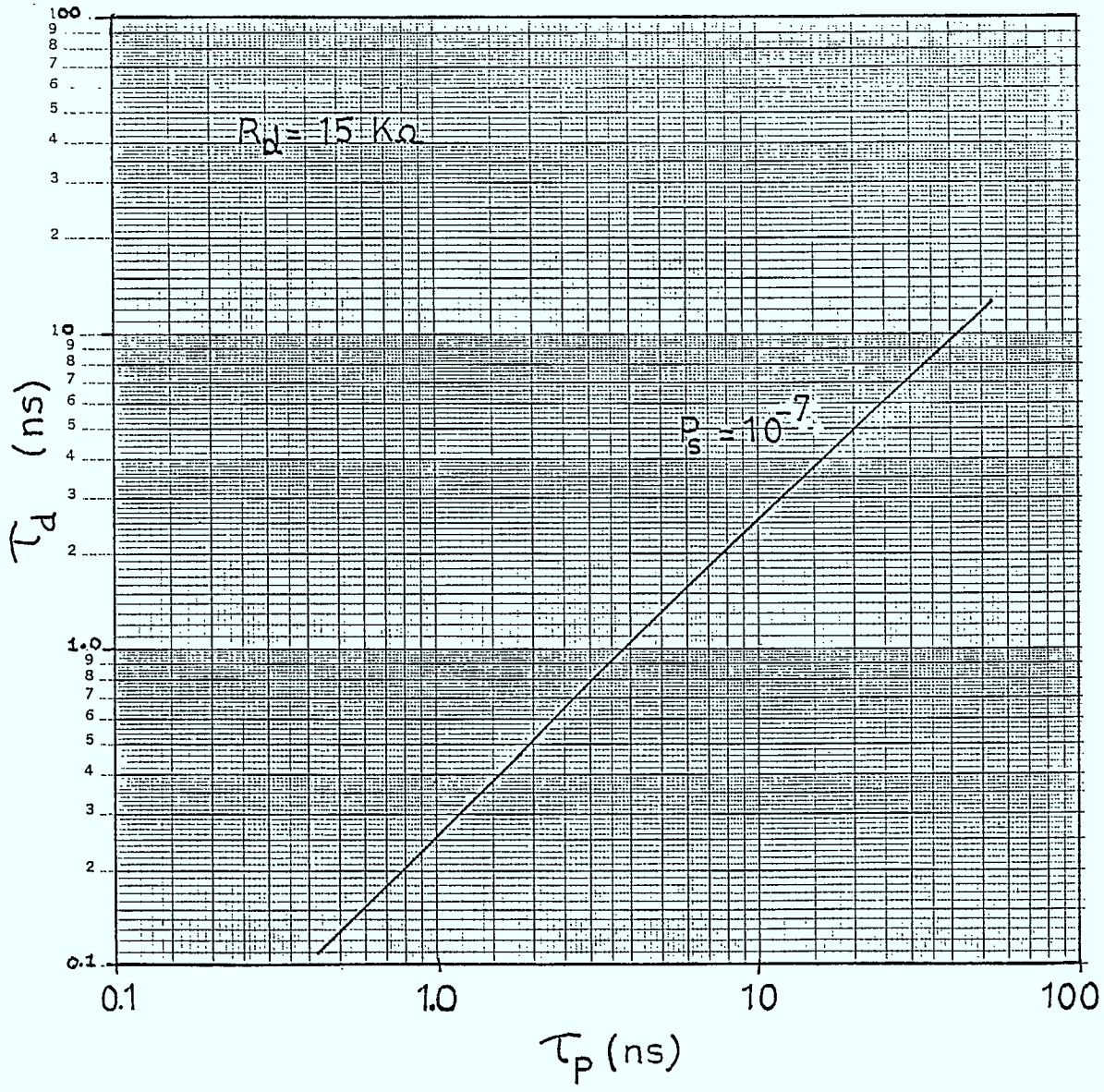


FIGURE 3.3.4: Optimum Detector Time Constant as a Function of τ_p

It is easy to show that under this assumption

$$Q = \frac{i_h}{\sigma_{\ell+h} + \sigma_\ell} \tag{3.3.9}$$

where $\sigma_{\ell+h}$ is the r.m.s. noise caused by the present plus past pulse at the present sampling time (τ_p).

So that

$$Q^2 = \frac{i_h^2}{(\sigma_{\ell+h} + \sigma_\ell)^2} = \frac{S_{pp}}{N_T} \tag{3.3.10}$$

Note that the definition of S_{pp}/N_T is now slightly different from (3.3.8).

In order to make our model a little more realistic at higher bit rates, we shall further assume that R_d is not a constant for all τ_d but is given by

$$R_d = 15\ k\Omega \times \frac{\tau_d}{6.82\ nsec} \tag{3.3.11}$$

down to 50 Ω.

This will conform approximately with a number of values found from the manufacturer for various values of τ_d.

Using the same parameters as in Eqn. (3.2.11) we get the results plotted in Figures 3.3.5 to 3.3.8. The results are substantially similar to the previous ones even though the S_{pp}/N_T is slightly less and the optimum τ_d slightly larger. The main advantage of using Assumption 2 however is that we can compare our results with those of Smith & Personick. In order to do that, however, we have to put our results in their form.

Figure 3.3.9 shows the results that Smith & Personick obtained using their calculations where the plot is in $\eta\bar{P}$ vs. bit rate. Here η is the quantum efficiency while $\bar{P}$ = average power of pulse for on and off

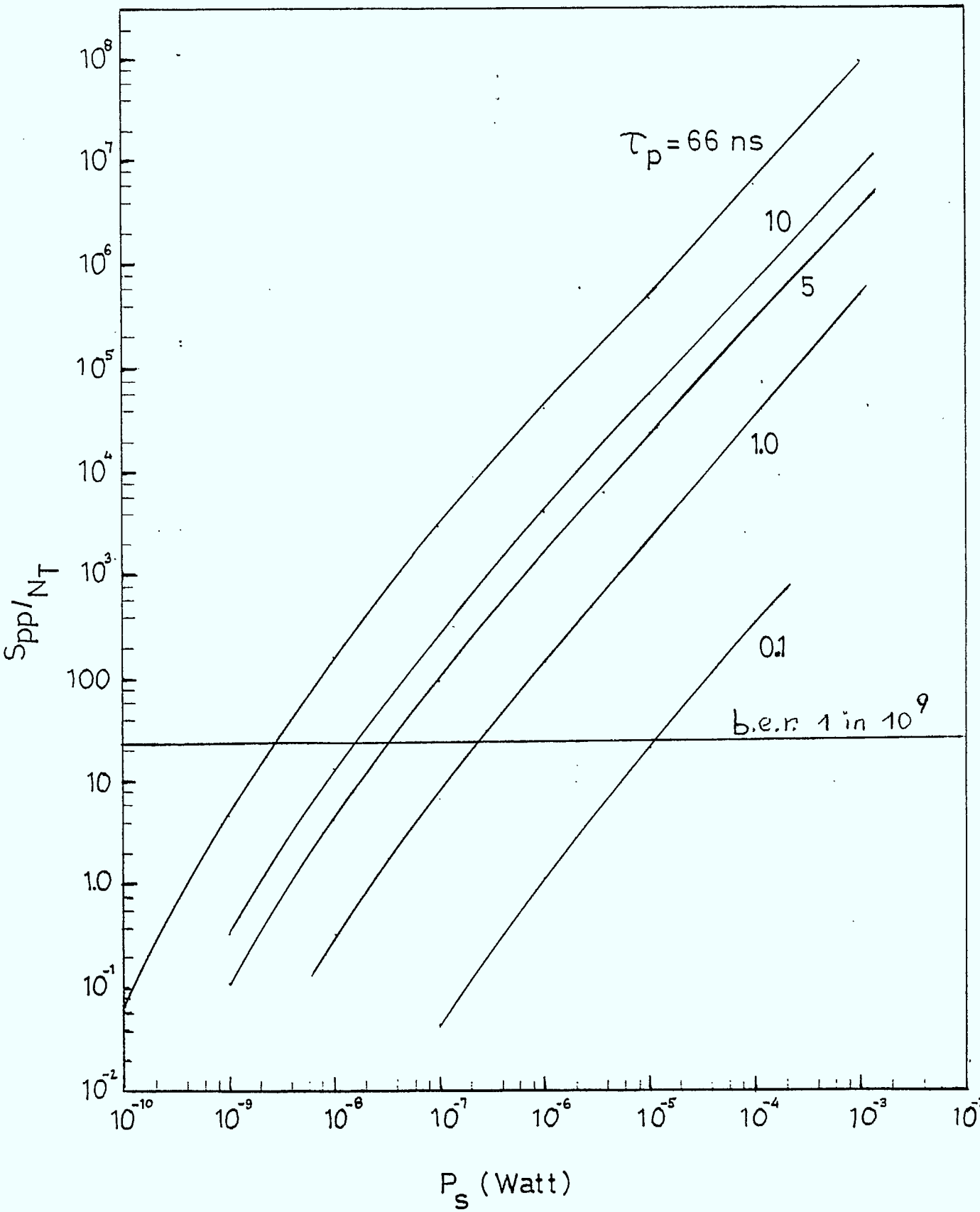


FIGURE 3.3.5: Optimum Signal to Noise Ratio as a Function of the Signal Power

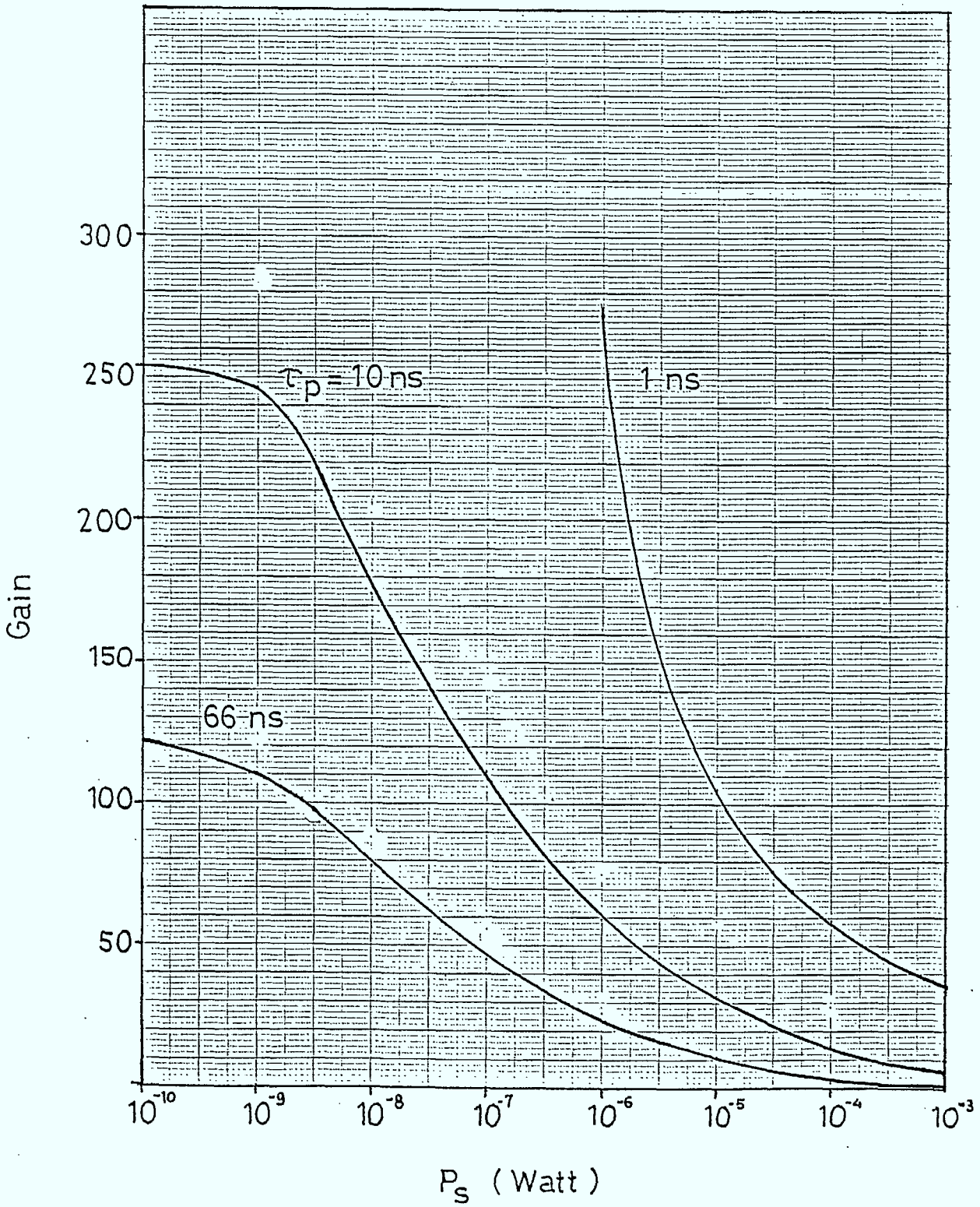


FIGURE 3.3.6: Optimum Gain as a Function of the Signal Power

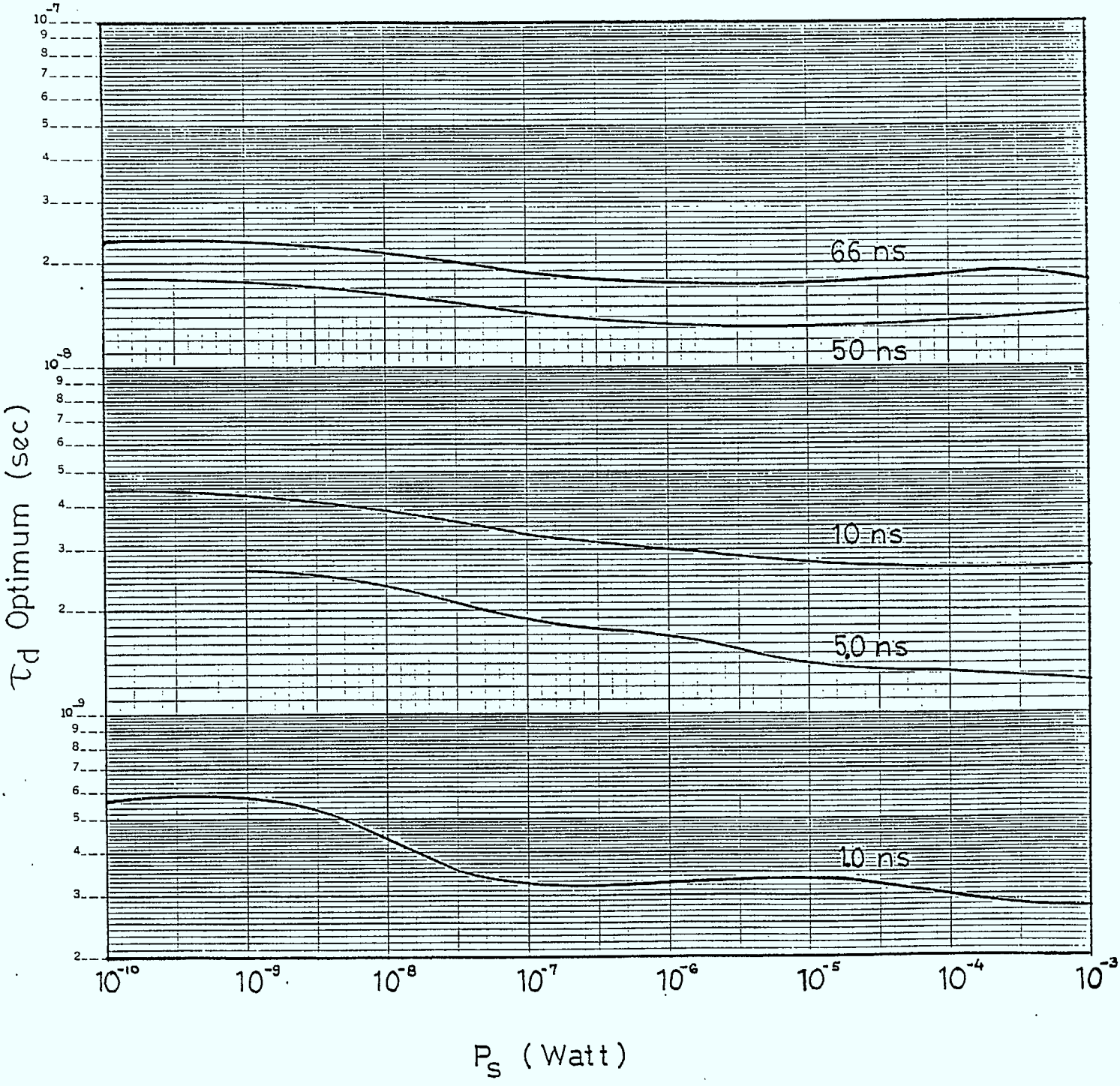


FIGURE 3.3.7: Optimum Detector Time Constant as a Function of the Signal Power

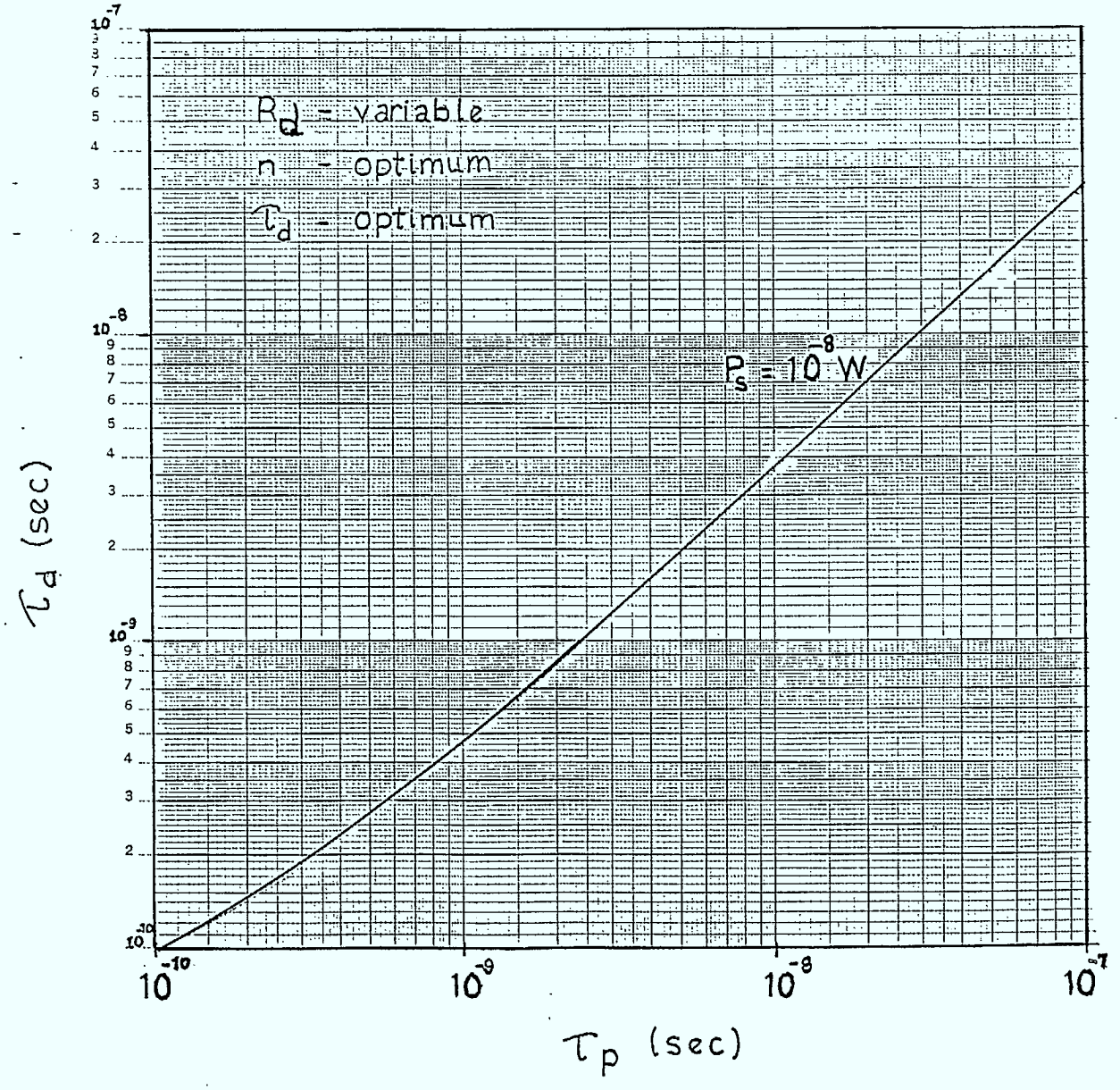


FIGURE 3.3.8: Optimum Detector Time Constant as a Function of τ_p

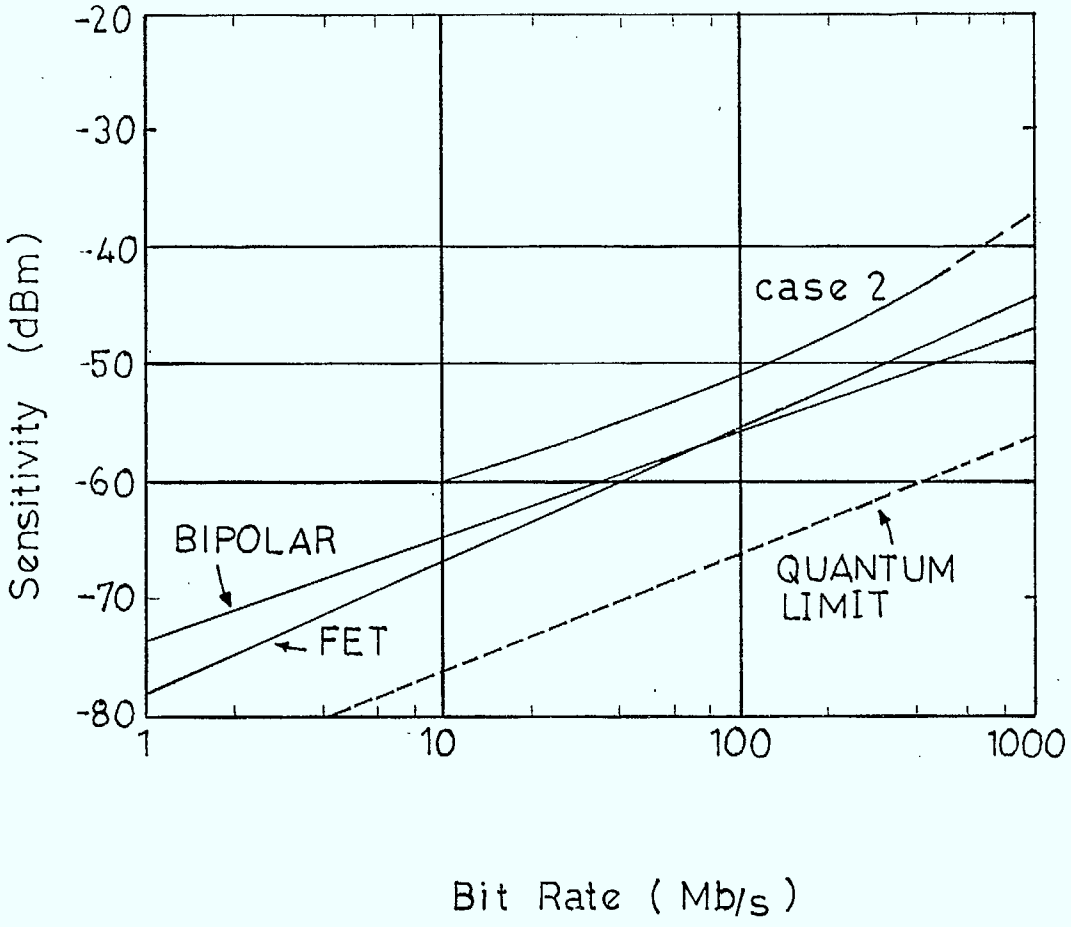


FIGURE 3.3.9: System Performance as a Function of the Bit Rate

pulses. Thus $\bar{P} = P_R/2$.

To conform to the values they have chosen, we further make the following slight modification to our parameters:

$$R_o = .6 \text{ A/W}$$

$$\eta = .9$$

$$\text{and } k_e = .03$$

Reworking the optimization, our results can then be plotted on that same graph as shown in Fig. 3.3.9. It is seen that our result is only about 3 - 4 dB above theirs. To account for the difference, it should be noted that we calculated the exact noise that includes both the time independent noise plus the time dependent noise while Smith & Personick made the assumption that the time dependent noise is negligible. If we look at Fig. 3.2.2 however, we see that the time dependent shot noise could go as high as 3 dB above the time independent noise for that case. Other cause for the slight discrepancy is that Smith & Personick assumed the use of an equalizer. Also, some detector parameters not explicity given in this paper, might be slightly different from ours. Note also that such a module is not available at present for bit rates above 500 M bit/sec. The results are thus shown in dotted line in that region.

3.3.3 Case 3 - Optimization of Fast Detectors Without Preamplifiers

It is generally desirable to have a preamplifier built as an integral part of a detector module. This allows better matching of impedances and reduction of stray capacitances. As the bandwidth requirement increases, however, the effective load resistance has to decrease more and more until it reaches 50 Ω or lower. Under these conditions, transmission line techniques are advantageous. In fact, all fast detectors, with rise time

$\lesssim$ 1 nsec are usually supplied without preamplifiers. We therefore have to derive theoretical models for such detectors that have to be connected directly to commercial low-noise preamplifiers.

Using the model shown in Fig. 3.3.10 for the detector-amplifier combination, it can be shown that the impulse function is

$$F(t) = A \frac{R_d}{R_L} \left(\frac{1}{\tau_d - \tau_a} \right) \left(e^{-t/\tau_d} - e^{-t/\tau_a} \right) \tag{3.3.12}$$

where A = Voltage amplification of amplifier (at low frequency)

$\tau_d = R_d \cdot C_T$ = time constant of detector

τ_a = Time constant of amplifier

C_T = Total input capacitance to amplifier.

The time development of the output signal is thus

for $t < \tau_p$

$$i_{s_{out}}(t) = F_s(t) = R_o MPA \frac{R_d}{R_L} \left(\frac{1}{\tau_d - \tau_a} \right) \left\{ \tau_d \left[1 - e^{-t/\tau_d} \right] - \tau_a \left[1 - e^{-t/\tau_a} \right] \right\} \tag{3.3.13}$$

and for $t > \tau_p$

$$i_{s_{out}}(t) = F_s(t) = R_o MPA \frac{R_d}{R_L} \left(\frac{1}{\tau_d - \tau_a} \right) \left\{ \tau_d e^{-t/\tau_d} \left[e^{\tau_p/\tau_d} - 1 \right] - \tau_a e^{-t/\tau_a} \left[e^{\tau_p/\tau_a} - 1 \right] \right\} \tag{3.3.14}$$

The noise model is as shown in Fig. 3.3.10.

The time development of the shot noise can now be shown in this case to be[3.4]

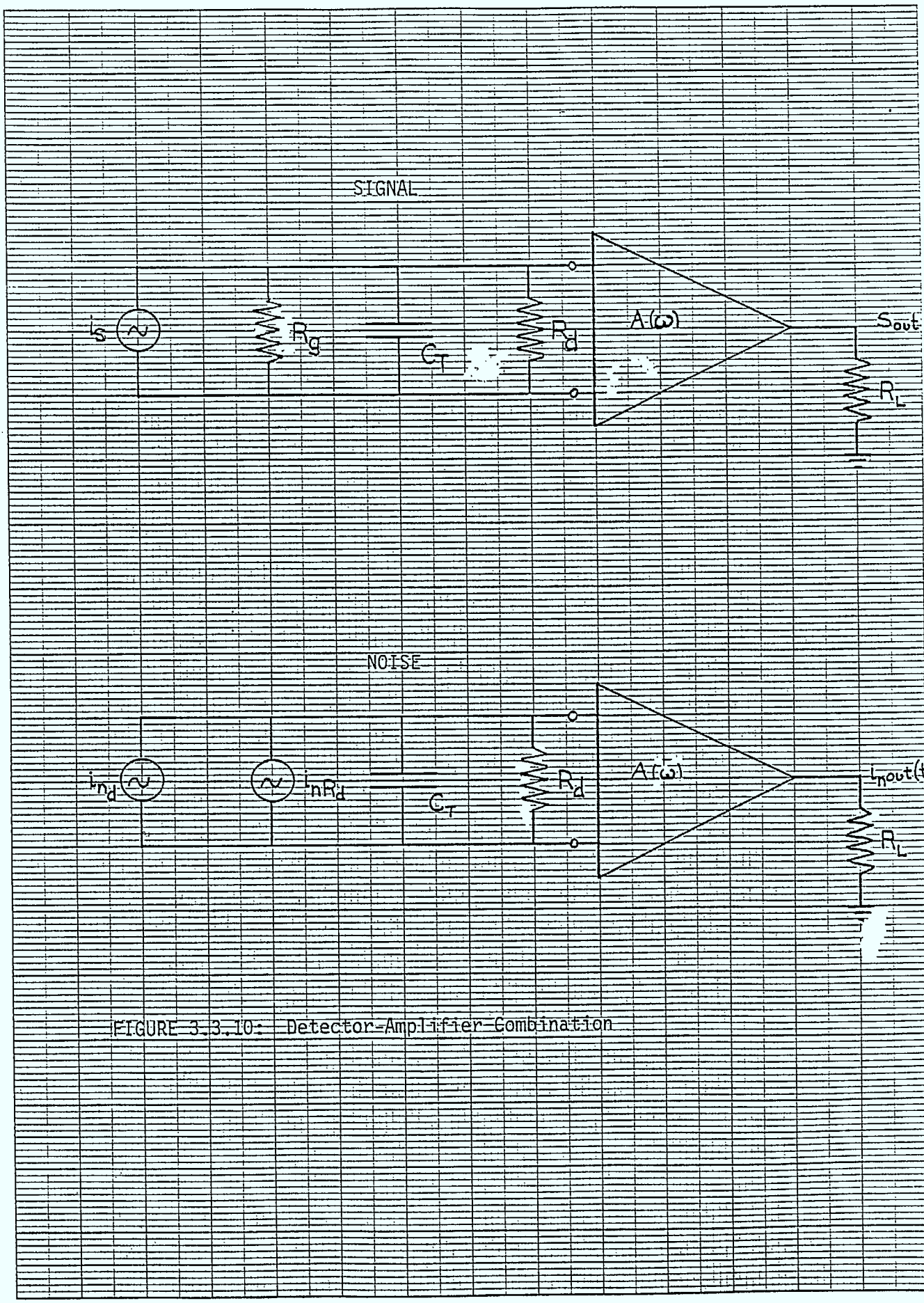


FIGURE 3.3.10: Detector-Amplifier-Combination

For $t < \tau_p$

$$i^2_{n_{out}}(t) = qM^2 E R_o P \left(A \frac{R_d}{R_L}\right)^2 \left(\frac{1}{\tau_d - \tau_a}\right)^2 \cdot$$

$$\left\{ \frac{\tau_d}{2} \left[1 - e^{-2t/\tau_d}\right] + \frac{\tau_a}{2} \left[1 - e^{-2t/\tau_a}\right] - 2\tau_e \left[1 - e^{-t/\tau_e}\right] \right\} \qquad (3.3.15)$$

where $\frac{1}{\tau_e} = (\frac{1}{\tau_d} + \frac{1}{\tau_a})$.

While for $t > \tau_p$

$$i^2_{n_{out}}(t) = qM^2 E R_o P \left(A \frac{R_d}{R_L}\right)^2 \cdot \left(\frac{1}{\tau_d - \tau_a}\right)^2 \cdot$$

$$\left\{ \frac{\tau_d}{2} e^{-2t/\tau_d} \left[e^{2\tau_p/\tau_d} - 1\right] + \frac{\tau_a}{2} e^{-2t/\tau_a} \left[e^{2\tau_p/\tau_a} - 1\right] - 2 \cdot \tau_e e^{-t/\tau_e} \left[e^{\tau_p/T_e} - 1\right] \right\} \qquad (3.3.16)$$

to this must be added of course the c.w. component of the noise which is now given as:

$$i^2_{n_{out}}(c.w.) = \left(A \frac{R_d}{R_L}\right)^2 \cdot \frac{1}{4} \left\{ 2q \left[\left(i_{db} + i_b \right) M^2 E + i_{ds} \right] + \frac{4kT}{R_{Ld}} \right\} \cdot$$

$$\left(\frac{1}{\tau_d + \tau_a}\right) + kT(NF - 1) \cdot \frac{A^2}{R_L} \cdot \frac{1}{\tau_a} \qquad (3.3.17)$$

where NF = Noise figure of preamplifier

To optimize for minimum bit-error-rate, we use the same definition of Q given in Eqn. (3.3.9) substituting the corresponding values obtained from Eqn. (3.3.13 → 3.3.17) and using the same values for the parameters as given in the previous case, except that

$$R_d = R_L = 50\ \Omega$$

$$NF = 2\ dB,$$

we allow the computer to maximize S_{pp}/N_T. Plotting the results and comparing with Smith & Personick, we obtain Fig. 3.3.11.

It is noted that this system is worse than the previous system at low bit rate but approaches it as the bit rate increases. Past 500 Mbit, it is better and in any case the only combination commercially available.

3.4 Heterodyne Detection Systems

Fundamentally, heterodyne detection requires a local oscillator mixing with the incoming signal through a mixer. When this is done, it is easily shown that the detector signal current at the difference frequency is given by:

$$i_{if} = 2\, R_o M \sqrt{P_R \cdot P_{LO}}\, \cos \omega_{if} t \qquad (3.4.1)$$

where P_{LO} is the L.O. power at the detector and ω_{if} the intermediate or I.F. frequency.

The noise term in a heterodyne system can be simplified substantially by noting that the only noise term of importance is that due to the L.O. shot noise. Hence the noise term at the output of the I.F. filter can be simplified to

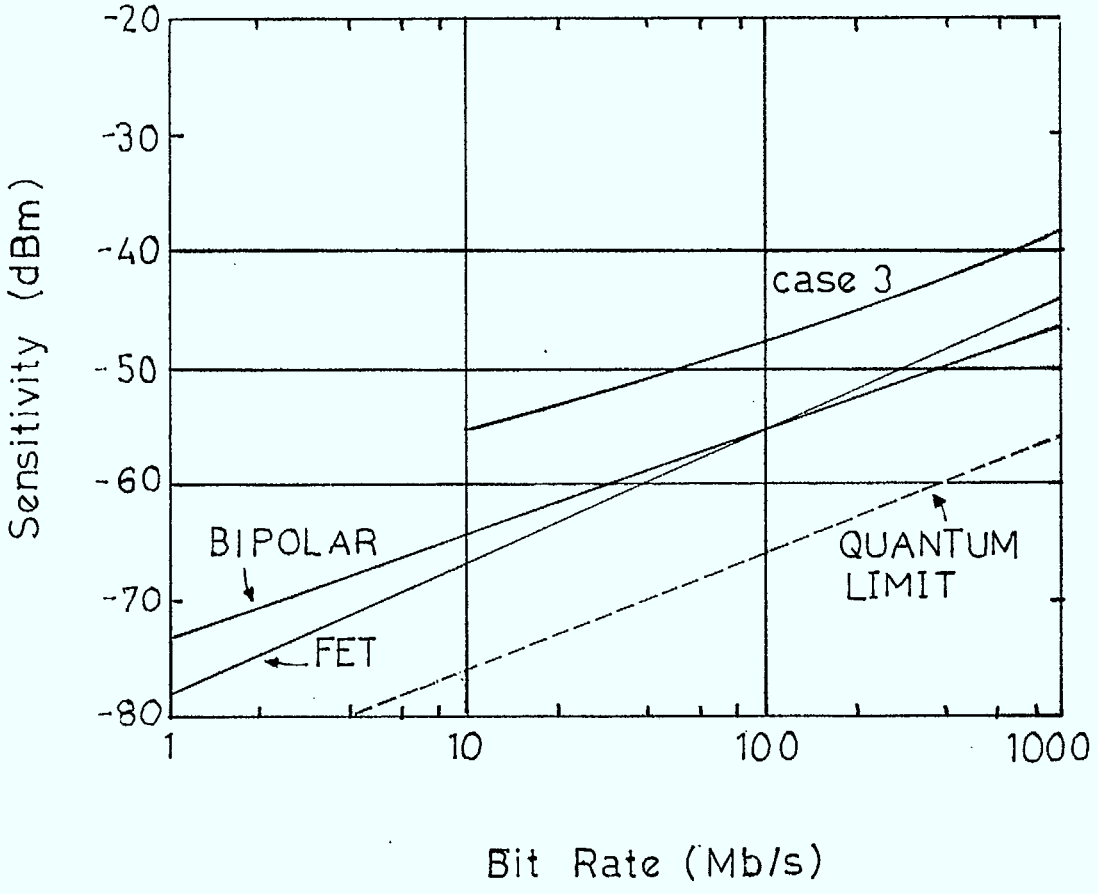


FIGURE 3.3.11: System Performance as a Function of the Bit Rate

$$i^2_{n_{if}}(\omega_{if}) = 2q\, P_{LO} \cdot R_o\, M^2\, E\, B_{if} \tag{3.4.2}$$

where B_{if} is the effective bandwidth of the I.F. amplifier.

Note that there is no time dependent noise term as in the envelope detection. We shall now consider two main modulation-demodulation schemes which will be representative of many others - namely, a heterodyne coherent O.O.K. and heterodyne-envelope O.O.K. system.

3.4.1 Heterodyne Coherent OOK

Figure 3.4.1 depicts the essentials of a heterodyne coherent OOK system. The received laser signal beats with the L.O. on the detector to produce a first i.f. current i_{if} after it passes through a bandpass preamp combination. The current is then heterodyned in turn with a reference current i_R on a second mixer which, after passing through a lowpass filter, becomes the output signal current i_{S_o}.

It can be shown that the time dependence of the output signal power (Watt) is given by

$$S_o(t) = i^2_{S_o}(t) = i^2_R\, R^2_o\, M^2\, P_{LO}\, P_S\, G^2_T(t) \tag{3.4.3}$$

where $G_T(t)$ is the time function part of the output current taking into account the signal pulse shape and the transfer functions of the detector, preamp bandpass filter and lowpass filter.

For simplicity, we assume that as was done in the previous section, the transfer function for the detector-preamplifier-bandpass filter can be put in the form

$$g_i(S) = \frac{1}{1 + \tau_i S} \tag{3.4.4}$$

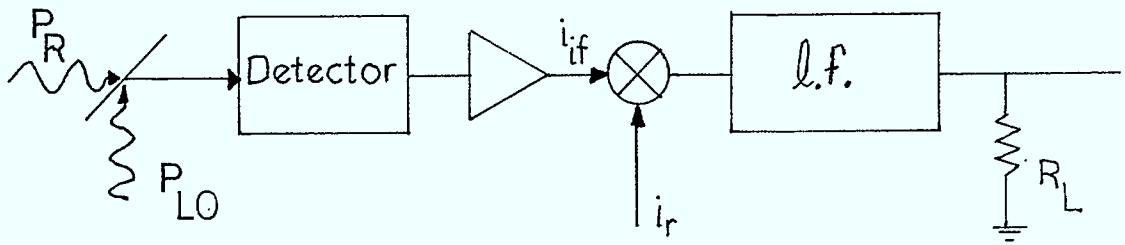


FIGURE 3.4.1: Heterodyne Coherent OOK System

and the transfer function of the output lowpass filter

$$g_o(S) = \frac{1}{1 + \tau_o S}$$

where $\tau_i + \tau_o$ are the respective time functions. It was shown in Eqn. 3.3.13 that under these conditions, the combined time function is for $t \leq \tau_p$.

$$G_T(t) = \left(\frac{1}{\tau_i - \tau_o}\right)\left\{\tau_i\left[1 - e^{-t/\tau_i}\right] - \tau_o\left[1 - e^{-t/\tau_o}\right]\right\} \qquad (3.4.5)$$

The noise output at the end of the lowpass filter is given by

$$i_{n_o}^2 = 1/4\, q\, P_{LO}\, R_o\, M^2\, E{\cdot}i_R^2\, B_o \qquad (3.4.6)$$

where B_o is the baseband filter bandwidth.

Hence

$$\frac{S_o(t)}{N_o} = \frac{4P_S \eta G_T^2(t)}{h\nu E \cdot B_o} \qquad (3.4.7)$$

Bit-Error-Rate

Using Assumption 2 as before, we get

$$Q^2 = \frac{P_S \eta G^2(\tau_p)}{h\nu \cdot E \cdot B_o} = 1/4\, \frac{S_o(\tau_p)}{N_o} \qquad (3.4.8)$$

where use has been made of the fact that for heterodyning $\sigma_{\ell+h} = \sigma_\ell$.

Also, in this case $B_o = \dfrac{1}{4(\tau_i + \tau_o)}$ (3.4.9)

Computer optimizations of the parameters for minimum BER gives results such as shown in Fig. 3.4.2 where $S_o(\tau_p)/N_o$ is plotted against P_R. Picking as before the S_o/N_o required for a BER of 10^{-9}, the sensitivity of such a receiver is plotted in Fig. 3.4.3. The comparision with our previous systems and Smith & Personick reveals that this heterodyne receiver is substantially better than all others and in fact approaches the quantum limit. A slight improvement over that system would be found with a coherent FSK which in essence is two heterodyne OOK back-to-back each with its own frequency shift.

3.4.2 Heterodyne Envelope Detection System (OOK)

The heterodyne envelope detection system is simpler physically than the heterodyne coherent system since there is no reference oscillator and a second mixer (see Fig. 3.4.4). On the other hand the noise theory is more complicated. This comes from the fact that the statistics are not Gaussian but rathe follow a Rician or Rayleigh distribution. The bit error rates are then given by:

For large S/N when i_ℓ is very small:

$$BER = 1/4 \text{ erfc } \frac{i_h}{2\sqrt{2}\sigma} + 1/2 \exp \frac{-i_h}{8\sigma^2} \qquad (3.4.10)$$

where $$i_h = 2\sqrt{P_S \cdot P_{LO}}\, R_o \cdot M\, G_T(\tau_p) \qquad (3.4.11)$$

is the envelope signal

and $$\sigma = 2\, q\, P_{LO}\, R_o\, M^2\, E\, B_{IF} \qquad (3.4.12$$

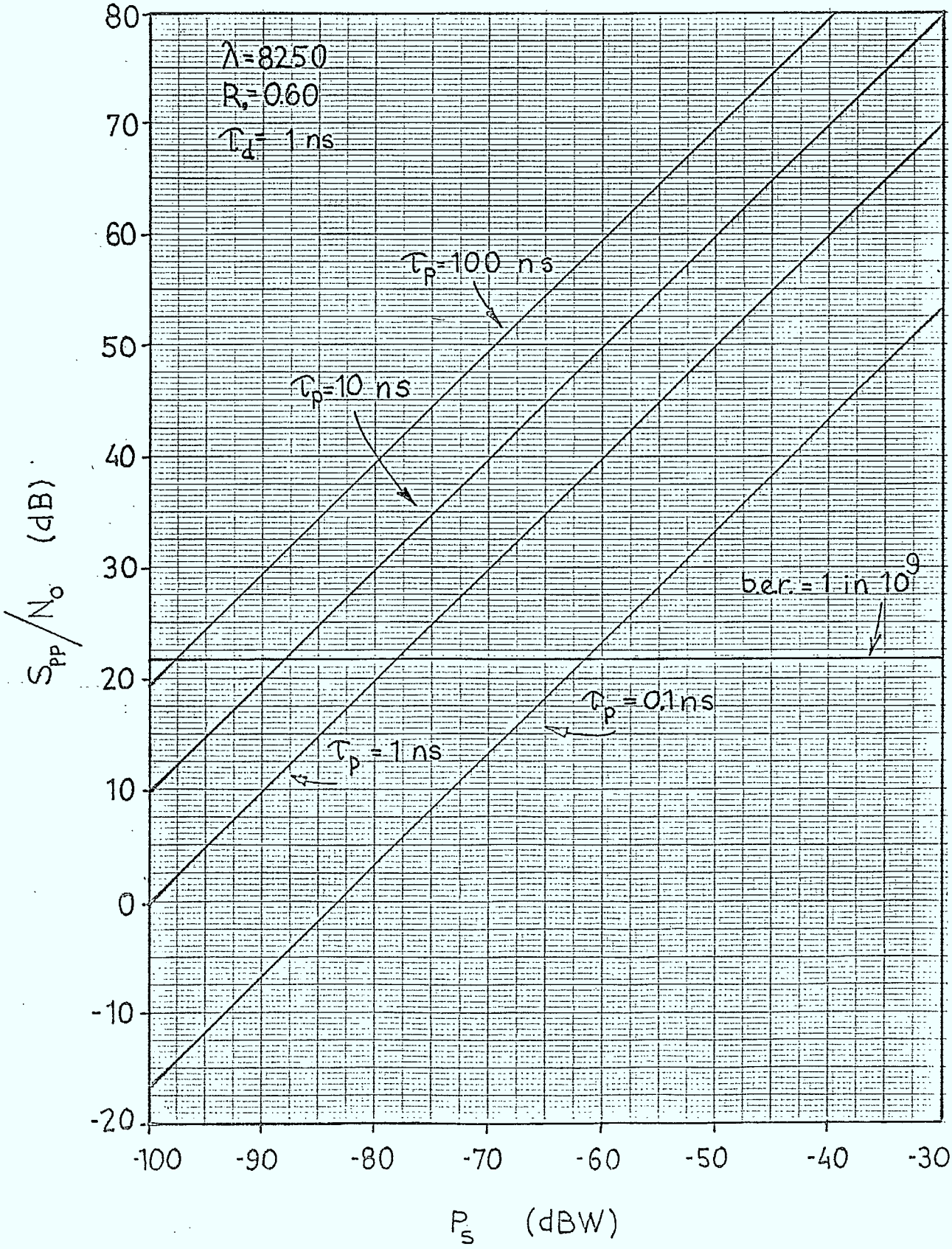


FIGURE 3.4.2: Optimum Signal to Noise Ratio as a Function of the Received Power

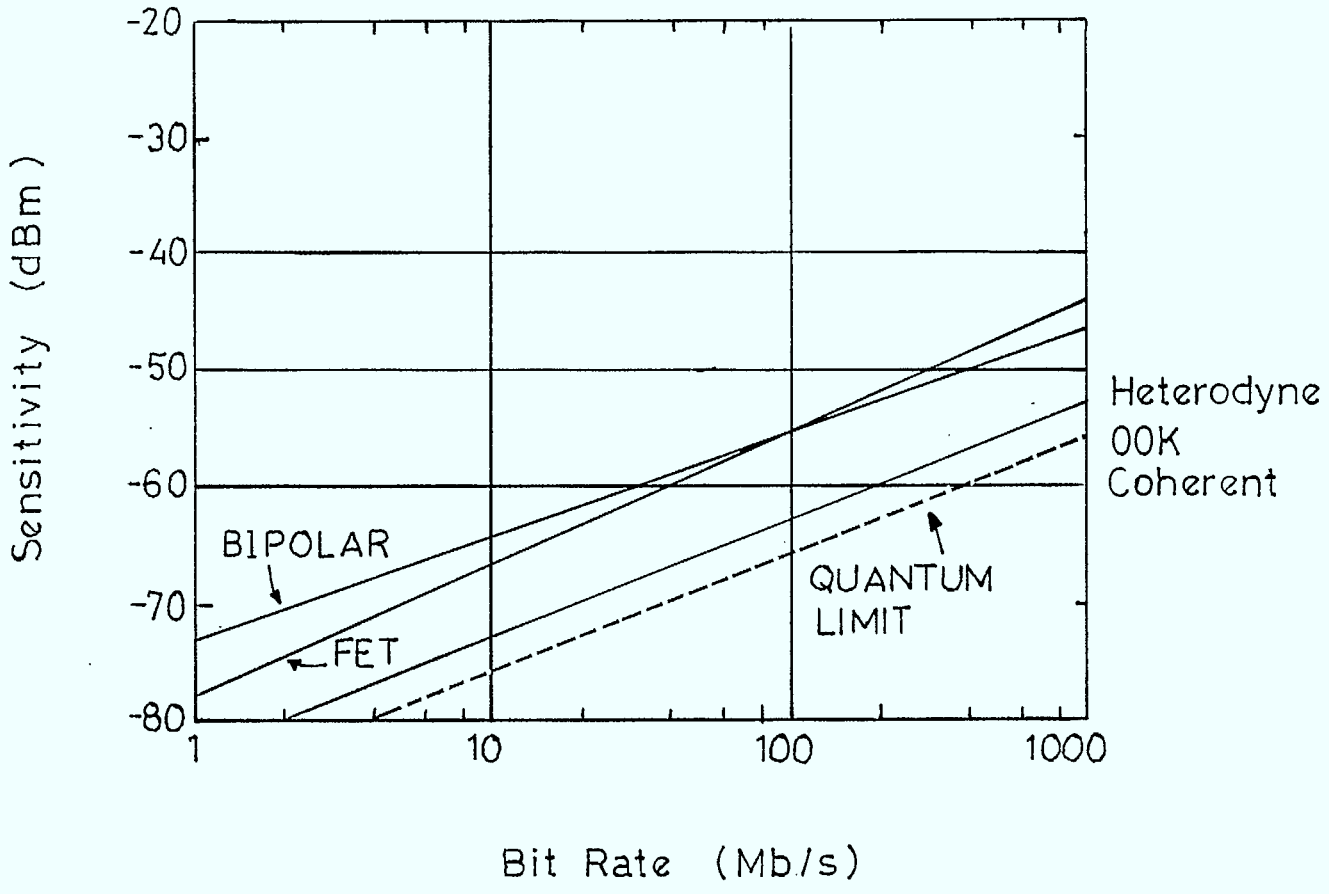


FIGURE 3.4.3: System Performance as a Function of the Bit Rate

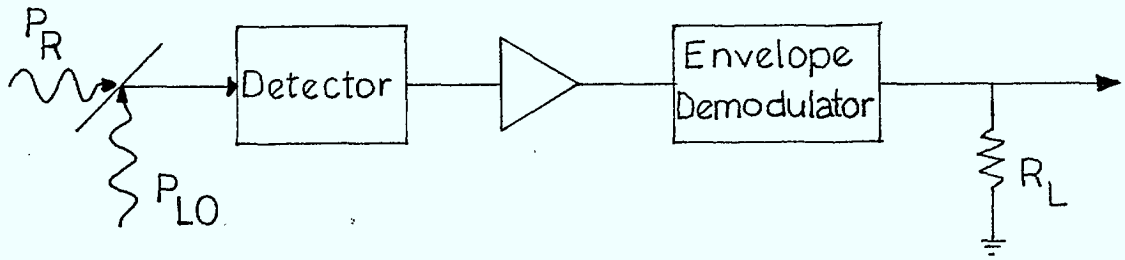


FIGURE 3.4.4: Heterodyne Envelope OOK System

when i_ℓ is substantial and cannot be ignored ($i_\ell/\sigma > 3$) then

$$\text{BER} = 1/2 \text{ erfc} \left(\frac{i_2 - i_1}{2\sqrt{2}\sigma} \right) \tag{3.4.13}$$

where $i_2 = i_h + i_\ell$

$i_1 = i_\ell$

and $i_\ell = 2\sqrt{P_S P_{LO}}\, R_o\, M\, G_T(2\tau_p)$.

Using condition (2), it can be shown that

$$\text{BER} \doteqdot 1/2 \text{ erfc} \left(\frac{1}{2\sqrt{2}} \left(S_{o\ell}/N_{o\ell} \right)^{1/2} \right) \tag{3.4.14}$$

where $S_{o\ell}/N_{o\ell}$ is the S/N at the output of the output filter. Computer optimization of the parameters for max $S_{o\ell}/N_{o\ell}$ minimum BER are plotted in Figs. 3.4.5, 3.4.6. It is seen that the sensitivity of the receiver for this case is intermediate between square law detection and heterodyne coherent detection.

It can be noted that the slope of sensitivity of direct detection (Figure 3.3.11) is smaller than that for the heterodyne cases. The implications are that, as the bit rate increases, the advantages of the heterodyne detection schemes decrease. It is expected therefore, that at some bit rate higher than 1 Gbit, the advantage of heterodyne detection might be lost. This same trend was observed in connection with HeNe and CO_2 laser systems that we have studied. More theoretical work is required to explore more fully the condition at which it is better to use direct detection at very high bit rate.

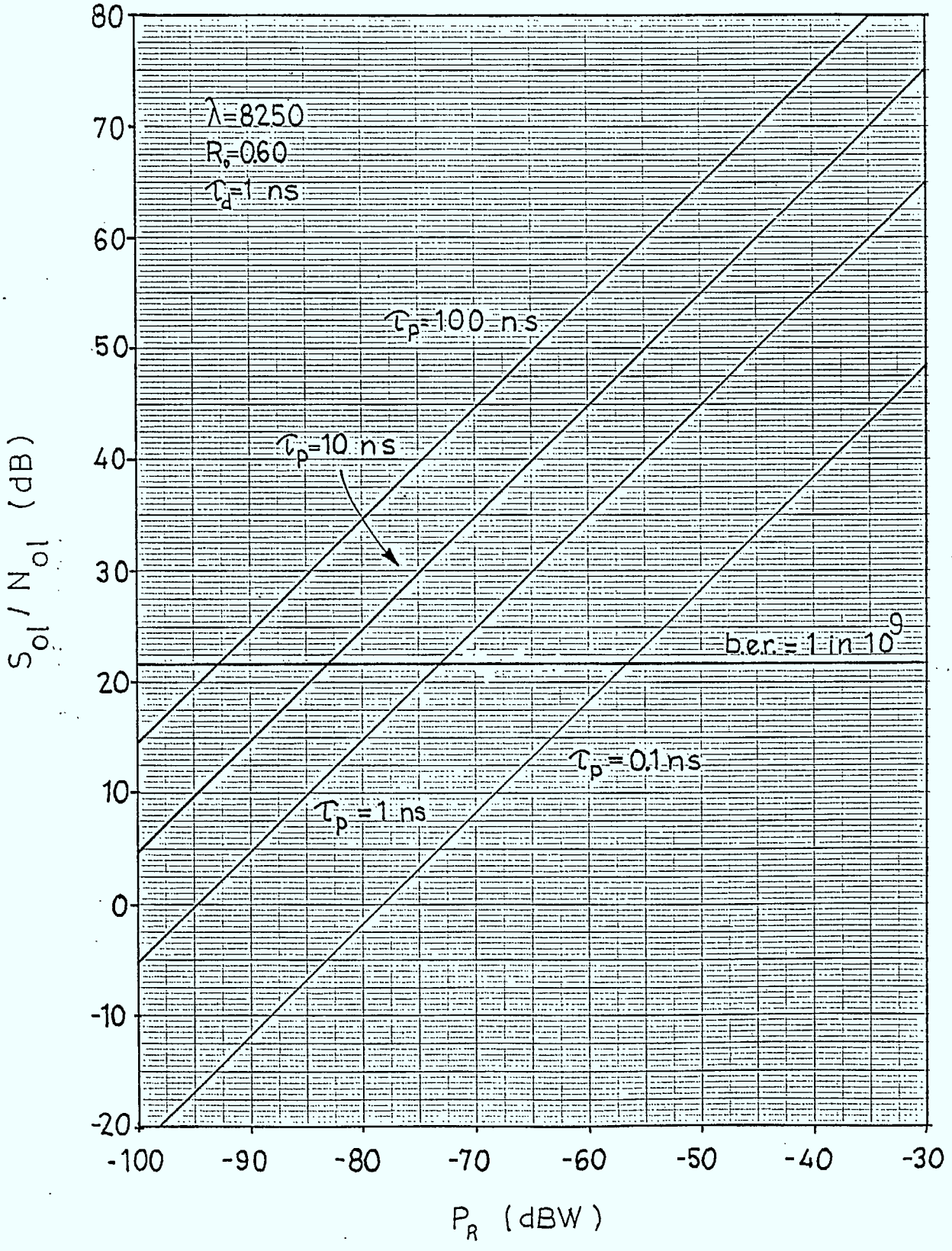


FIGURE 3.4.5: Optimum Signal to Noise Ratio as a Function of the Received Power

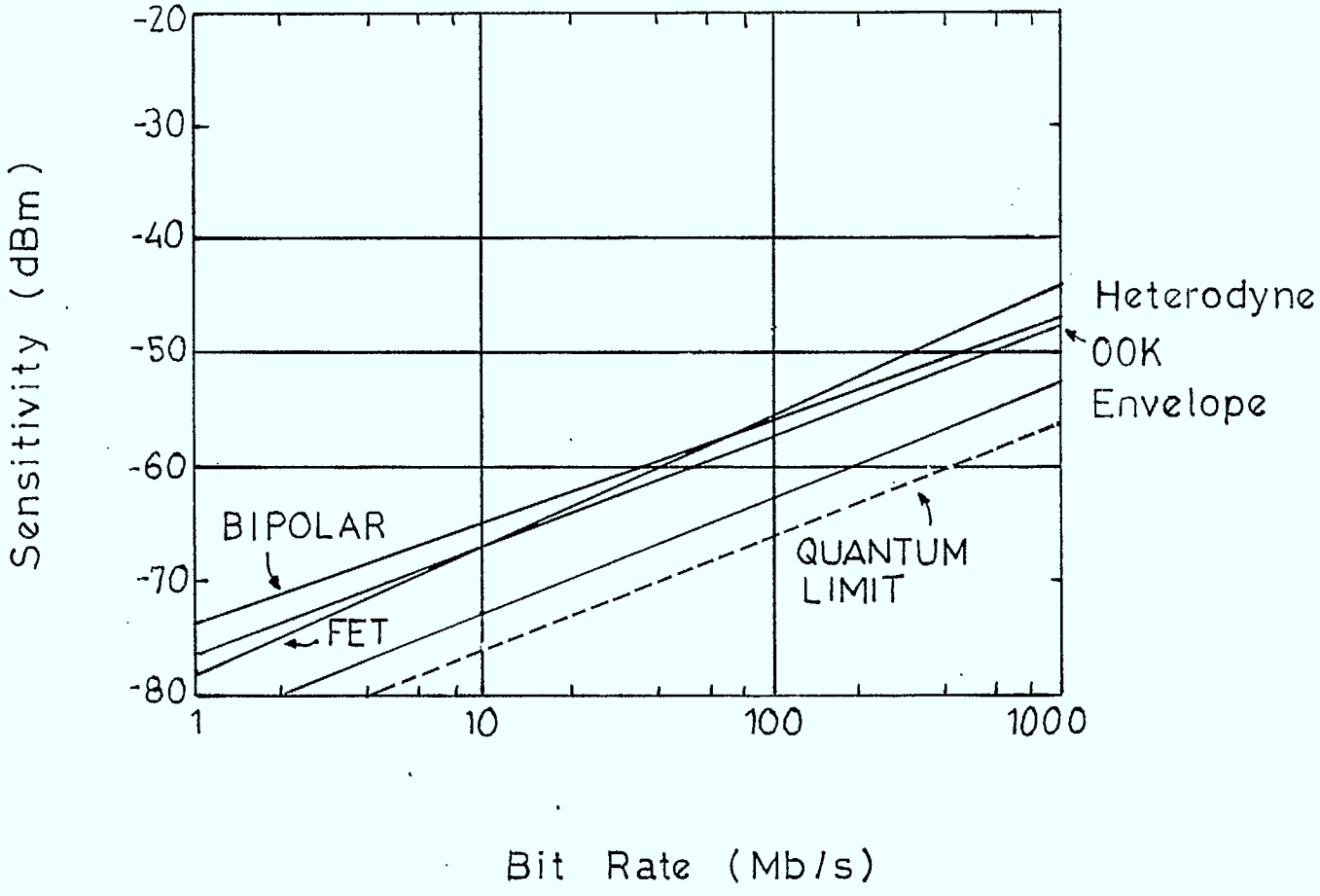


FIGURE 3.4.6: System Performance as a Function of the Bit Rate

4.0 MODULATION FORMATS

Optical carrier frequencies are typically three orders of magnitude greater than microwave carrier frequencies. This results in one of the major advantages offered by optical communication systems - a larger available signal bandwidth. Such bandwidths allow the exploitation of digital pulse code modulation techniques. In a pulse code modulation format the message signal is quantized in both time and amplitude. It is well known that the use of digital encoding techniques offers many advantages over analog modulation techniques. These include relative insensitivity to transmission noise and interference, ability for efficient regeneration and the possibility of a uniform format for different kinds of baseband signals.[4.1] Since it is generally accepted that digital communications will be used in satellite optical communication links, (in fact digital techniques are used in almost all optical telecommunication systems), especially if data communication is involved,this report will be concerned only with evaluating candidate pulse code modulation formats.

Two kinds of digital systems will be considered. These are baseband direct detection systems and coherent detection systems. Coherent detection systems can further be classified as homodyne and heterodyne detection systems. Modulation formats commonly considered for use in direct detection systems are direct intensity modulation (IM), also known as pulse intensity modulation (PIM), on-off keying (OOK), or amplitude shift keying baseband detection (ASK baseband), and pulse position modulation (PPM). Coherent detection systems utilize variations of three

fundamental modulation formats. They are amplitude shift keying (ASK), frequency shift keying (FSK), and phase shift keying (PSK). Both coherent and envelope detection techniques can be used in the coherent systems.

4.1 Direct Detection

An intensity modulation direct detection system represents the simplest configuration. A basic system block diagram is shown in Figure 4.1.1. In this format, the laser is turned on by a current pulse when a binary 1 is sent, and is turned off when a binary 0 is sent. In practice, the laser is never turned completely off in order to avoid laser turn-on delays. The zero level is usually set just above laser threshold such that the AM extinction ratio is maximized. PIM transmitters are generally designed so that the average output power level is kept constant by an electro-optic feedback loop. Laser spectral purity is generally not an important factor unless long distance transmission is performed over optical fiber, or the detector is preceeded by a narrowband optical filter. A photodiode-preamp combination of sufficient gain and bandwidth is used as a detector. The electrical signal out of the detector is then demodulated by a threshold detector. System performance is limited by the thermal noise in the receiver electronics and the sensitivity of the photodiode.

Pulse position modulation is another type of direct detection format presently being considered for intersatellite links.[4.2] In pulse position modulation one pulse is sent per bit period. The exact time slot within the bit period that the pulse is in represents the data transmitted in a corresponding

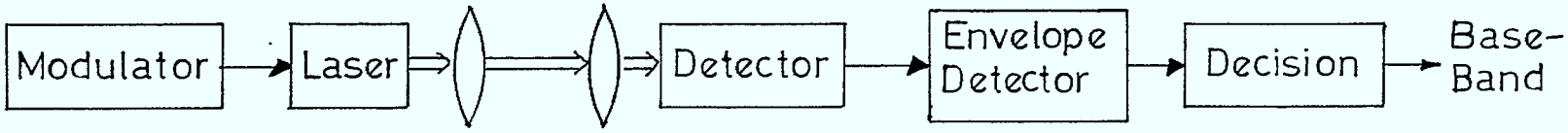


FIGURE 4.1.1: Basic Block Diagram for an Intensity Modulated/Direct Detection System

number of baseband bits. For example, binary PPM sends one pulse per baseband bit at an equivalent bit rate, quarternary PPM sends one pulse per pair of baseband bits at one half the baseband bit rate, etc. This scheme is illustrated in Figure 4.1.2. Although the basic system block diagram for the PPM system is the same as for the IM system, the individual blocks are slightly more complex. The most notable differences are in the modulator where a baseband tc QPPM encoder is needed, and in the decision circuit. The most notable advantage of PPM over IM is that no adaptive threshold circuit is required, rather the signal can be demodulated by suitably delaying the received signals in each time slot and using a "greatest of" circuit. A disadvantage of PPM is that it requires an increase in system bandwidth of at least a factor two over an equal capacity PIM NRZ system.

An analysis has been performed comparing the performance of OOK and low-order PPM modulations in optical communications using APD-based receivers.(4.3) The results showed a 3 dB improvement in receiver sensitivity for a QPPM system over that of an equivalent data rate OOK system. Receiver performance curves are shown in Figure 4.1.3 for the example communications link nominal parameters. Two important qualifications should be noted. First the absolute sensitivities are highly dependant on the physical system device parameters and secondly the PPM improvement will be slightly degraded by the increase in thermal noise caused by the increase in receiver bandwidth for PPM. Again various simplifying assumptions are made in the analysis, however it is useful as a relative comparison of the performance of the two modulation schemes. It can be clearly seen that there

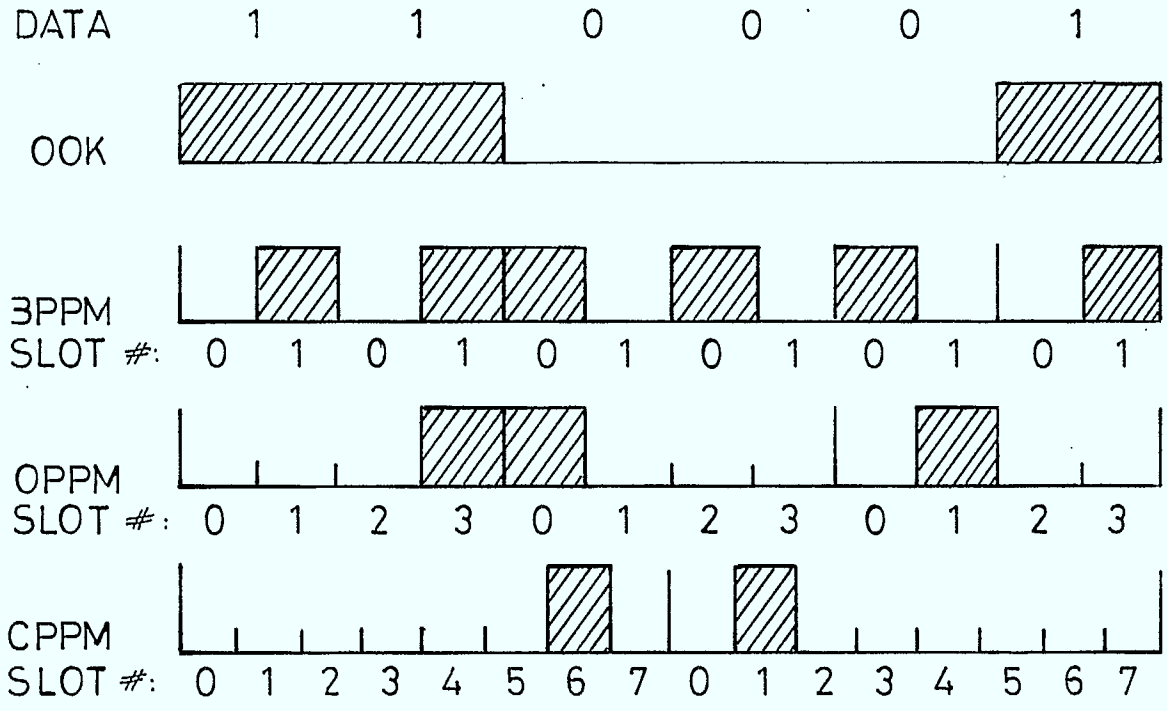
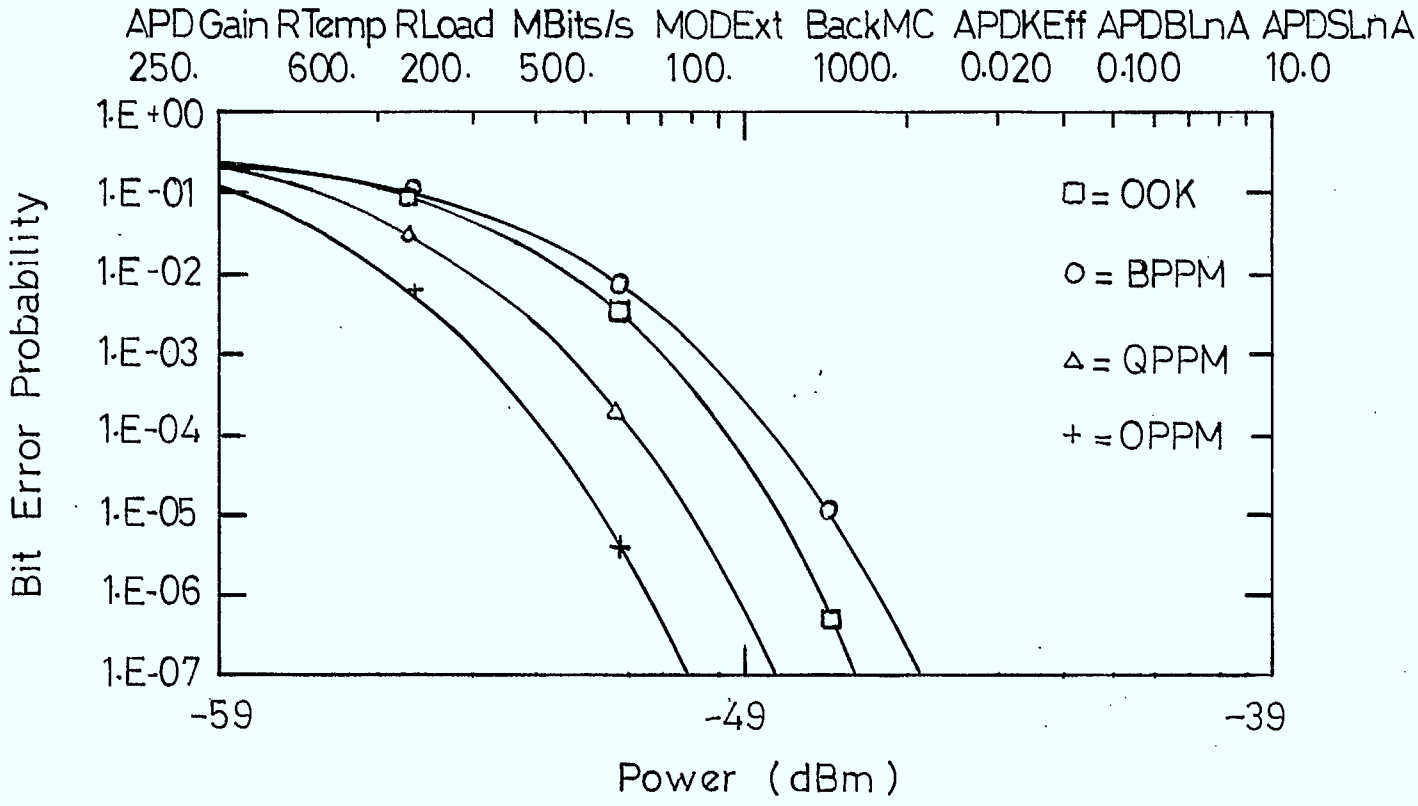


FIGURE 4.1.2: Bit Streams in OOK and PPM Modulation Schemes



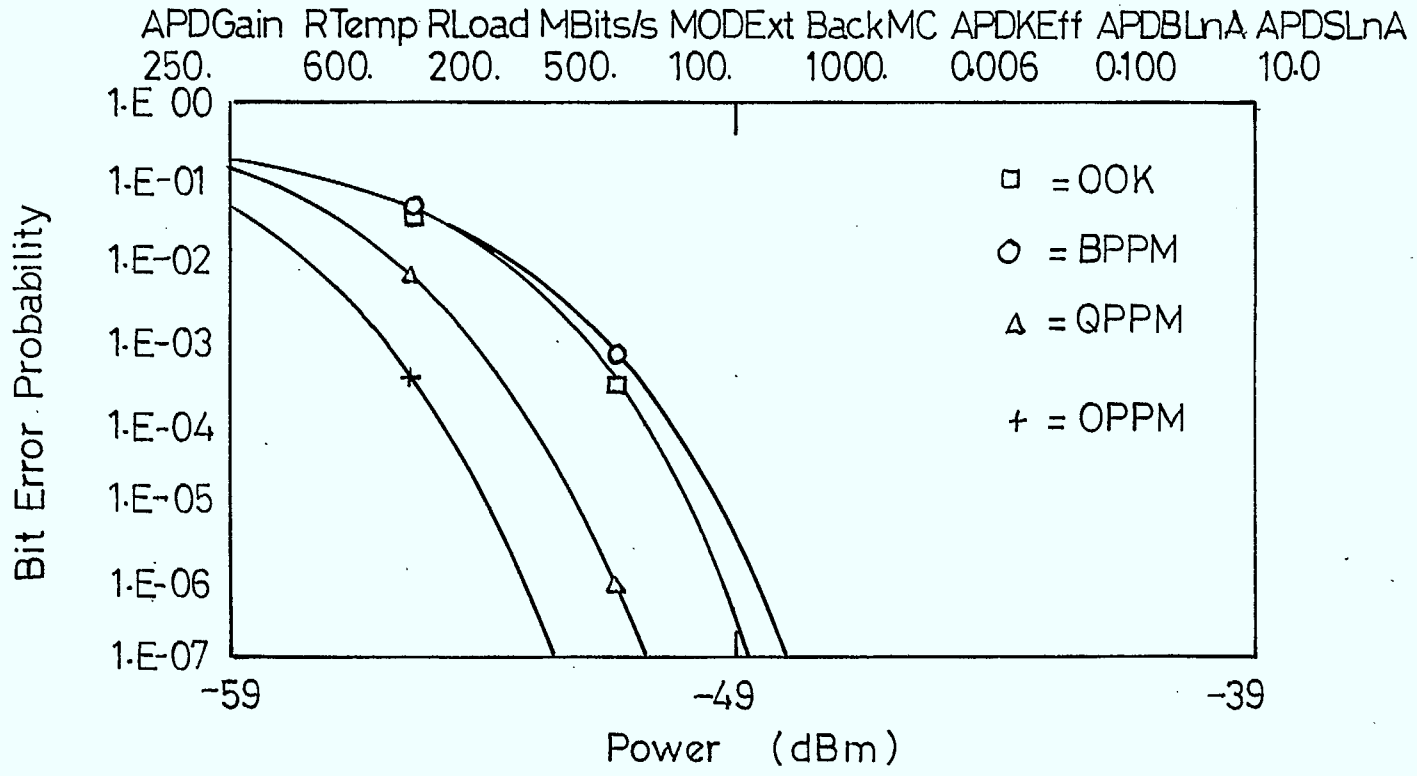


FIGURE 4.1.3: Receiver Performance Curves for the Given Parameter Values. The Second Set is for a Lower Noise APD.

is no performance advantage in choosing a BPPM modulation format over an OOK modulation format, however the 2-3 dB gain offered by a QPPM modulation format may well justify the marginal increase in complexity.

4.2 Coherent Detection

The various coherent modulation formats and demodulation techniques are summarized in Figure 4.2.1. In this figure, reference is made to two stages of detection: optical and electronic. Optical detection describes the difference frequency between the transmitter laser and the local oscillator resulting in either a homodyne or a heterodyne system. Electronic detection describes the signal demodulation after optical-to-electronic conversion. In homodyne schemes baseband is directly recovered while in heterodyne schemes either coherent detection can be performed wherein baseband is recovered by mixing the signal with an electronic local oscillator at the IF frequency, or conventional envelope detection may be used. Automatic frequency control of the optical local oscillator is an important aspect of heterodyne receiver design. Similarily, automatic phase control of the electronic oscillator is an important aspect of coherent receiver design. In homodyne systems this phase control is extended to the optical local oscillator.

Amplitude shift keying (ASK) involves the modulation of the laser's amplitude or intensity. Such a format is conceptually simple in that either the laser is pulsed or not. However, complications may arise in the laser spectrum when modulated at high powers with a large modulation index. A

FORMAT	OPTICAL DETECTION	ELECTRONIC DETECTION
ASK	Homodyne Heterodyne	 Coherent Envelope
FSK	Heterodyne	Coherent Envelope
PSK	Homodyne Heterodyne	 Coherent
DPSK	Heterodyne	Coherent

FIGURE 4.2.1: Summary of Coherent Modulation Formats and Demodulation Techniques

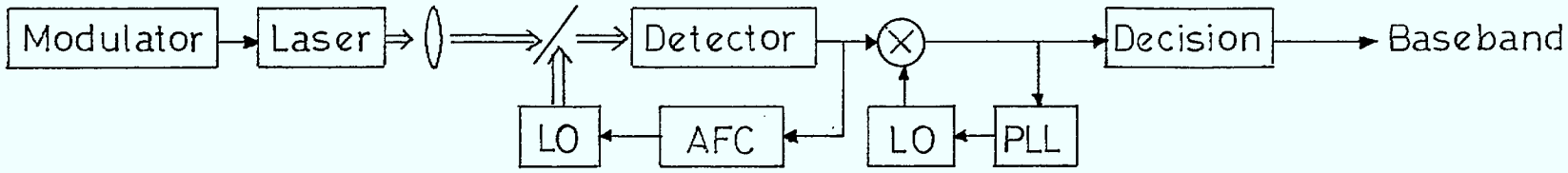
certain amount of frequency modulation is to be expected but this should not pose a serious problem as long as frequency drift is minimized. Four types of ASK receivers are shown in Figure 4.2.2. Of these four, heterodyne-envelope detection is the easiest to implement. ASK heterodyne-coherent requires the added complexity of establishing a phase tracking loop for the IF local oscillator. This task could prove to be difficult if there is substantial phase noise in the signal, as will be discussed. In an ASK homodyne system this phase tracking must be performed on the local oscillator laser with tolerances much finer than in the frequency tracking loop. This phase locking may be performed electronically using a phase lock loop (PLL) or by injection locking the local oscillator with the incoming signal.

Frequency shift keying (FSK) involves the modulation of the laser frequency. This modulation can be induced by the direct application of a current pulse of a few mA amplitude. Therefore, a small amount of amplitude modulation can be expected as well. A second method of performing the FM modulation is to use an external modulator. However, this adds to the transmitter complexity and at the same time reduces transmitter power by introducing an insertion loss into the system. Various FSK receivers are shown in Figure 4.2.3. Homodyne detection is not possible for FSK modulation. Of the schemes indicated the first two are the most common with envelope detection again being the easiest to implement. Single filter envelope detection is identical to ASK heterodyne-envelope detection and so does not give the FSK performance gain. Considerations regarding phase and frequency control affect the FSK receiver in a manner similar

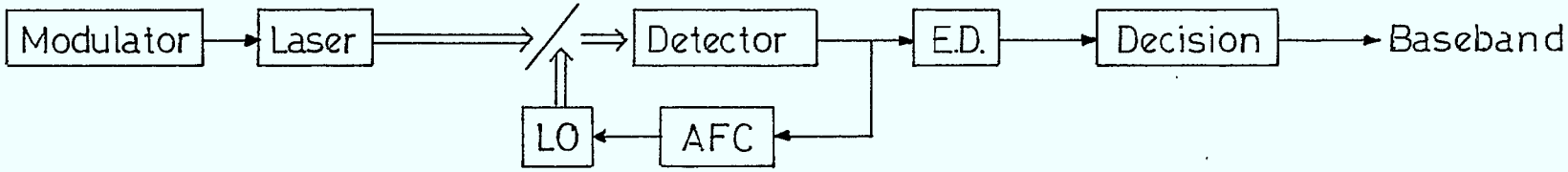
1 ASK

i) coherent

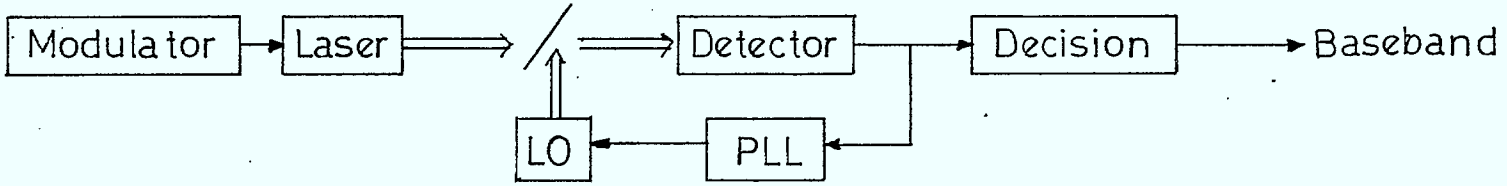
FIGURE 4.2.2: Basic Block Diagrams for an ASK Coherent Detection System



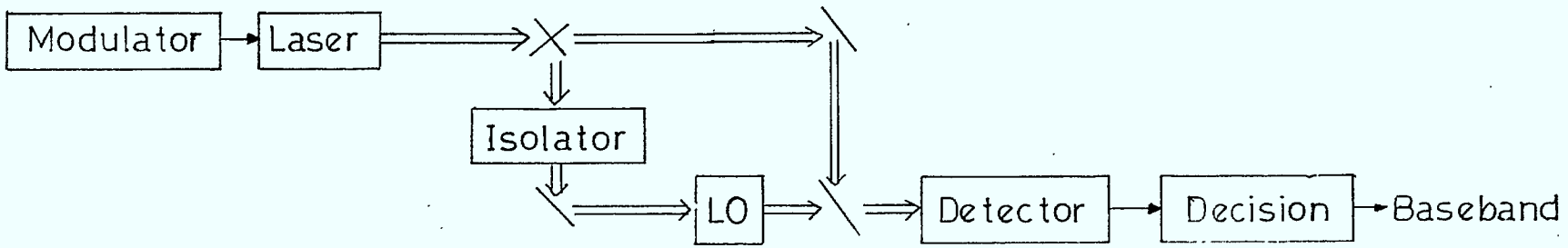
ii) envelope



iii) homodyne (optical PLL)



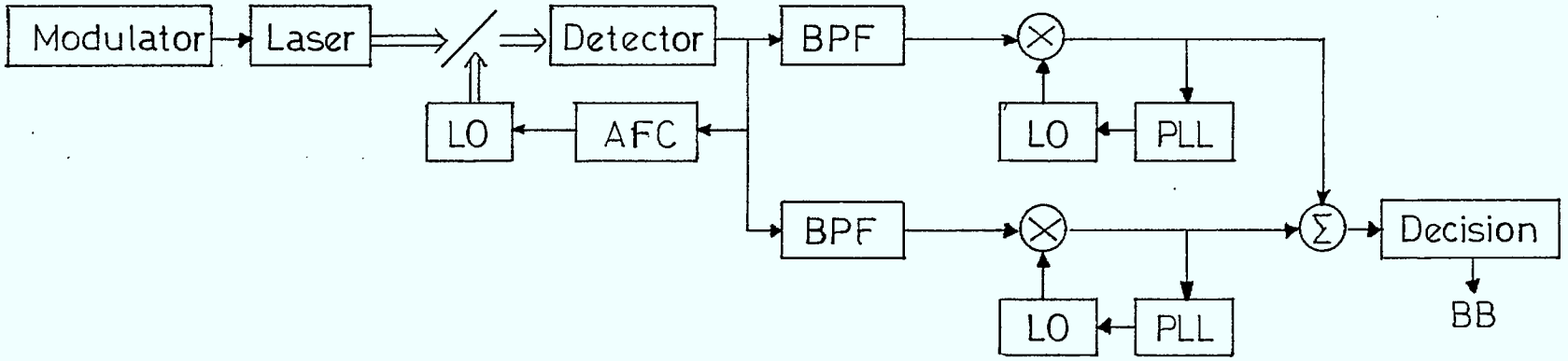
iv) homodyne (injection locking)



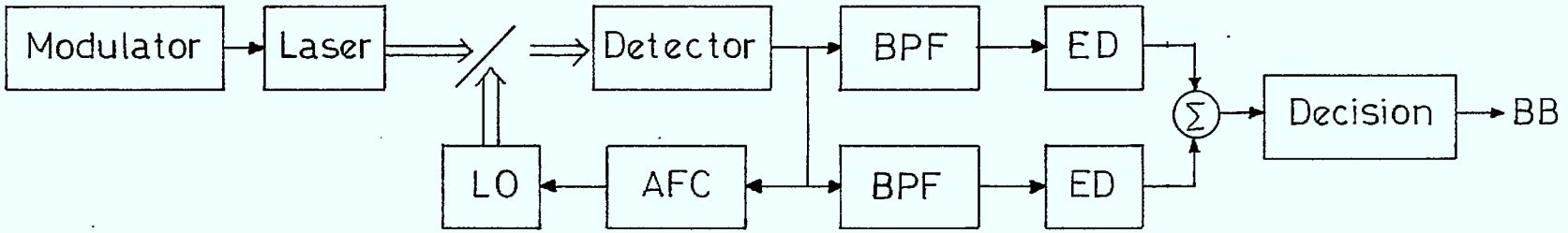
2-FSK

FIGURE 4.2.3: Basic Block Diagrams for an FSK Coherent Detection System

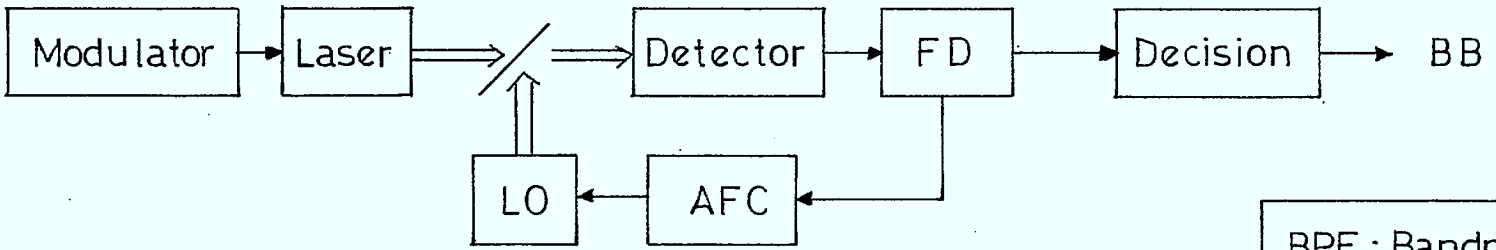
i) coherent



ii) envelope (dual filter)



iii) discriminator



BPF : Bandpass Filter
ED : Envelope Detector
FD : Frequency Discrim.

iv) envelope (single filter)

– same as ASK envelope

to the ASK receiver. One notable feature of the FSK receiver is the requirement for wide bandwidth. For example, operation at a transmission rate of 1 Gbit/s, could require a receiver bandwidth approaching 10 GHz depending on the detection scheme employed. In the envelope detection scheme, performance sensitivity to laser linewidth is reduced with increasing bit rate. It is the linewidth/bit rate ratio that is important, rather than absolute linewidth. However large linewidths may degrade the operation of the AFC loop.

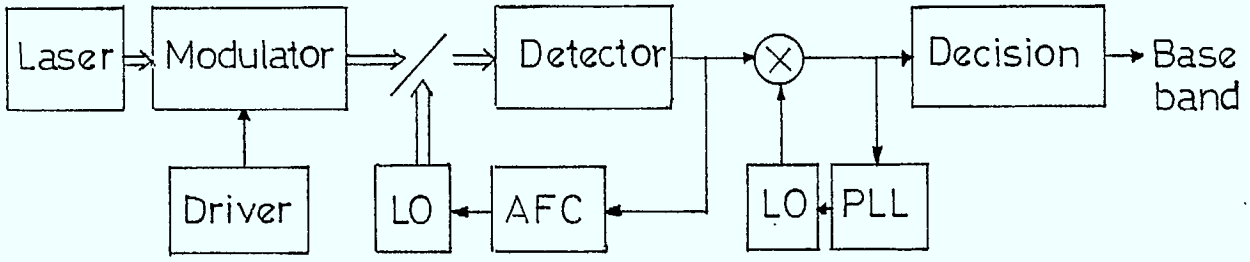
Phase shift keying (PSK) involves the modulation of the laser's phase. Although semiconductor lasers can be directly phase modulated if they are first weakly injection locked,[4.4] phase modulation is usually performed by the use of external phase modulators. Such a modulator increases transmitter complexity and introduces an optical loss source. It is imperative that this insertion loss does not negate the gain in receiver sensitivity. Three PSK receivers are diagrammed in Figure 4.2.4. The first receiver uses coherent detection much like the ASK and FSK coherent detection receivers. A receiver unique to phase shift keying is the differential phase shift keying (DPSK) receiver where the local oscillator is replaced by the signal itself delayed by one bit period. PSK homodyne is also possible, however establishing the necessary phase lock on the local oscillator is very difficult. Of these schemes DPSK is currently the most popular.

4.3 Comparison

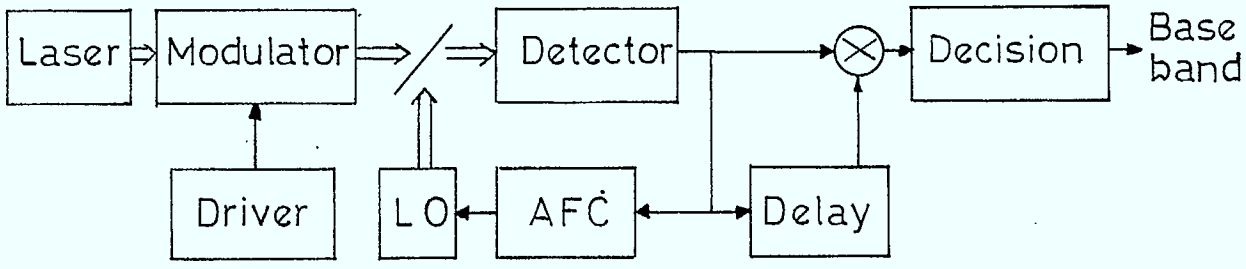
The relative format performance comparison given in this section is based upon an analysis performed by Okoshi et al.[4.5]

3-PSK

i) coherent



ii) DPSK



iii) homodyne

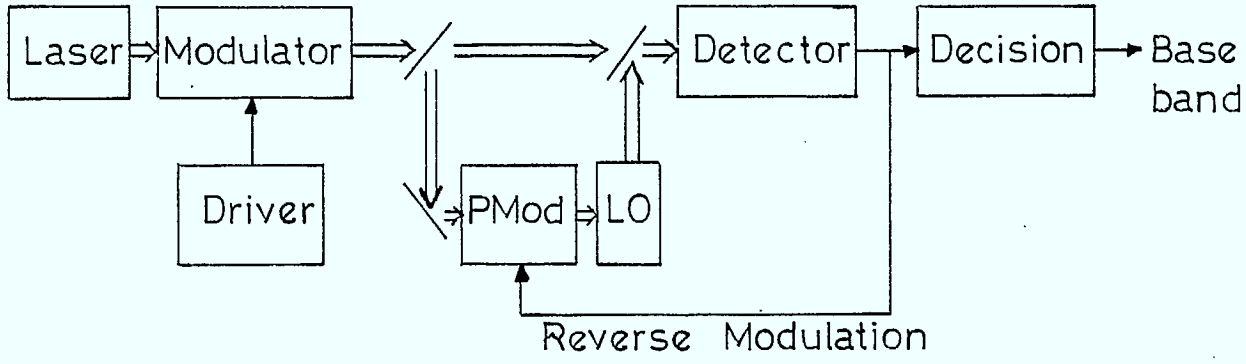


FIGURE 4.2.4: Basic Block Diagrams for a PSK Coherent Detection System

Okoshi's analysis follows from that of Yamamoto[4.6] and Smith and Personick.[4.7] This analysis uses a GaAs:FET amplifier after the detector while Yamamoto uses a Si:bipolar amplifier. Although the absolute receiver sensitivities are very much a function of the receiver electronics (i.e. thermal noise, wavelength, detector parameters, etc.), the theoretical performance gains attributed to the various heterodyne schemes are independant of these operating conditions. Okoshi's analysis also did not take into account degradations due to signal waveform distortion, laser amplitude noise and laser phase noise. Intrinsic laser noise can become very important at frequencies near the laser resonance peak which is typically a few gigahertz.

Figure 4.3.1 gives a comparison of an ASK heteroyne system with an IM/DD system. The figure shows minimum detectable signal power versus bit rate in order to maintain a bit-error-rate of 10^{-9}. The IM/DD curve assumes a dark current of 1 nA. APD gain is set at optimum. The dashed line shows the performance for an IM/DD system when maximum APD gain is limited to 100. Further, this plot is generated for a short wavelength system using a silicon APD with an excess noise factor of $x = 0.2$. Operating temperature is 300 K. Figure 4.3.1 illustrates a definite convergence of performance curves at high bit rates to within 1 - 2 dB. The actual improvement is very much a function of the operating parameters and front end receiver design. For example at high bandwidths it becomes much more difficult to achieve shot noise limited operation of the APD. Also the design of wide band detectors with high quantum efficiencies along with low noise electronics is very important. The heterodyne gain is greater at

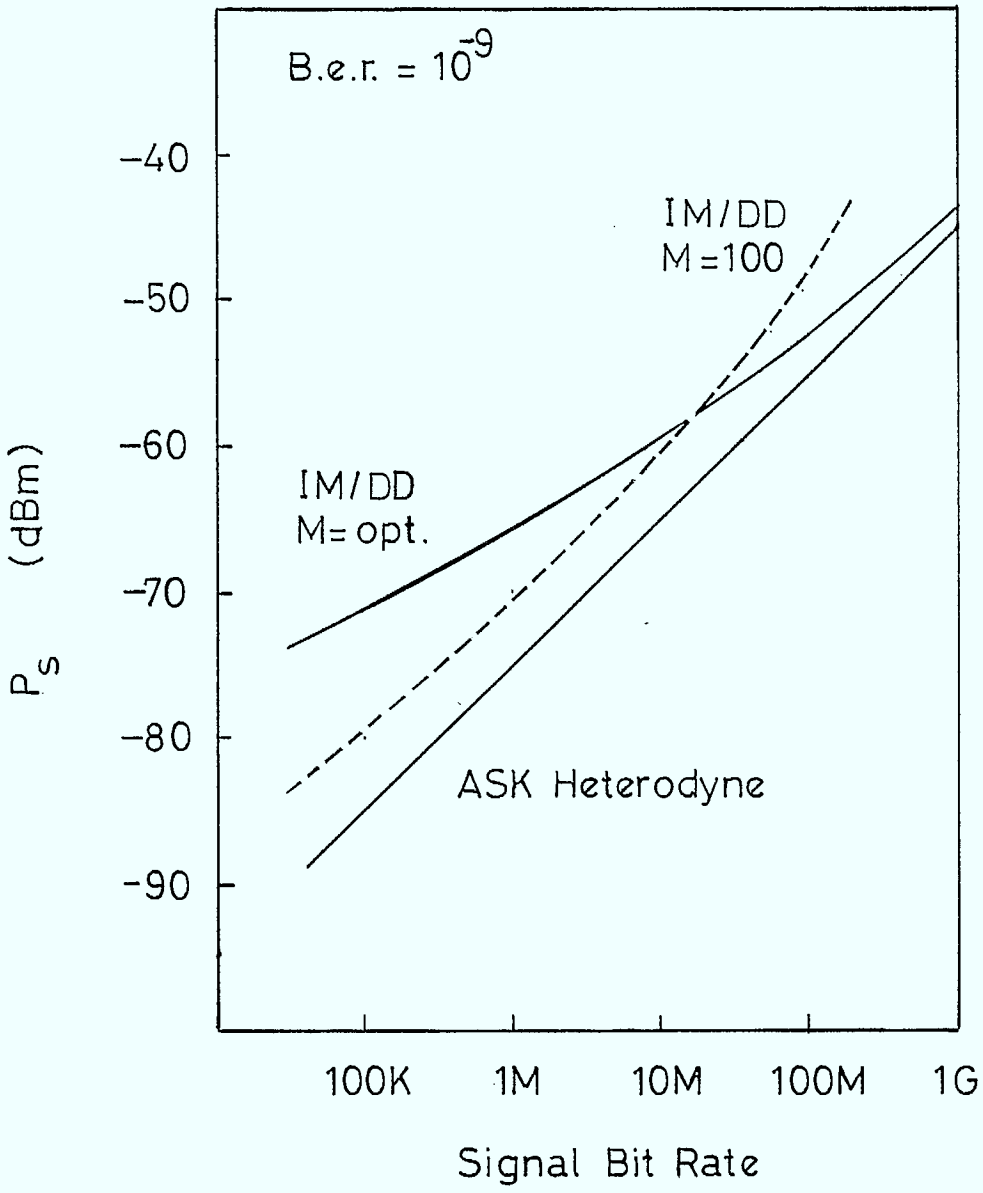


FIGURE 4.3.1: Comparison of an ASK Heterodyne System with an IM/DD System

longer wavelengths, primarily due to the higher excess noise factors of the materials used to make the APD's at these wavelengths.

Chapter 3 of this report gives a rigourous comparative analysis of IM/DD, ASK-envelope and ASK-coherent systems. Figure 4.3.2[4.5] tabulates the BER calculations and gives the theoretical performance gains. Homodyne detection gives a gain of 3 dB over like heterodyne detection by virtue of a factor two decrease in bandwidth. The analysis shows a 3 dB increase in sensitivity for FSK over ASK, however it should be noted that the same BER is achieved with the same average signal power in both because in the ASK system the signal power is zero in the spaces. The FSK system has another advantage in that its decision level is easily optimized. Phase shift keying gives an additional 3 dB gain over FSK systems although this advantage will quickly deteriorate with transmitter phase fluctuation. Little performance improvement is gained in using heterodyne-coherent detection over heterodyne-envelope detection, especially at higher signal-to-noise ratios where $\exp-(x)^2 \cong \operatorname{erfc}(x)$.

The general performance analysis identifies PSK homodyne detection as the most sensitive system. However, at present, technical difficulties in frequency and phase stability make it largely impractical to implement. The optimum performance/ease of implementation trade off appears to be with the FSK envelope and DPSK systems. Potential improvement is 3 - 6 dB above that of the ASK gain. These gains could reach 10 - 25 dB above IM/DD systems depending on the nature of operation.

Kikuchi et al,[4.8] later analysed the degradation of the

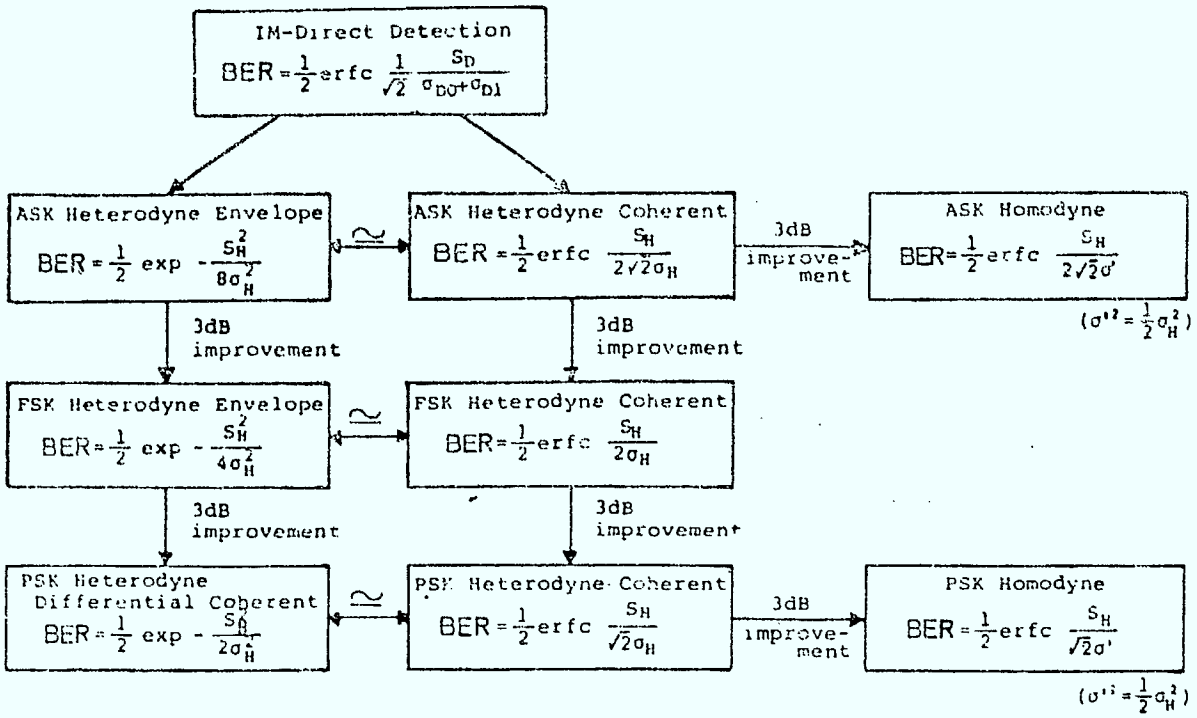


FIGURE 4.3.2: BER Calculations and Theoretical Performance Gains for Various Modulation/Demodulation Formats

bit-error rate in coherent systems due to the transmitter and local oscillator phase noise. Results of this analysis are shown in Figure 4.3.3 where BER vs SNR for various phase error is given where

$$SNR = \frac{S^2}{2\sigma^2}$$

$$s(t) = 2\sqrt{P_s(t)P_L}$$

$$\sigma^2 = 2e\,\frac{en}{h\nu}\,P_L B$$

The variance $\sigma\phi^2$ is determined by the bandwidths of the IF phase fluctuation $\phi n(t)$ and the phase-tracking circuit. The variance has been calculated as follows:

$$\sigma\phi^2 = \frac{\delta f_T + \delta f_L}{f_x}\ [\text{rad}^2]$$

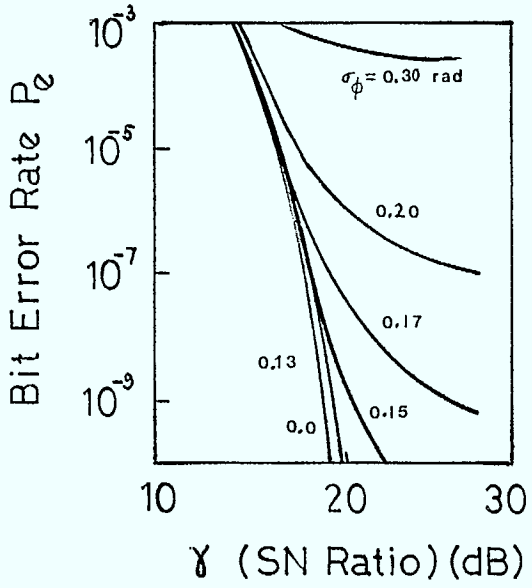
where δf_T = transmitter linewidth (FWHM)

δf_L = local oscillator linewidth (FWHM)

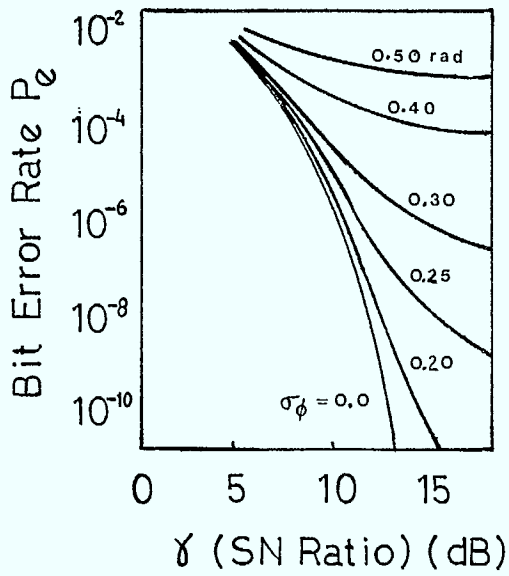
and $f_x = 2f_c$ in both heterodyne coherent and optical PLL homodyne scheme (f_c is the locking bandwidth of the PLL).

$f_x = f_s$ in a DPSK scheme (f_s is the signal bandwidth)

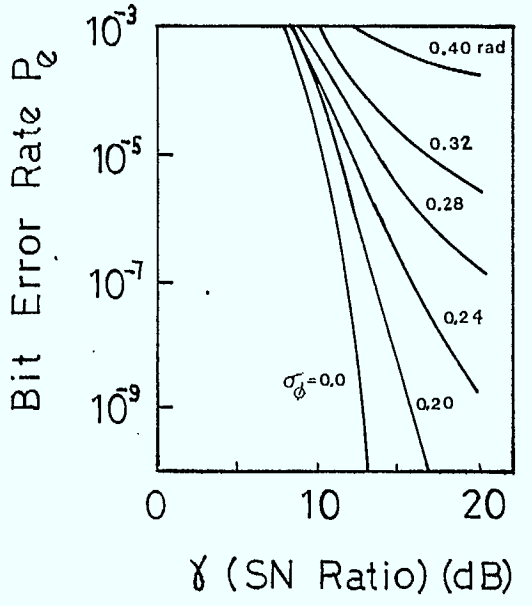
$f_x = 2f_i$ in an injection locked homodyne scheme (f_i relates the degree of injection locking)



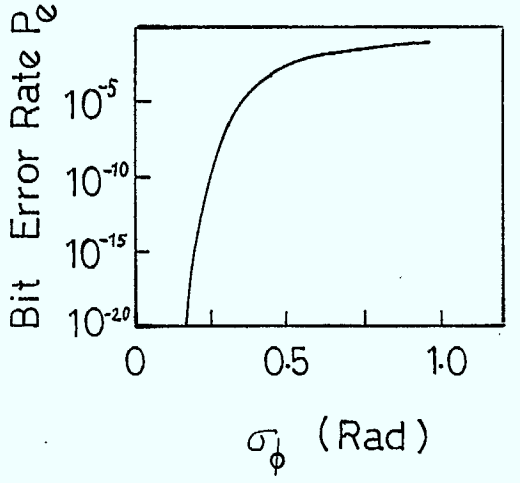
(a) ASK System



(b) PSK System



(c) DPSK System



(d) PSK and FSK System With Infinite S/N

FIGURE 4.3.3: Results of the Analysis of the Degradation of the BER in Coherent Systems due to Laser Phase Noise

The curves in Figure 4.3.3 show that for PSK schemes, the phase error must be less than 0.2 rad to prevent a S/N degradation of 3 dB at a BER of 10^{-9}. For ASK operation the phase variance restriction is much more severe. For $\delta f_T = \delta f_2 =$ 10 MHz for a GaAlAs laser with an output power of 10 mW,[4.8] f_x must be greater than 500 MHz for $\sigma < 0.2$ rad. This dictates a PLL bandwidth of greater than 250 MHz for heterodyne PSK and optical PLL homodyne systems, a signal bit rate greater than 3 Gbit/s for heterodyne DPSK and a signal power greater than -37 dBm for injection-locking homodyne. These calculations demonstrate the difficulty in implementing heterodyne coherent detection and homodyne detection techniques. In fact, the injection-locking homodyne technique cannot be implemented since the receiver sensitivity must be better than -37 dBm to be useful. Figure 4.3.3 also shows a theoretical BER curve for PSK and FSK schemes as a function of σ in the limit of infinite S/N. This curve shows that even for infinite S/N, the phase noise σ must be less than approximately 0.25 rad for a BER of 10^{-9} to be achieved.

Therefore, unless the linewidths of the semiconductor lasers can be substantially reduced, it would appear that from an ease of implementation standpoint, heterodyne-envelope detection receivers are the most attractive. This is certainly the case for high S/N where the theoretical performances are similar.

4.4 Signal Multiplexing

As mentioned in Section 2.4, frequency division

multiplexing techniques are applicable to optical heterodyne systems. However, before FDM can become practical the problem of laser frequency stabilization must be solved. The transmitter lasers would need to be frequency locked together, or have sufficient guard bands, to ensure that one laser does not drift into the frequency window of another. Also local oscillator tuning must be well established in order to lock onto the proper carrier. The advantage of FDM technique is increase in link capacity that it will allow. However, because of the increase in complexity that it adds to the system it is likely that TDM at high bit rates will be fully exploited before FDM techniques are used. It is also likely that wavelength division multiplexing (WDM), which is a crude form of FDM, will find application before FDM simply because of ease of implementation.

It is expected that optical ISL heterodyne links will accommodate conventional TDMA communications in certain scenarios. One such scenario would have conventional microwave TDMA communications between LEO (low earth orbit) or satellites, aircraft or earth stations and geostationary communication or data transfer satellites. Optical links could then be established between a number of these relay satellites to form global communication networks. The optical transmission technology should certainly be capable of operating in the burst mode of operation that such a system would demand. The optical bus between these satellites would be constantly maintained. However, it is unlikely that TDMA would be feasible if optical links replaced the microwave links used in communication with the relay satellites. The narrow laser beam is incapable of covering

an area large enough to include many targets, therefore multiple optical links are needed to establish this communication. These links would require sophisticated acquisition, tracking, splitting and combining subsystems. Thus, it is anticipated that strictly optical TDMA links may only be implemented with a high level of complexity, however the optical transmission technology can be operated in burst mode which will accommodate TDMA operation if earth/satellite communication uses conventional techniques.

5.0 ACQUISITION AND TRACKING

The problem of pointing acquisition and tracking in an optical satellite communication link is very difficult. This is essentially because to make full use of the advantage of a laser communication system, a laser link must have an extremely narrow beamwidth (in the sub. microradians). This puts a major constraint on the mechanical pointing requirement of the link. In addition, because the required cone of the beam is much narrower than the cone of uncertainty of the position of the other satellite, a special stragegy must be adopted during acquisition to widen the laser beam to encompass this uncertainty. The result however is a decrease in S/N at a time when it is most needed.

A vast amount of theoretical studies have been published on the subject. For completeness in this report, we shall touch upon the subject to bring out the problems and possible solutions. We shall refer the interested reader to the literature (See References 5.1 - 5-13) for more comprehensive discussions.

5.1 Acquisition

Acquisition, in the context of optical communication, can be defined as the process whereby two satellite stations search for each other using the optical link until acquisition is obtained and then proceed towards a fine tracking mode that will allow two-way communication.

Some "a priori" information is assumed initially as to the approximate location of the other satellite station. The mutual search must therefore cover the total solid angle of uncertainty

in order to permit the acquisition and then the process of establishing contact with the other station.

The main objective to be achieved is to cover the complete area of uncertainty in a reasonable time and that once acquisition is made, to switch to a tracking mode fast enough so that the contact is not lost during the process. Problems are different if acquisition is between LEO (low earth orbit) or GEO (geosynchronous earth orbit). Also a LEO to GEO link requires acquisition periodically since the line of sight is broken every time the LEO is eclipsed by the earth.

Positional uncertainty of one satellite with respect to the other includes a) altitude uncertainty of the satellite itself, b) orbital ephemeris of the position and velocity of the other satellite with respect to the first one, c) relative timing errors, d) point ahead errors, e) attitude uncertainty of the satellite itself, f) vibration noise of the satellite.

It was found in a study(5.1) that satellite attitude errors dominated the other sources and could be of the order of .25°, 3 axis rms. It can therefore be assumed that the angular uncertainty that could be used in systems analysis could be taken as .5°, describing the full uncertainty cone from one satellite to the other.

Analysis showed(5.1) that acquisition time was directly proportional to the uncertainty solid angle and inversely proportional to laser power if acquisition is done by array or quadrant detectors.

The possible acquisition systems are many, each with its own advangtages and disadvantages. Fundamentally, it consists

of a satellite sending a beacon in the probable direction of the other satellite. This could be done as wide as the cone of uncertainty, or narrower, but made to scan the full cone. At the receiving end, acquisition of the beacon could be achieved by using an array detector to cover the full cone of uncertainty, or alternatively using a quadrant detector, or a single detector, but have the receiving optics scan the entire area of uncertainty. One advantage of this latter scheme is that the same transmitter/receiver optical subsystem can be used for both acquisition and then communication. A reprint of a study which we have made for NASA is included for convenience at the end of the section.

Whatever technique is used for acquisition requires a predetermined protocol of a "handshake" such that when one satellite acquires the other, it transmits that information to the other. Each satellite can then narrow its transmitter beam to its final narrow communication beam thus boosting up its S/N. Wideband communication can then begin. A tracking mode must be counted on to keep the two transmitting beams continuously pointed towards each other during regular communication.

5.2 Pointing

A systems study described in Ref. 5.1 reveals that with present technology, servo system should allow a 1 μrad pointing and tracking accuracy using gyro stabilized or other approaches.

Such high pointing accuracy requirements put very high demands on the electromechanical control system of the gimbal mounts. A large number of error sources have to be considered in the analysis such as sensor noise, spacecraft base motion

disturbances, friction torque disturbances, point ahead errors, gyro noise, dynamic following error,etc. Detailed analysis [5.1] however shows that careful design could result in pointing errors of less than 1 μrad.

Efficient optical communication telemetry format analysis, has shown [5.1] that the digital pulse position modulation is the better scheme to use. The bit rate requirement for pointing control is of the order of 10 Kbps. With 8 time slots, the time per slot is 37.5 μsec. Typical laser diodes giving out 20 mW average power and APD are adequate for this use.

5.3 Burst and System Errors

The origin of bit error rate can be categorized as coming from the communication systems channel or from burst error. The former source is that which is usually calculated in a communication channel and occurs from the standard noise sources such as shot noise, Johnson noise, etc. The probability of error is then derived from the signal to noise ratio during sampling. The burst error on the other hand, is associated with the probability that the servo system will maintain pointing within some predetermined error limit. Burst error could be the cause of the loss of many bits since the response time of the pointing system, being mechanical, is much longer than the bit time. Therefore a sizeable amount of effort is required to reduce the burst error to a minimum.

5.4 Summary

The result of a number of studies seem to confirm, that although acquisition and tracking is a difficult problem for optical ISL, it should nevertheless be possible to design and

build such systems both for LEO and GEO terminals. This would be possible for data rate of the order of 500 Mbs and should not impose undue penalty to the host satellites.

Reprinted from IEEE TRANSACTIONS
ON *AEROSPACE AND ELECTRONIC SYSTEMS*
Volume AES-6, Number 3, May, 1970
Pp. 407–409

A Dual Scan Acquisition Technique for a Laser Communication System

Abstract

A dual scan acquisition system is described in which both the transmitter and the receiver optics scan, although at different frequencies. The condition where only one station scans appears as a special case of the general approach. Curves of acquisition time versus various parameters are calculated. The effect of range is also discussed.

Introduction

With the advent of satellite communication systems, interest in the possible use of lasers as a carrier is developing rapidly. One of the first problems that will have to be solved, in such a system, is that of acquisition. The problem arises because a certain amount of uncertainty exists as to the exact location of each satellite. For acquisition, then, a spatial search routine has to be established. One of the simplest and most elegant ways of accomplishing this was proposed by McAvoy *et al.*[1] This consists essentially of spreading the transmitting beam of one of the satellites to cover the whole region of uncertainty in the position of the other satellite. The latter satellite then accomplishes a spatial scan with its receiver until it acquires the wide laser beam. The critical problem is caused by the fact that a wide laser beam implies very little power received. This problem is compounded by the fact that the sensitivity of the superheterodyne system is lower prior to acquisition, because a certain amount of uncertainty exists on the relative laser frequencies between transmitter and LO. The bandwidth of the receiver must therefore be widened to accommodate this uncertainty with a resulting loss in sensitivity.

We shall consider here a generalized approach, in which both the transmitter and the receiver scan, although at different frequencies. The transmitted beam is, however, narrower, in general, than that required to cover the whole region of uncertainty. It will be found that this system, when optimized, can provide substantial reduction in acquisition time. The method proposed by McAvoy *et al.* is a special case of this generalized approach.

Theory of Acquisition

Geometrical Considerations

Let A and B be two satellites trying to acquire each other as shown in Fig. 1. The general approach we propose is then as follows. Transmitter A produces a beam in space of angle α_A during acquisition. It is held momentarily in a fixed position. During this time, the optics of B, with receiving field of view α_B, scans one complete angular frame θ_B^2 in a raster fashion. Here θ_A and θ_B are the linear angles of uncertainty to be searched by A and B, respectively. At the end of the θ_B^2 frame scan, the optics of A deflects one unit angular beamwidth α_A. One complete frame is again scanned by B. This is repeated for each step deflection of A using a step-raster scan until the acquisition of A by B is completed. At this point, B stops its scan and sends a narrow beam (angle η_B) in the direction of A. The receiver of A, of field of view α_A, then detects this high intensity beam (because of the narrow field of view) and localizes the position of B. The optics of A is then aligned towards B and its field of view reduced to its final narrow angle η_A. At this point then both A and B receive and transmit on their communication narrow field of view η_A and η_B, respectively.

The total scan time required to complete one full search occurs when one complete raster of A has been entirely stepscanned. We shall call this time the equivalent period T_e.

Let the power received by B during communication be $P_{\eta_A}=KP_{o_A}$ where K is a factor taking into account the geometry optics and wavelength used and P_{o_A} is the power output of A. During acquisition, however, the received power is reduced to

$$P_{\alpha_A}=\left(\frac{\eta_A}{\alpha_A}\right)^2 P_{\eta_A}. \tag{1}$$

Let us now redefine angles in terms of scanning cells. Let $n_A=(\theta_A/\alpha_A)$ be the total number of scanning "cells" required to cover the angle α_A. Similarly, $n_B=(\theta_B/\alpha_B)$ is the total number of cells to cover the uncertainty angle α_B. In terms of these, we have

$$P_{\alpha_A}=P_{\eta_A}\left(\frac{n_A^2}{n_{Ao}^2}\right) \tag{2}$$

where $\eta_{Ao}=(\theta_A/\eta_A)$ is the maximum number of linear cells possible into θ_A; this occurs with the narrowed field of view η_A which is diffraction limited.

Electronic Considerations

It is shown above that during acquisition the power received by the satellite can be reduced considerably if the angular uncertainty in position is large. In general this would reduce the SNR of the received signal close to or below threshold. Since no communication is required during acquisition, one of the simplest

Manuscript received October 20, 1969; revised December 18, 1969.
This project was supported in part by the Defense Research Board of Canada's Industrial Research Program under Grant 5501-55.

[1] N. McAvoy, H. L. Richard, J. H. McElroy, and W. E. Richards, Goddard Space Flight Center, Rept. X-524-68-206, May 1968.

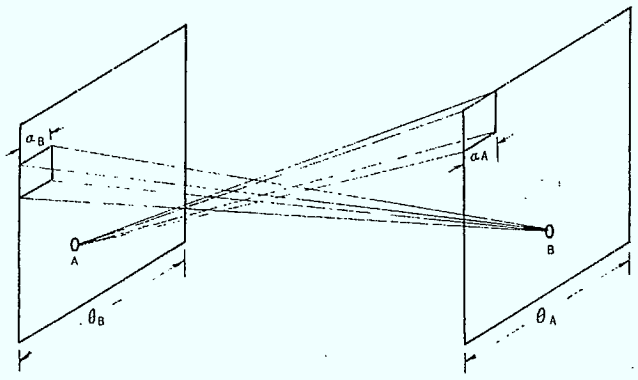


Fig. 1. Geometrical configuration.

methods to increase the SNR to tolerable level is to reduce the total bandwidth of the system. In a superheterodyne system, which we shall assume, this can be achieved by passing the output of the mixer through a narrow-band filter of bandwidth b centered at the IF frequency f_0. If the exact transmitter frequency is not known at the receiving end, the IF produced by an erroneously tuned LO can be so large as to be outside the bandwidth of this filter. It is evident then that the higher the noise in the system, the narrower the filter must be, which in turn could put very stringent tolerances on the correct tuning of the LO. This can be particularly difficult in a laser communications system where absolute stabilities of the order of 1 part in 10^8 could be required. One of the simplest techniques that can be used to decrease this requirement is then to get the transmitter (or LO) to perform a frequency search by employing a sawtooth frequency modulation for one of the lasers. The frequency swept must be large enough to include the complete range of possible frequency uncertainty between the two lasers (Δf). If we assume the acquisition bandwidth b to be small compared to Δf, a pulse will be produced through the filter for each transmitter (or LO) sweep (assuming optical acquisition). If the sweeping frequency is f_s, then the rate of frequency sweep is $\dot{f}=\Delta f \cdot f_s$. The pulse width is then $(b/\dot{f})$. Now the risetime of the filter is approximately $(1/b)$. We must therefore impose the condition that the risetime of the filter is sufficiently short that the signal can build up during the window opening, i.e.,

$$\frac{b}{\dot{f}} = \frac{1}{b} \tag{3}$$

and, by substitution,

$$b^2 = \Delta f \cdot f_s. \tag{4}$$

In terms of the two satellites we are considering, the receiver which is being discussed is the fast scanning or B receiver. Its scanning frame consists of n_B^2 cells. The sweep frequency is related to the frame period T_B by $f_s=(n_B^2/T_B)$. This ensures that during the spatial search, at least one frequency scan will occur during the momentary optical alignment with the other satellite.

By substitution we obtain

$$b^2 = \frac{n_B^2 \Delta f}{T_B}. \tag{5}$$

Now the power SNR for a superheterodyne receiver is given by

$$\text{SNR} = \frac{P_{\alpha_A}}{(\text{NEP}) \cdot b} \tag{6}$$

where NEP is the noise equivalent power of the superheterodyne receiver. Solving for T_B we obtain

$$T_B = n_B^2 \Delta f \left[\frac{(\text{NEP})(\text{SNR})}{P_{\alpha_A}}\right]^2. \tag{7}$$

Finally, the equivalent scanning period T_e (or total period per complete search) is given by

$$T_e = n_A^2 \cdot T_B = \left(\frac{n_B}{n_A}\right)^2 \Delta f \left[\frac{(\text{NEP})(\text{SNR})}{P_{\eta_A}}\right]^2 \left(\frac{\theta_A}{\eta_A}\right)^4. \tag{8}$$

We recall that $n_B=(\theta_B/\alpha_B)$. To reduce T_e, we would like to make n_B as small as possible [from (8)]. This can only be done by reducing the uncertainty θ_B. It would not be advisable to reduce n_B by increasing α_B which is a receiving field of view. This would reduce the resolution and increase the background noise. We shall assume then that $\alpha_B=\eta_B$, i.e., a diffraction-limited field of view. Other ways of reducing T_e for a given SNR comprise a reduction in Δf and in NEP, or an increase in power transmitter. Also it should be noted that a reduction in the uncertainty θ_A would help in a reduction of T_e by a fourth power.

Finally, we note that an increase in the number of transmitting cells n_A would reduce the acquisition time.

In what follows, then, we shall discuss how T_e can be minimized by optimizing the value of n_A. We shall therefore assume that all other terms in (8) have already been optimized for lowest possible T_e.

Since (8) dictates the highest possible value in n_A for low T_e, let us find out what are the limiting conditions.

1) *Bandwidth Limit:* The bandwidth of the filter cannot usefully be larger than the system's bandwidth. This implies that

$$T_e \geq \left(\frac{\Delta f}{b_{\max}^2} n_A^2 n_B^2\right). \tag{9}$$

2) *Mechanical Limit:* The mechanical optics scanning frequency of B is given by $f_m=(f_s/n_B)$. It is evident that mechanical limitations will impose a maximum value for f_m and so, by substitution,

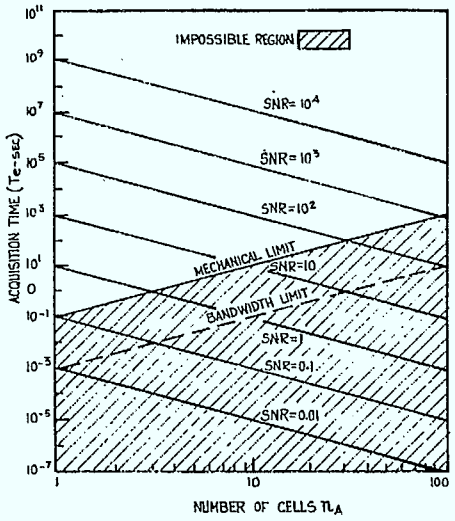


Fig. 2. Effective scan period versus number of cells n_A for various SNR.

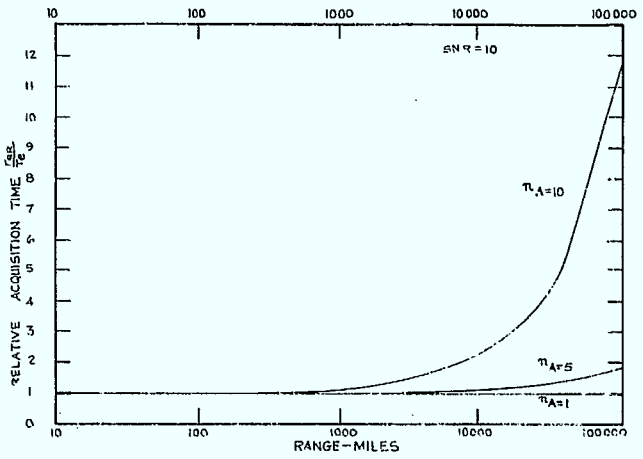


Fig. 3. Effect of range on acquisition time.

$$T_e \geq \frac{n_A^2 n_B}{f_{m_{max}}}. \tag{10}$$

In order to gain some insight into the theory, we shall discuss a typical case that could be met in space by using reasonable assumptions.

Let the communications system be a CO_2 laser superheterodyne system at 10.6 μ with an uncertainty in the IF, $\Delta f = 10$ MHz, a power NEP of 10^{-19} W/Hz, narrow fields of view of $\eta_A = \eta_B = 0.002°$, and also $\theta_A = \theta_B = 0.2°$, and $P_{n_A} = 10^{-10}$ watts. Using (8), we can then plot a family of curves giving the effective acquisition time T_e versus the number of cells n_A for various SNR. This is shown in Fig. 2.

In order to impose limits 1) and 2) we shall further assume that $b_{max} = 10$ MHz and that $f_{m_{max}} = 1$ kHz. Equations (9) and (10) are then plotted onto the same graph. Since the regions under these curves are impossible, it is evident that in the case we chose, the system is mechanical rather than bandwidth limited. This is generally true. We also note, that for a given SNR the shortest acquisition time occurs in general for the case where $n_A > 1$. In words, this means that although it might appear that spreading the search beam to cover the complete region of uncertainty (θ_A^2) woud allow a faster acquisition time than to narrow the beam and search, in actual fact this is usually not true. The reason is that although the wide beam system gains in search time on geometric grounds by the factor n_A^2 compared to a narrow beam system, it loses on power grounds by a factor of n_A^4.

Effect of Range on Acquisition Time

So far we have ignored the effect of transit time on acquisition. To take this into account, we note that the frame period T_B should be increased to include the transit time required for the beam of A to reach B at the beginning of each step scan. Similarly, that same transit time should be added at the end of each step scan. This is to cover the case when the acquisition cell happens to be the last B cell that would occur before A step scans again.

Time should then be given for the acquisition information to travel from B to A thus stopping the step scanning action. If R is the range between the two satellites, then an extra time of $(2R/c)$ should be added to T_B, where c is the velocity of light. In addition, we shall add an extra constant period T_c to cover the small time required to step scan from one cell to the next. The total effective frame period T_{Be} is thus

$$T_{Be} = T_B + \frac{2R}{c} + T_c; \tag{11}$$

thus

$$T_{eR} = n_A^2\left[T_B + \frac{2R}{c} + T_c\right] \tag{12}$$

where T_{eR} is the effective scan time when the range is then taken into account.

To get an insight into the effect of range, we have plotted in Fig. 3 the relative increase in acquisition period, or (T_{eR}/T_e), as a function of range for a SNR = 10 and various values of n_A. This was done for the same numerical values of the parameters that were used in the calculations of Fig. 2. It is seen that for higher n_A, the effect of range occurs sooner, and that below 1000 miles, its effect is minimal. For higher distances, when the discrepancy becomes appreciable, (12) should therefore be used in calculating the effective acquisition time.

Conclusion

We have examined a laser acquisition system using heterodyne detection in which both stations scan spatially to acquire the other. This generalized system was compared with one in which any one of the stations searches spatially while the other remains passive and illuminates the entire search area. It was found that for a given SNR, the former method can usually provide substantially shorter acquisition time.

A. Waksberg
RCA, Ltd.
Research Laboratories
Montreal, Quebec, Canada

6.0 SYSTEM DESCRIPTION

The fundamental elements of a semiconductor laser heterodyne communication system were presented in Figure 2.2.1. Each element will be discussed briefly below. The first five elements are located in the transmitter, while the remaining seven can be found in the receiver.

1) Modulator: The modulator accepts the input data signal and performs any special encoding or multiplexing functions that may be required. The modulator then supplies the laser driver circuit with a trigger pulse appropriate for the modulation format that has been selected. The major technical problems encountered with the modulator are high speed design (GHz bandwidths) and pulse jitter.

2) Laser Driver: Actual current modulation of the laser is performed by the laser driver circuit. However, if external modulation of the laser beam is performed, then the laser drive would supply the external modulator with a voltage signal. Further discussion will consider only direct current modulation. The driver must supply current pulses of the proper amplitude and format. These pulses may be very small as would be the case if FSK were used, or they may be very large as may be the case in an ASK system. Whatever the system chosen, the driver must have the specified power capability and the necessary bandwidth. It is also necessary that the pulses display low jitter and distortion.

3) Laser: For the reasons discussed previously, GaAlAs lasers operating in the .8 μm - .85 μm region seem to be the most

promising. The advantages offered by these lasers for space systems include small size and weight, low complexity, direct modulation capability, power efficiency, high capacity for integration, reliability and relatively low cost. However, these lasers also possess certain disadvantages such as relatively low output power, especially for lasers of the required spectral purity, lasing frequency in stability and beam nonuniformity. Promising solutions to these problems have been proposed, although further research and development is still required to bring them to fruition.

4) Laser Support: Laser support functions refers to all subsystems required to ensure successful operation of the laser and yet are not involved in the signal generation. These subsystems include a stable, low noise, tunable bias current supply, laser package temperature controllers, and any necessary laser frequency stabilization controls. These subsystems may be electronic, optical, mechanical, or some combination thereof.

5) Transmitting Optics: Special optics are needed to collimate the highly divergent laser beam. The emitting aperture of a semiconductor laser is generally a few tenths of microns in the transverse direction by up to ten microns in the lateral direction. Thus, the farfield pattern in the planes perpendicular and parallel to the junction experience different degrees of divergence resulting in a elliptic beam profile. Typical divergence values are

$10\text{-}15^{\circ} \times 25\text{-}30^{\circ}$. Such a condition places two important constraints on the transmitter optics. First, the collimating lens must have a high numerical aperture in order to collect a higher fraction of the emitted laser radiation. This results in a short focal length and working distance for the lens and hence the need for a very stable lens mounting assembly to guarantee stability of the laser/lens separation. Secondly, there may be a strong requirement for beam circularizing elements so that the beam spot at the receiver is as uniform as possible. This condition is required to obtain efficient mixing with the local oscillator at the beam combiner and so minimize amplitude noise caused by pointing error. Following collimation and circularization, beam expansion is needed to minimize the extent of diffraction of the laser beam. Expansion from beam diameters of 5 mm to 20 cm may be considered typical. Some spatial filtering of the beam may be easily incorporated if necessary. In addition, pointing and tracking optics will be required which will certainly include scanning elements. Further modification will be needed if certain optical frequency stabilization techniques are used and if the system is designed for full duplex operation (i.e. transmitter/receiver sharing the same or parts of the same optical assembly). Care must be taken to prevent any reflections from re-entering the laser cavity. This necessitates the antireflection coating of certain critical optical components. If this problem proves to be serious, then an optical isolator may

be needed in the optical assembly. Finally, the lenses should be designed for diffraction limited operation at the lasing wavelength.

6) Receiving Optics: The receiving optical assembly must be designed to collect as much signal power as feasible and after mixing with the local oscillator beam, focus it onto an optical detector. Many of the same considerations that apply to the transmitting optics apply also to the receiving optics with the following additions. It is important that the entrance aperture of the receiving telescope be as large as possible, within physical limitations. This is necessary in order to collect as much signal power as possible. In an effort to restrict background optical radiation from impinging on the detector and thus becoming a source of noise, a "narrow-band" optical filter is incorporated into the receiving optics. Baffles would likely be incorporated in order to restrict the field-of-view of the receiver and hence limit exposure to bright point sources such as the sun, moon, earthglow, etc. The receiver optical assembly would also include the collimating optics, etc., for the local oscillator.

7) Beam Combiner: The optical signal is mixed with the local oscillator at the beam combiner. The simplest beam combiner is a conventional beamsplitter cube. More sophisticated beam combiners can be made using fiber optic couplers or integrated optic waveguides. The beam combiner must provide good spatial and polarization

alignment of the two beams. Mismatch of the wavefronts must be minimized for efficient mixing to occur.

8) Detector/Preamp: A square-law optical detector is used to recover the message signal from the optical beam. The most attractive devices are semiconductor APD and pin photodiodes. An attractive feature of APD's are their intrinsic gain. The pin photodiode is restricted to unity gain, however it is a generally higher speed device. To compensate for the unity gain limitation, it is common in fiber optic receivers to integrate the detector with a GaAs FET amplifier. The two most critical requirements imposed on a detector for high speed, long distance communications is large bandwidth and high sensitivity. Unfortunately, high bandwidth necessarily means lower sensitivity. Therefore, optimum compromises must be found. For coherent communications approaching 1 Gbit/s, detector bandwidths approaching 10 GHz may be a requirement. Silicon and GaAs detectors are used to match detector peak quantum efficiency with laser wavelength. Optical alignment tolerances can be eased with the use of large area detectors, however this usually implies a reduction in bandwidth. If the availability of high voltage sources in satellite application is low, the APD's may be at a disadvantage since they typically require bias voltages ranging from 100 V - 450 V. Regardless of which detector structure is chosen, a low noise front end amplifier immediately following the photodector will be a requirement. Although high

impedance front ends offer superior sensitivity, they may not be capable of operation at the required high speeds. Therefore, transimpedance front ends may prove to be a better choice. IF filters, centered at the heterodyne frequency, would also be used.

9) Demodulator: The demodulator recovers the baseband signal from the received signal. Therefore the demodulator will require clock recovery circuitry, decision/regeneration circuitry and any decoding or demultiplexing circuitry. A major technical problem, similar to that encountered with the modulator, is the capability of the electronics to operate at the necessary high speeds and levels of integration.

10) AFC Loop: The automatic frequency control loop is necessary in achieving the required long-term stability in the IF signal. The loop adjusts the local oscillator frequency in accordance with the drift in the received signal carrier. Implementation is performed with a frequency discriminator and a low pass filter. The error signal derived from the frequency discriminator provides a reference signal to the local oscillator bias electronics. The local oscillator is then current tuned in such a manner as to keep the IF signal centered within the band pass of the receiver IF filter.

11) Local Oscillator: Another laser similar if not identical to the transmitting laser is used as the local oscillator. The transmitting laser and the local oscillator should be well matched in terms of wavelength, mode pattern, and

beam quality to enable efficient mixing. Thus, most of the discussion regarding the transmitter laser also applies to the local oscillator. The most significant difference, however, is in the method of operation. The transmitter laser must be capable of high speed modulation while the local oscillator will be operated in a cw fashion, requiring only sufficient bandwidth to be able to accept the feedback signal from the AFC loop. Thus, although the laser chips themselves may be identical, their packaging may be entirely different.

12) L.O. Support: The local oscillator support subsystems are the same as those discussed for the transmitter laser. Although actual implementation may differ, the same considerations apply.

6.1 Experimental System

The experimental system was assembled such that it could accomodate ASK baseband and heterodyne shift keying and FSK shift keying modulation formats. The heterodyne mode of operation used noncoherent, or envelope detection. Effort was concentrated on the FSK heterodyne single filter envelope detection, although ASK could be employed by a change in the transmitting laser operating condition. The ASK baseband mode of operation could be implemented simply by blocking the local oscillator beam.

A testbed schematic is given in Figure 6.1.1. The AFC loop components included the filter, the frequency discriminator, the LPF and amplifier. The frequency discriminator components

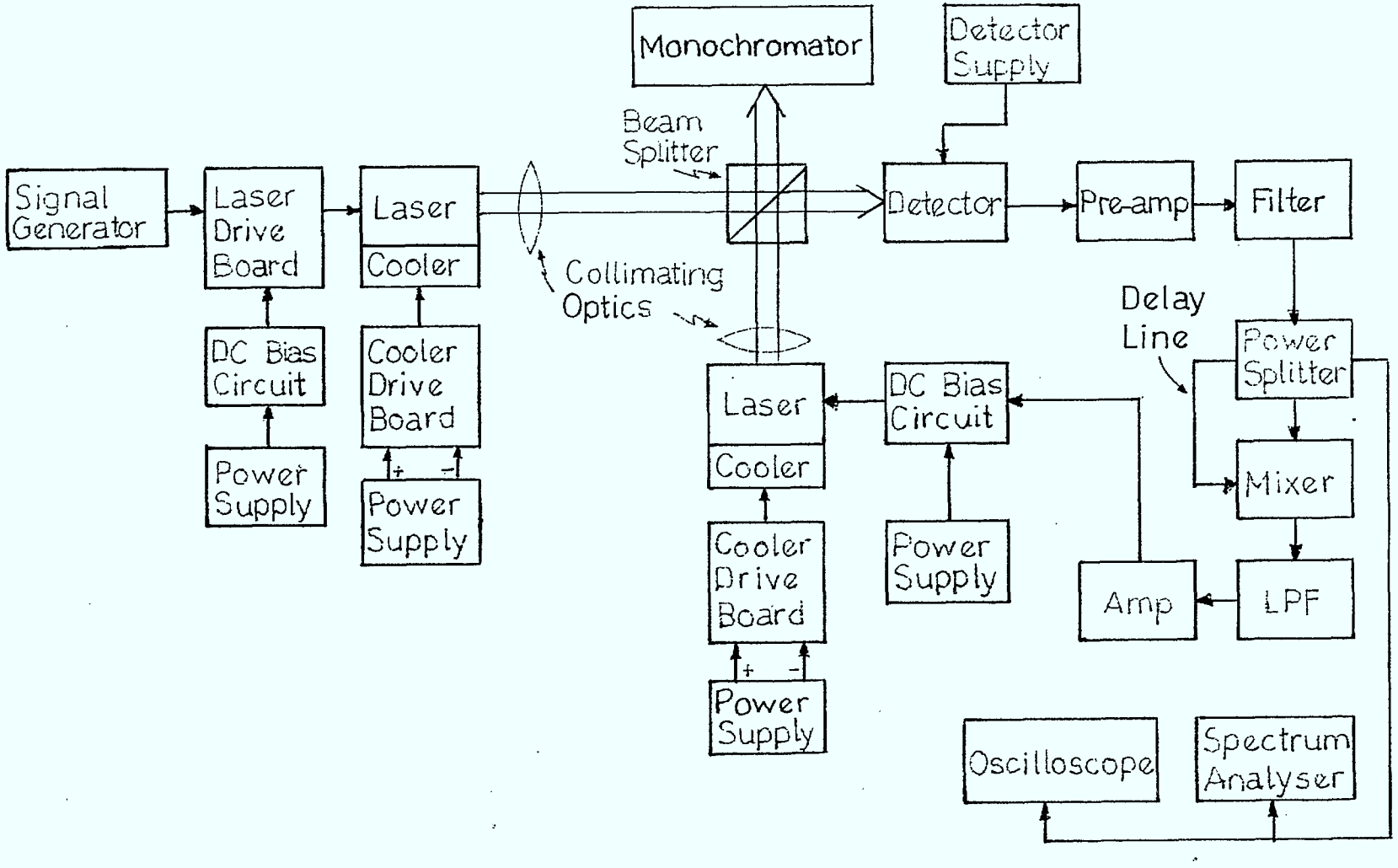


FIGURE 6.1.1: Testbed Schematic Diagram

included the power splitter, the mixer and the delay line. For measurements performed without the AFC, the output of the preamp fed directly into either the oscilloscope or the spectrum analyzer. The spectrum analyzer served as the envelope detector in the system. The collimating optics and the detector were mounted on precision three-axis positioners. Three baseplates were employed; one at the transmitter, one at the local oscillator and one at the beamsplitter/detector assembly. The beamsplitter was mounted on a precision z-translator and rotator. The signal generator served as the modulator. It produced a square wave pulse of adjustable bit rate, duty cycle and amplitude. Each component part is identified in Table 6.1. Further detailed discussion is given in Chapters 7 and 8. The entire experiment was supported by a floating granite optical table. Photographs of the system are shown in Figure 6.1.2.

6.2 System Operation

The transmitting laser was modulated by a 500 Mbit/s, 50% duty cycle, pulse from the signal generator. This pulse stream simulated a binary 1010 1Gbit/s NRZ message signal. The pulse amplitude and laser dc bias correspond to the desired modulation format. For ASK operation, the bias was set close to threshold, while the pulse amplitude was at its maximum (i.e. about 20 mA for this pulser). For FSK operation, the bias was set high to maximize power and minimize linewidth, while the pulse amplitude was attenuated to the milliamp level. Outside of the difference in implementation, there is essentially no difference in the ASK detection and in FSK single filter envelope detection as far as the receiver is concerned. ASK and FSK spectrums are shown in

TABLE 6.1: Parts Listing

Element	Part Number	Description
Signal Generator	Colby PG 1000A	Pulse Generator
Laser Drive Board	*	
DC Bias Circuit	*	
Laser	Hitachi HL8311E Hitachi HLP1400	
Tx Cooler	Melcor CP 1.4-71-10L	
Cooler Drive	*	
Tx Collimating Optics	NSG W-1.8-0.23-B/BC-0.83 Melles Griot GLC-002 Melles Griot GPA-001 Melles Griot GBX-001	Grin-rod lens Collimating lens Anamorphic Prism Pair Beam Expander
Beamsplitter	Optikon 33 5520	Beamsplitter Cube
LO Collimating Optics	NSG W-1.8-0.23-B/BC-0.83	Grin-rod lens
Local Oscillator	Hitachi HLP1400	
LO Cooler	Melcor CP 1.4-31-10L	
Cooler Drive	*	
DC Bias Circuit	*	
Detector	Antronics AR-S5 Antronics AR-S2 RCA C309506	APD PIN APD
Pre-Amp	Miteq AFD3-020040-30	Low Noise Amplifier
Filter	RLC F10-4000-R	DC-4 GHz Filter
Power Splitter	R.F. Z43PD-4	
Mixer	R.F. ZFM-4212	Double Balanced Mixer
LPF	*	
Amplifier	*	
Positioners & Mounts	Linetool Model A Model J Model A (LH) Newport T1	X-Y-Z Prism Table

Table 6.1: Parts Listing - cont'd.

Element	Part Number	Description
Measuring Equipment	HP 8569A	Spectrum Analyzer
	Tek 7834	Storage Oscilloscope
	7A29	High Speed Amplifier
	7B92A	Dual Time Base
	7S11	Sampling Unit
	7T11	Sampling Sweep Unit
	Scientech 360203	Power Meter
	Jarrel Ash Czerny-Turner	Scanning Spectrometer: 2 meters
Power Supplies	HP 6218A	Tx Laser Supply
	6234A	LO Supply
	Lambda LO0422FM	Cooler Supplies
	Power Designs Pacific 3K10	High Voltage Detector Supply

* - Denotes elements that were designed at MPB Technologies Inc.

Note that a variety of general lab supplies such as optics holders, sma connectors, digital voltmeters, etc. were used. Custom design laser mounts and a custom designed chamber for the local oscillator were also machined.

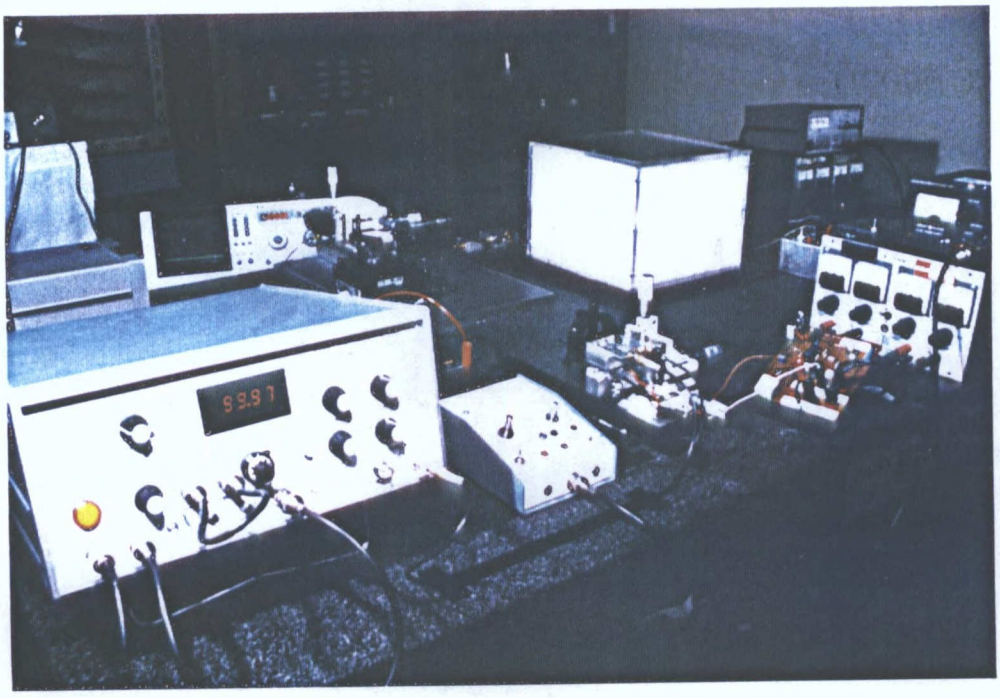


FIGURE 6.1.2(a): Experimental testbed. This photograph shows the transmitter in the foreground with bias circuitry in the box to the left of the laser assembly and the cooler control circuitry on the right. The local oscillator is housed in the white chamber. The detector is mounted on the three axis positioner immediately in front of the spectrum analyzer. The beamsplitter cube is immediately in front of the detector.

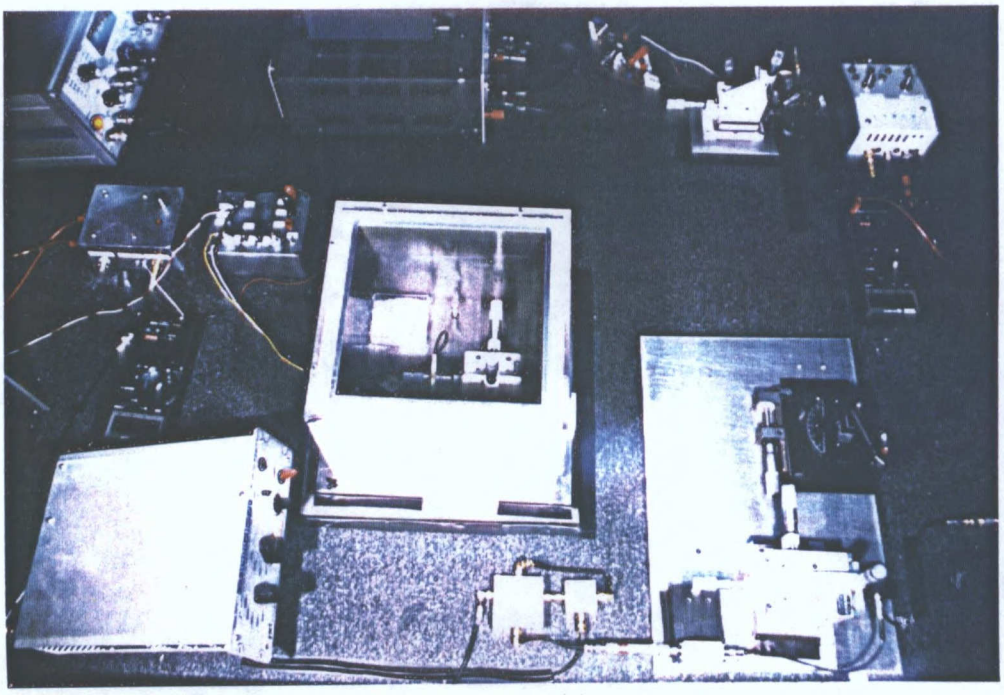


FIGURE 6.1.2(b): Experimental testbed (top view). The frequency discriminator components can be seen at the bottom of the picture next to the local oscillator chamber.

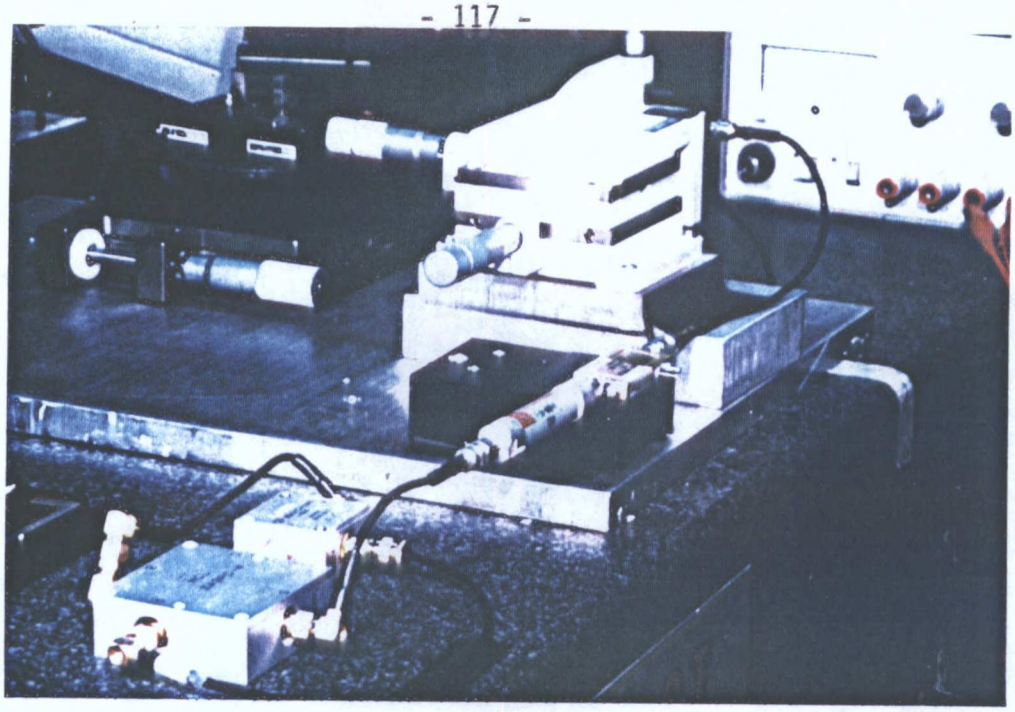


FIGURE 6.1.2(c) This photograph shows the elements of the frequency discriminator. Following the detector comes the pre-amp, the filter, the power splitter and finally the balanced mixer. The coaxial delay line between the splitter and the mixer can be seen to the left of these components.

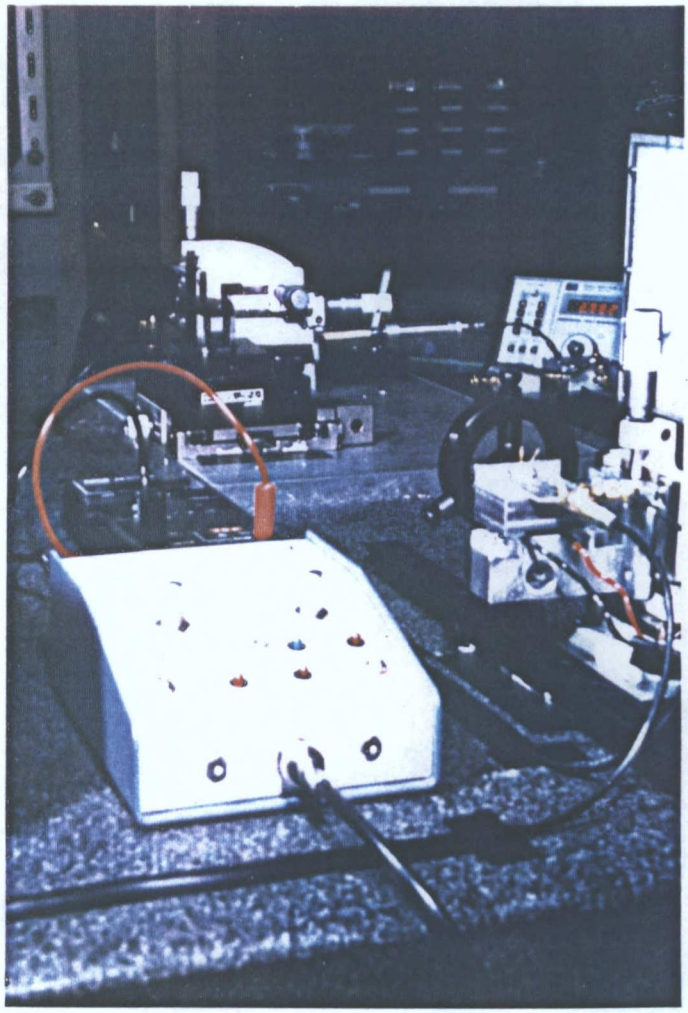


FIGURE 6.1.2(d): Experimental testbed looking near the optical axis from the transmitter side. The laser dc bias box is in the foreground.

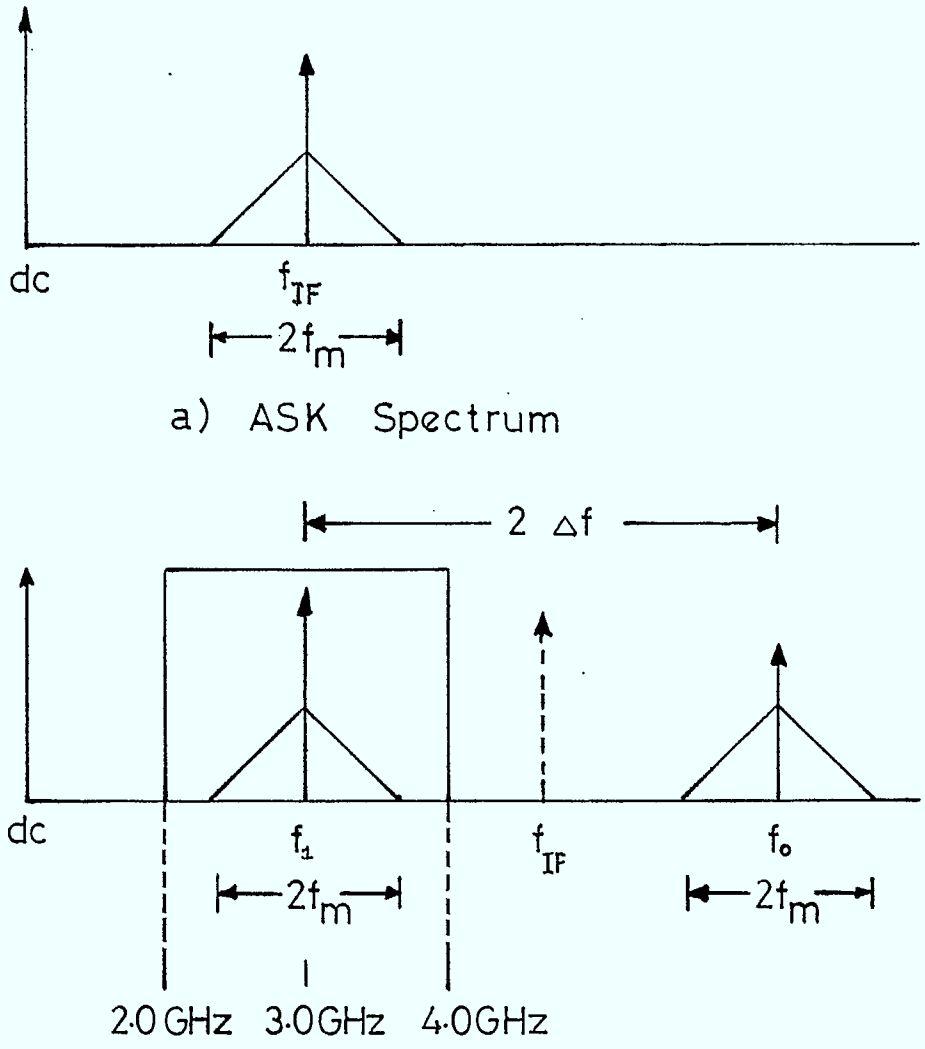


FIGURE 6.2.1: Theoretical Spectrum of the ASK and FSK Modulated Signal After Optical Detection

Figure 6.2.1. In the ASK scheme signal is present at the beat frequency when a one is transmitted. No signal is present when a zero is transmitted since the laser is turned off. Demodulation is performed by a bandpass filter with bandwidth corresponding to twice the message bandwidth (i.e. 2 GHz for a 1 Gbit/s bit rate) centered at the IF frequency and an envelope detector followed by a threshold detection circuit. In an FSK system, signal is always present as indicated in 6.2.1(b). The dotted f_{IF} spectral line corresponds to the beat frequency when no modulation is performed. When a one is transmitted the beat frequency shifts to f_1, while when a zero is transmitted the beat frequency shifts to f_0. There is a small difference in line strength due to a small AM modulation. For frequency discrimination or dual filter detection, receiver bandwidths corresponding to $2(\Delta f + fm)$ are required as determined by Carson's rule. However, if single filter detection is used the same demodulation technique as used for ASK can be employed. The trade-off is a degradation in performance to the ASK heterodyne level, in favor of ease of implementation. In order to avoid spectrum folding problems and problems in the baseband interference, the cw IF frequency should be set at least four times greater than the message bandwidth.[6.1] Therefore for dual filter, or frequency discrimination demodulation and a 1 Gbit/s transmission rate, a receiver bandwidth of at least 6 GHz would be required. In a practical system this figure would probably be closer to 10 GHz. However, for single filter detection the required receiver bandwidth would be reduced to 4 GHz. The detector used in this experiment has a measured bandwidth of 3.5 GHz. Therefore,

because of this reduced bandwidth requirement and ease of implementation, FSK single filter detection was chosen. The system was operated such that the f_1 frequency component occurred at 3.0 GHz. The operating window was 2.0 GHz to 4.0 GHz.

7.0 TRANSMITTER

The functional elements of the testbed transmitter were given in Figure 2.2.1 and briefly discussed in Chapter 6. In summary, these elements include the laser, dc bias drive, modulation electronics, cooler circuitry and optics. Also included in the transmitter were custom designed laser mounts and a laser bias board for combining the dc and modulation signals at the laser.

The transmitter is capable of operation in both ASK and FSK modes. Input pulse amplitude in ASK systems is limited to 20 mA or approximately .36 of threshold current. The objective of the transmitter design was to develop the simplest system capable of providing cw heterodyning with a high degree of flexibility with respect to the modulation format. Performance evaluation using signal-to-noise ratio measurements and general assessment of ease of implementation were prime objectives. Effort was largely concentrated on the laser diode behaviour since it represents the most critical component in the transmitter design.

7.1 Laser

It is important that the laser satisfy certain criteria as discussed in previous sections. These criteria result in the following fundamental laser specifications:

1. Maximization of output power
2. Single mode operation
3. Short wavelength operation
4. High speed packaging

5. Flat FM modulation in the region of interest.

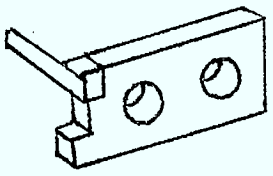
Based on these general criteria a transmitter laser was selected. Initially, the transmitter was operated with a Hitachi HLP 1400 laser diode. This laser was later replaced with a Hitachi HL8311E. Both lasers are essentially the same, differing only in their packaging. The laser packages and their dimensional outlines are shown in Figure 7.1.1. The HLP 1400 has the advantage of easy optical access to both laser facets. This ability facilitates operation in an external cavity configuration. The initial design of the transmitter was such that this configuration remained an option. However, because the diode is not hermetically sealed, it is exposed to a generally hostile environment (i.e. airborne volatiles, dust, humidity, damage from physical contact, etc.).

Both the HLP 1400 and the HL8311E lasers are fabricated from GaAlAs and have a channelled substrate planar structure as shown in Figure 2.4.3. These lasers thus tend to single mode operation by virtue of their index guiding property. They are rated for 15 mW cw operation corresponding to a current level of approximately $2I_{th}$ where I_{th} is the current threshold of the laser. Measured threshold levels are in the range of 55 mA at room temperature. Beam divergence in a plane parallel to the junction is specified at 10^{o} for an output power of 10 mW while the beam divergence in a plane perpendicular to the junction is specified at 25^{o}. Typical farfield patterns are shown in Figure 7.1.2

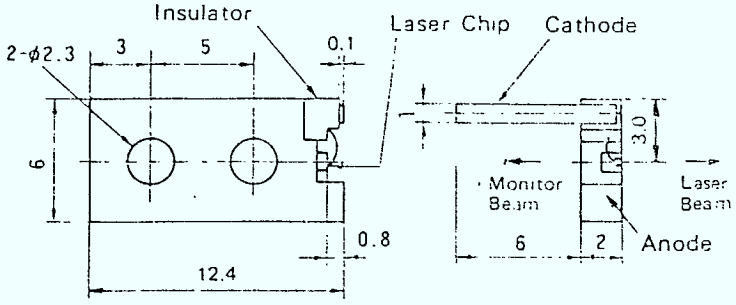
Continuous wave (cw) spectra for the HLP 1400 are shown in Figure 7.1.3. Spectra are given for operation just below

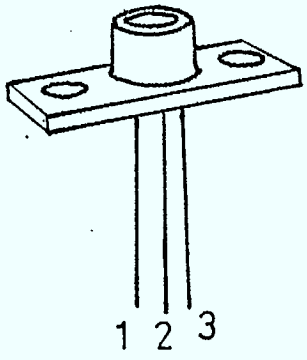
Package

Package Dimensional Outline (mm)



HLP1400





HL8311E

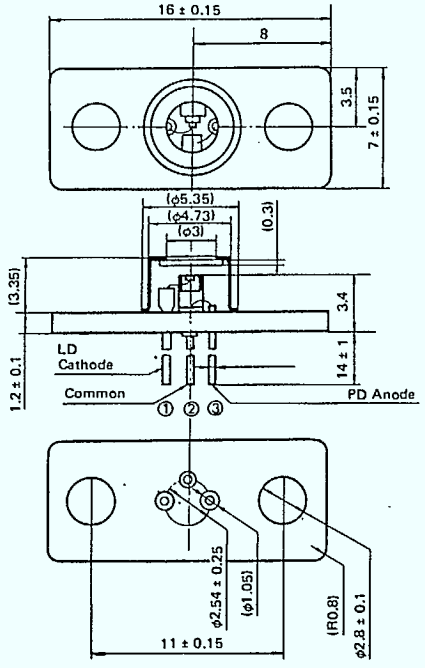


FIGURE 7.1.1: Laser Packages and Dimensional Outlines

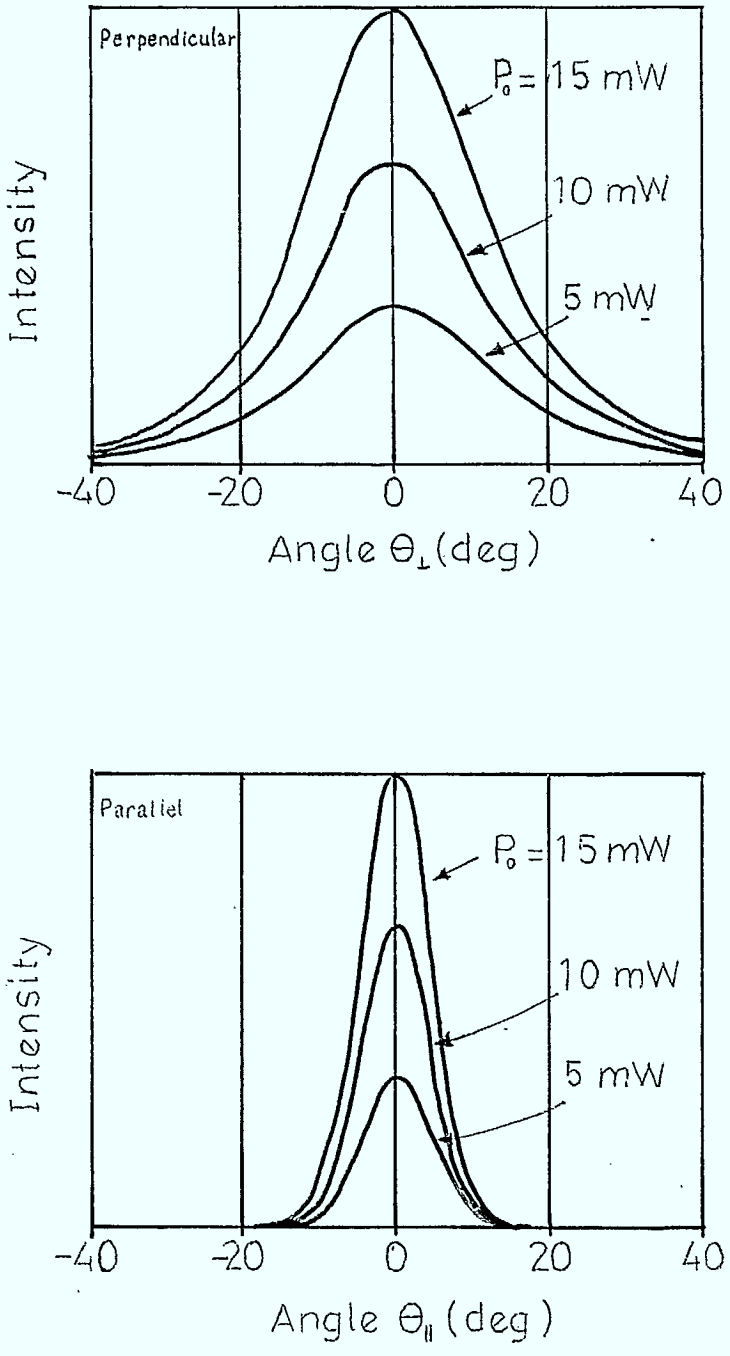


FIGURE 7.1.2: Typical Farfield Patterns for the HLP 1400 and HL8311E Lasers

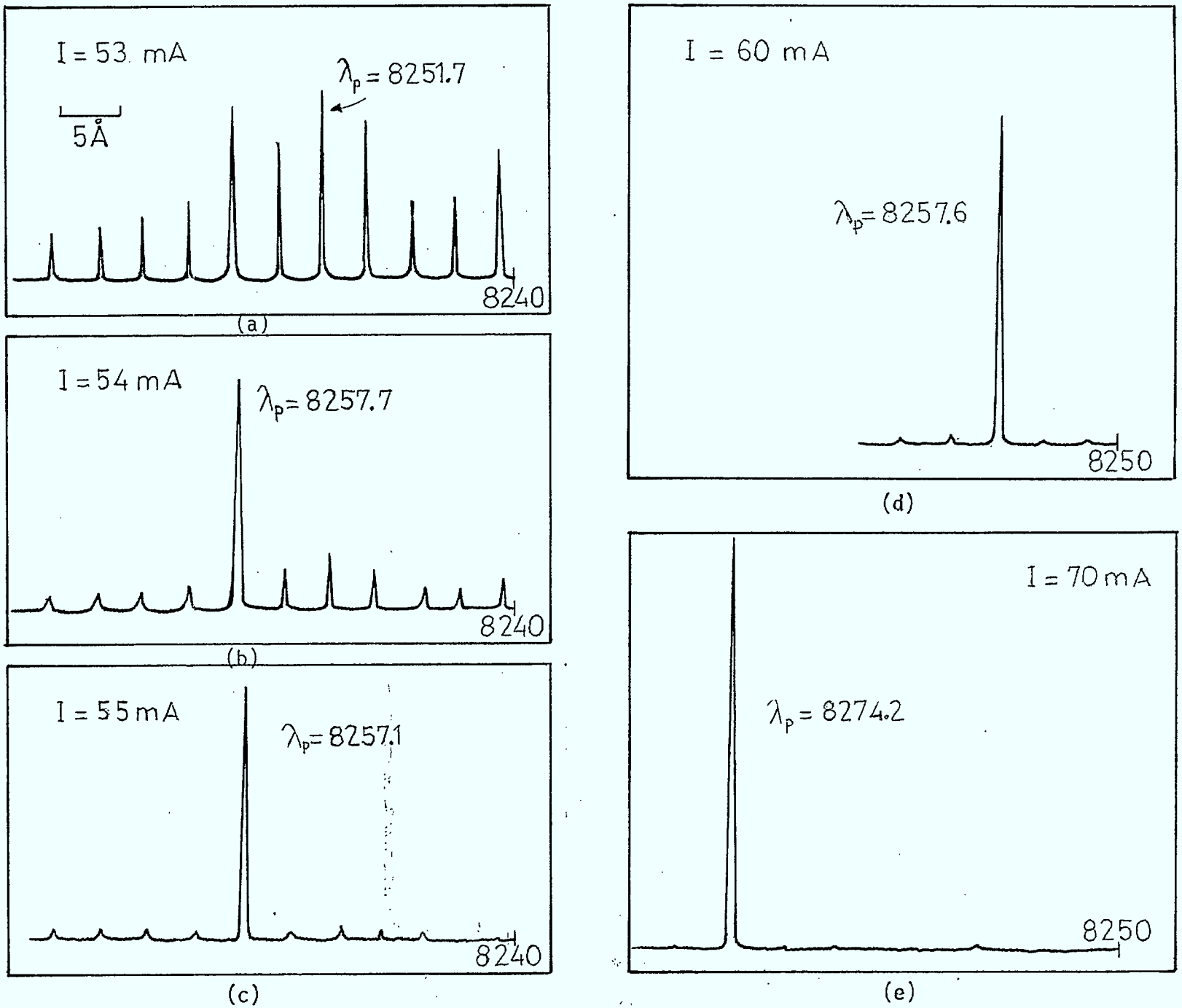


FIGURE 7.1.3: Continuous Wave (cw) Spectra for the HLP 1400

threshold, at threshold, and well above threshold. Near threshold the laser can be seen to display multimode behaviour. This is an inherently noisy region of laser operation due to extensive mode competition. However, at currents of 60 mA and 70 mA single mode operation has been well established. A curious mode shift occurs between traces (d) and (e). This behaviour was further investigated and the results are shown in Figure 7.1.4.

When single mode operation is first established at about 1.1 I_{th}, the lasing wavelength was measured to be about 8257 Å. As the drive current was increased, and hence the junction temperature, the lasing mode tended to longer wavelengths as expected from a drift in the gain spectrum caused by a decrease in bandgap due to increased junction temperature. At the same time the laser is experiencing a modification of its cavity refractive index, also stemming from the change in temperature. These drifts are about 20-25 GHz/^{o}C for these lasers. However, at a certain current level mode hops spanning four modes occur. This same phenomenon occurs again after a further current increase of about 10 mA. This behaviour is characteristic of CSP lasers, and has been attributed to a distortion in the gain spectrum caused by spectral and spatial hole burning due to a finite thermalization time of injected carriers.(7.1) This mechanism results also in a hysteresis effect so that the lasing wavelength is determined not only by the temperature and drive current, but also by the manner in which they were arrived at (i.e. current/temperature increased or decreased).

It is fairly well known that if the instantaneous current

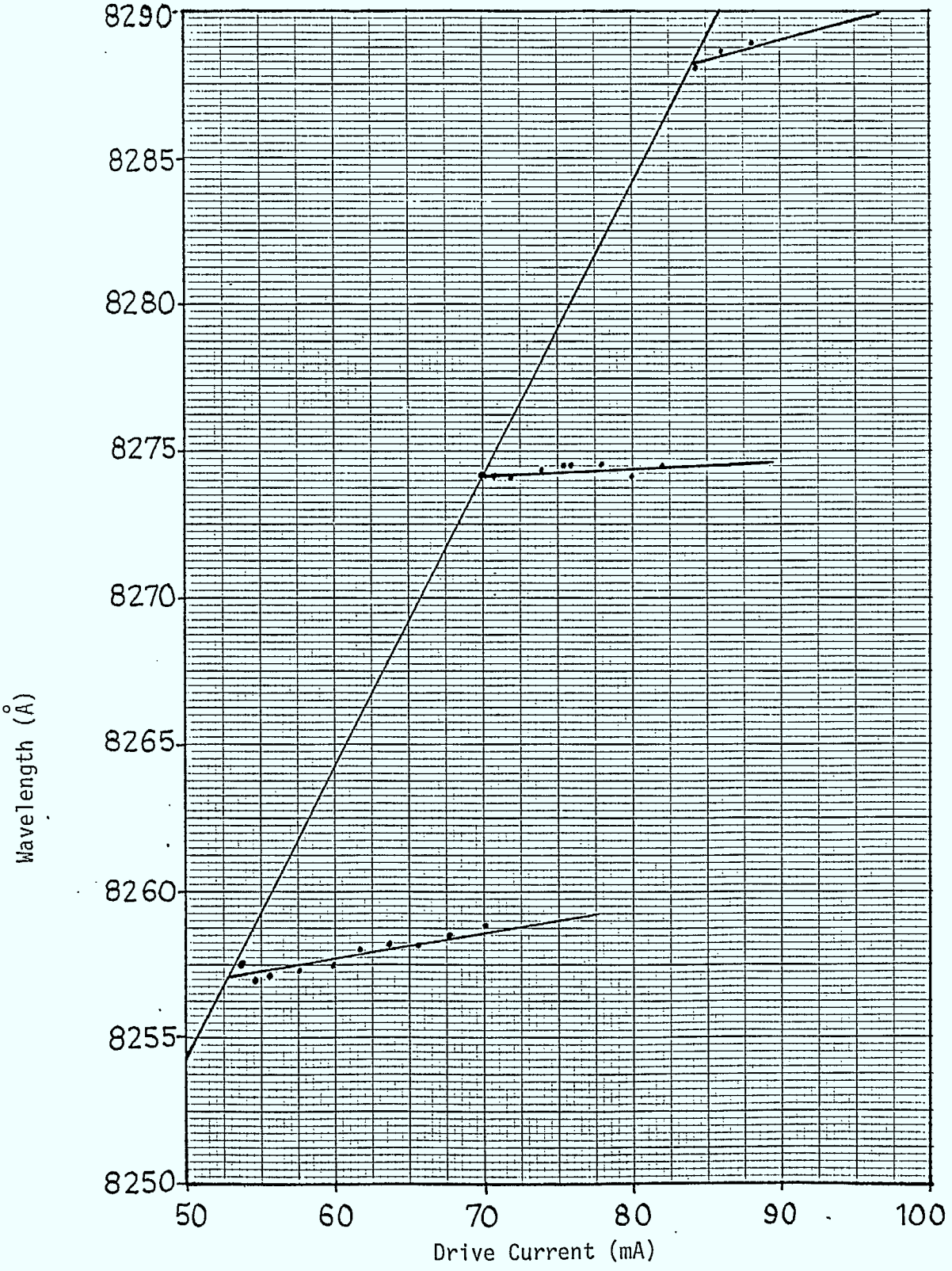


FIGURE 7.1.4: Peak Wavelengths for Increasing Bias Currents for the HLP 1400

of modulated lasers drops below threshold, then normally single mode lasers will display multimode behaviour. The extent of this behaviour is sensitive to the frequency and amplitude of modulation and the bias level. The behaviour can be attribtued to an overshoot of the threshold condition by the gain spectrum. When this happens many satellite modes may momentarily experience net gain and hence begin to lase. Gain overshoots are enhanced by high frequency modulation, high pulse current and low bias levels. At low frequency modulation (i.e. < 1 MHz), a temperature modulation of the gain spectrum may also result in a multimode spectrum. Conversely it has been observed that for a laser in cw multimode operation, the application of modulation can induce single mode operation.

This type of behaviour has been observed in the HLP 1400 laser as can be seen in the pulsed spectrum given in Figure 7.1.5. This figure shows pulsed spectra at three bias current levels and at three pulse frequencies. Series (a) shows spectra for bias near threshold. Comparison of the 50 MHz spectra with cw given in Figure 7.1.3, shows an enhancement of single mode operation, however at higher frequencies multimode operation is enhanced. At 1 GHz near the resonance peak, evidence of high frequency chirping can be seen which is caused specifically by the gain modulation effect. However, when the bias level is adjusted well above threshold level, as in series (b) and (c), there is a suppression of general multimode behaviour. At lower pulse frequency levels some two and three mode operation can be seen. This behaviour is a result of the mode hopping behaviour characteristic of the CSP laser as discussed earlier. When the

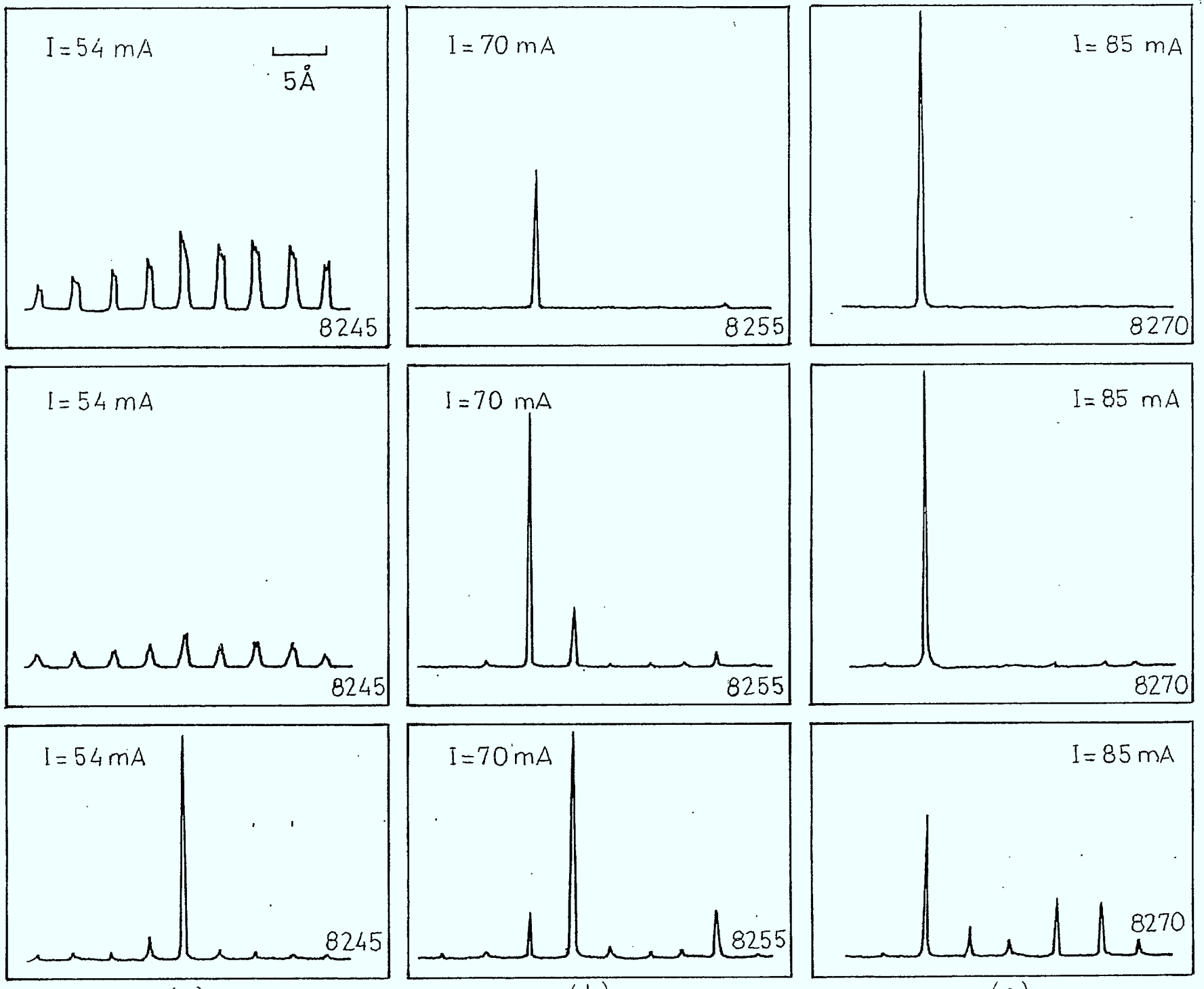


FIGURE 7.1.5: Pulsed Spectra

modulation is such that the current level drops from one plateau to another, there will be a shift in the lasing mode. The measured strength of this mode will be proportional to the time that the mode is lasing since it is time averaged spectra that has been measured. Lasing of a satellite mode adjacent to a peak mode such as seen in series (b), is most probably due to partition noise resulting from the gain peak being centered somewhere between the modes. Operation at high frequencies and at high bias levels and at pulse currents that do not drive the laser into different operating plateaus still retain the single mode behaviour of the laser as witnessed in Figure 7.1.5.

Heterodyne operation was performed in an FSK fashion, hence pulse current amplitudes were no more than a few milliamps, at laser bias currents near 80 mA.

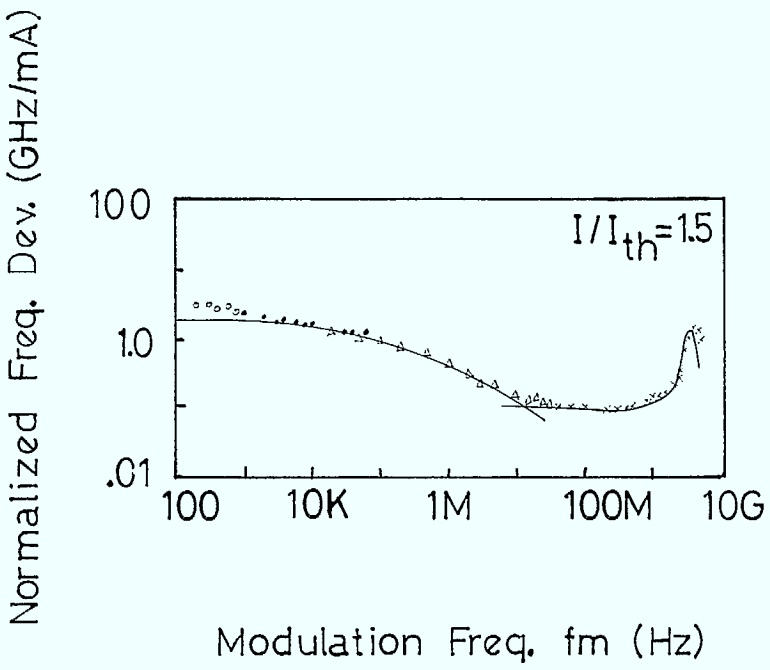
The dynamic behaviour of semiconductor lasers is governed by a pair of coupled rate equations which describe the carrier and the photon populations within the laser cavity. At modulation frequencies in excess of 1 GHz the equations become out of phase and thus display resonant behaviour. The peak of this resonance is very much influenced by the physical parameters of the laser, and also by the operating point of the laser. The resonance peak can be shifted to higher frequencies by an increase in the bias current. As noted, the refractive index is strongly influenced by the carrier concentration through the free carrier plasma effect and anomolous dispersion. Therefore the refractive index of the laser active region can be modulated in a fashion corresponding to the gain modulation, which leads to FM modulation of the laser. A peak in the frequency modulation

efficiency of the laser should be expected at the laser resonance frequency as it is in the amplitude modulation response. A frequency modulation curve for a CSP laser is shown in Figure 7.1.6.[7.2] Also shown is the shift due to increased bias current levels. The frequency modulation response is flat between approximately 10 MHz and 1 GHz. Below 10 MHz, temperature modulation of the refractive index is experienced. The modulation efficiency of the laser is reduced at higher bias currents. This behaviour is demonstrated in Figures 7.1.6(a) and (b).

AM and FM intrinsic laser noise resulting from quantum process inside the laser cavity (i.e. laser shot noise), is seen to display a similar spectrum as will be discussed later in the report.

For system modulation at frequencies in the GHz region this resonance peak can become a serious problem because of a not flat FM modulation response and noise enhancement. It is therefore important to force the peak to as high frequencies as possible and to reduce the intrinsic noise level perhaps through the use of dual balance detectors. The enhanced low frequency modulation response can be suppressed by the appropriate use of equalizers in the laser modulation circuit.

Equally important to heterodyne success is the frequency stability and linewidth of the lasing mode. Traditionally laser linewidth was predicted by the well known Schalow-Townes formula. This relation demonstrated a Lorentzian lineshape with a FWHM inverse power dependence. However, experimentally measured linewidths were constantly found to exceed the calculated value.



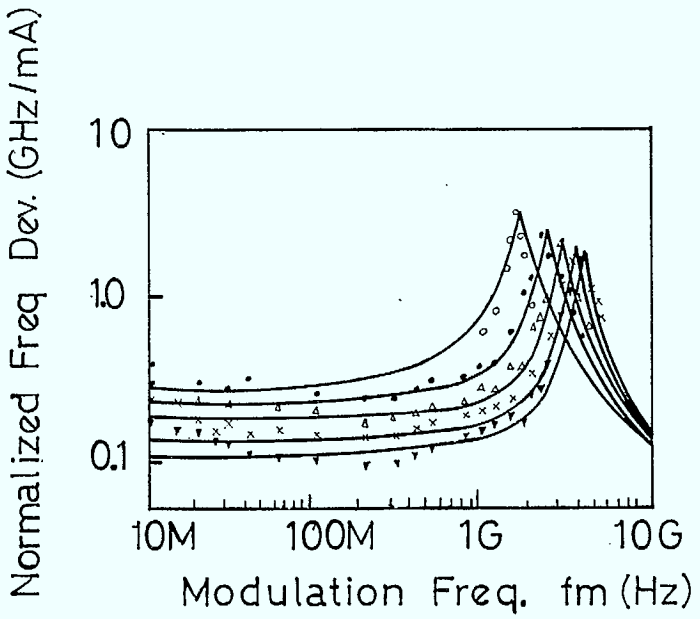


FIGURE 7.1.6: CSP Frequency Modulation Curves

Recently, Henry has demonstrated that the linewidth of semiconductor lasers is increased by a variation of the real refractive index with carrier density.[7.3] The width of the laser line can be thought of as being due to fluctuations arising from spontaneous emission events which discontinuously alter the phase and intensity of the lasing field. As Henry points out: "Besides the instantaneous phase change caused by spontaneous emission, there will be a delayed phase change resulting from the instantaneous change in field intensity. To restore the steady-state field intensity, the laser will undergo relaxation oscillations which last about 1 ns. During this time, there will be a net gain change $\Delta g(t) = (-2\omega/c)\Delta n''(t)$, where $\Delta n''(t)$ is the deviation of the imaginary part of the refractive index from its steady-state value. The change in n" is caused by a change in carrier density, which will also alter the real part of the refractive index n'. The ratio of these changes is:

$$\alpha = \frac{\Delta n'}{\Delta n''} \qquad 7.1.1$$

A change in n' during a limited period of time results in an additional phase shift of the laser field and in additional line broadening." The term α is referred to as the linewidth enhancement factor. The laser linewidth can now be calculated as follows:

$$\Delta f = (1 + \alpha^2)\Delta f_{ST} \qquad 7.1.2$$

Δf_{ST} is the linewidth calculated using the Schalow-Townes formula. An estimation of the linewidth enhancement factor of AlGaAs lasers can be made by a correlation measurement between

the AM and FM noises.[7.4] The value of CSP lasers has been estimated to be 2.2 - 2.8 using this method. The factor is dependant on the particular laser structure.[7.5] Natural linewidths in the range of 10 MHz - 30 MHz have been reported for GaAlAs CSP lasers.

As discussed in Chapter 2, techniques currently used for reducing laser linewidths concentrate on the uses of special laser structures that predominantly increase laser cavity Q and/or provide selective optical feedback. The structures receiving greatest attention are the distributed feedback laser, the cleaved coupled cavity laser, and the external cavity laser. The use of these structures has reduced laser linewidths to well below 1 MHz. The effect on system performance of laser linewidth with respect to modulation format was discussed in Chapter 4.

Finally an important source of laser noise external to the laser itself should be mentioned. Both the laser power and frequency stability have been found to be very sensitive to optical feedback by reflections from optical components in the system.[7.6] These reflections can occur in the near field or in the far field. Near field reflections occur when the point of reflection is only on the order of a few centimeters from the laser facet. This case can be treated theoretically as a double cavity.[7.7] Fluctuations in laser output intensity and wavelength occur when there is a change in the interference condition of the double cavity which can be caused by a change in the external cavity length due to physical movement of the optics or by a change in the effective laser cavity length due to temperature and gain effects caused by changing laser carrier

concentration. Generally farfield reflections are considered to be reflections from points from the laser where the round trip time of the external cavity is well beyond the coherence time of the laser. When these conditions prevail it has been found that the light output intensity varies in proportion to the feedback amplitude.(7.8) There is a possibility for these reflections to mode lock the laser causing it to exhibit noisy self pulsations. In addition line broadening and degradations in spectral purity and stability can be induced by reflections. The reflection effects become pronounced when "the reflection coefficient with respect to amplitude is larger than the relative amount of spontaneous emission in the lasing modes".(7.9) This implies that single mode lasers may be particularly sensitive to these reflections. For this reason it is important that certain critical optical components be antireflection coated to reduce the reflection coefficient. Many experimental heterodyne systems have further included optical isolators for the specific purpose of reducing laser noise induced by optical reflection.

7.2 Electronics

7.2.1 Laser Bias Board

The dc bias signal and the modulation signal are combined on the laser board before application to the laser. The board is mounted directly on the laser mount behind the laser package. A circuit schematic is shown in Figure 7.2.1. In order to accommodate the high speed operation required by the experiment, microstripline was used for the laser board. The stripline was designed for 50 Ω impedance so that impedance matching was maintained. The stripline was terminated at the laser by the

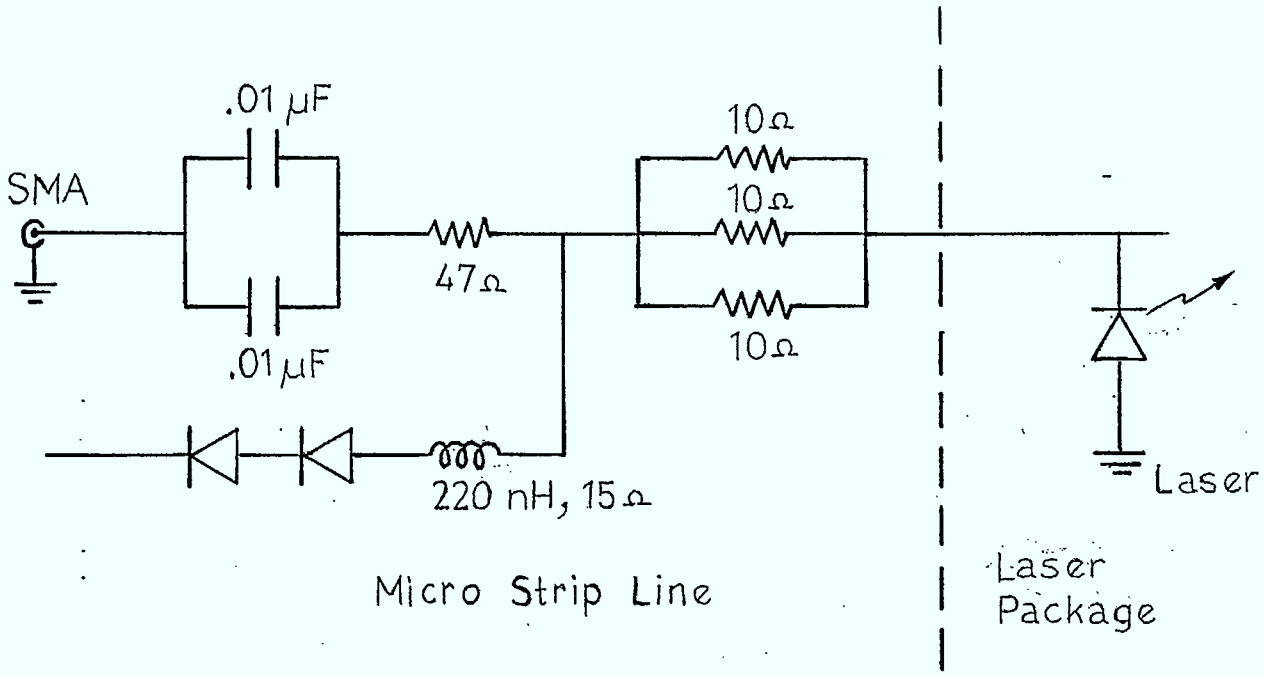
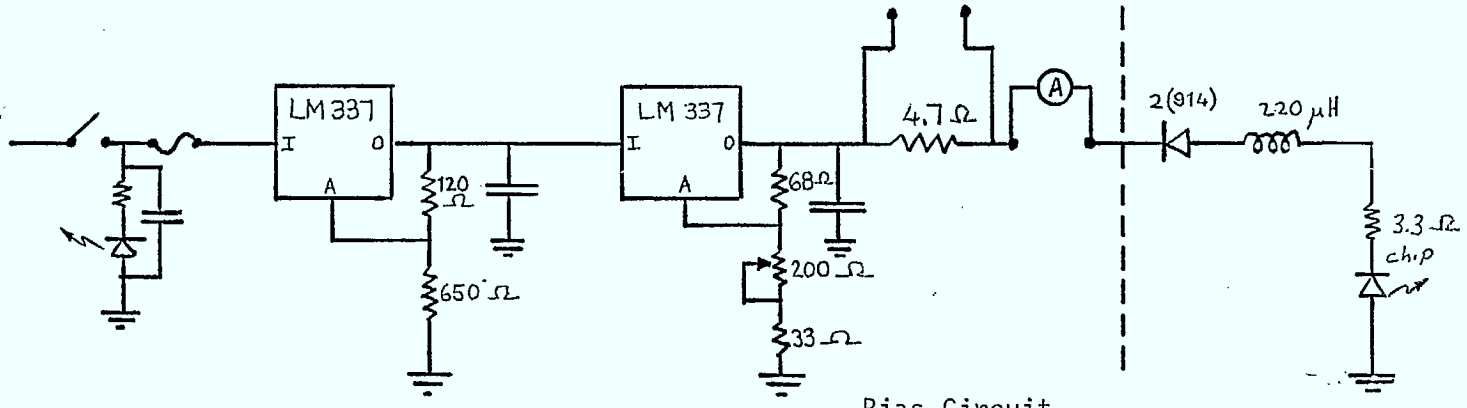


FIGURE 7.2.1: Laser Board Schematic

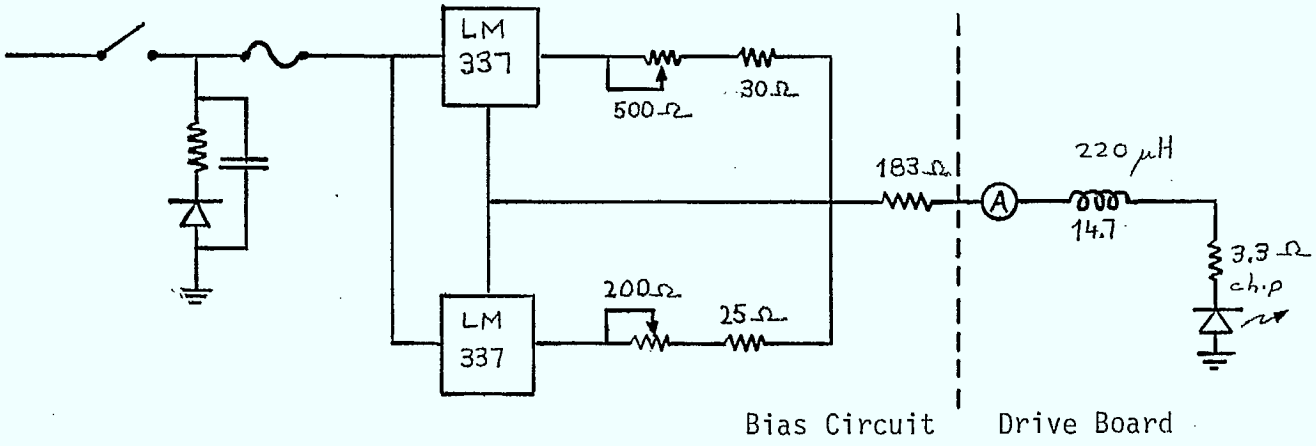
chip resistors and chip capacitors. An SMA connector was used to bring the signal onto the board. Chip capacitors were used to prevent the flow of dc current into the modulator. The dc bias current was supplied through the IN914 diodes and the inductor. The inductor was required to prevent the loss of modulation signal into the bias circuit. The silicon diodes were included to prevent the possibility of back bias on the diode. These diodes were removed for the HL8311E laser. This laser board configuration was such that the modulator drove the laser directly. The critical design consideration for this board is the ensurance of proper impedance matching of the stripline and its termination such that the signal is delivered to the laser with a minimization of reflected power. Distortion of the current pulse through the laser can seriously degrade the performance of the system. In later work, should any transmitter equalization circuitry be required, it can easily be incorporated onto the laser bias board.

7.2.2 DC Bias Circuit

Two dc bias circuits were used during the course of this study. The circuit shown in Figure 7.2.2(a) provides a regulated voltage output from the LM 337. This circuit was used in conjunction with the HLP 1400 laser. The circuit shown in Figure 7.2.2(b) provides a regulated output current. This circuit was used to drive the HL8311E laser. Maximum output currents of -90 mA and -92 mA were obtained with these drivers. Drive currents were normally monitored with in line digital multimeters, however a voltage signal was obtained across the 4.7 and 183 resistors in order to drive an X-Y plotter when light



(a) HLP 1400 Driver



(b) HLP 8311E Driver

FIGURE 7.2.2: Laser Bias Circuits

current curves were measured. The drive circuits must be capable of supplying the required current range, have sufficient protection built in to guard against laser damage (e.g. turn on transients or component failure), and must be sufficiently stable in order not to induce current noise into the laser which would resultin line broadening. Short term stabilities of these drive circuits were generally within 0.1 mA as monitored by the DMMs. For prototype systems it may be worthwhile to re-evaluate the current stability from these circuits, especially long term stability which are affected by factors such as temperature. In addition, it may be valuable to incorporate a power feedback control into the bias circuit.

7.2.3 Modulator

No effort was put into the design of modulation circuitry; a laboratory signal generator was used instead. As already mentioned, the generator could supply a 1 Vpp signal into 50 at repetition rates ranging from 1 MHz to 1 GHz. Duty cycles were adjustable within this range. In order to preserve pulse shape when frequency modulating the laser, the output was set at maximum and attenuators were inserted between the pulser and the laser bias board. This pulser adequately performed as the modulator for purposes of this study.

7.2.4 Cooler

A Peltier effect thermoelectric cooler was used to provide cooling for the laser heat sink. This cooler was driven by the circuit shown in Figure 7.2.3. Cooler drive current was adjustable between 0 and $\sim$ 500 mA. Temperature feedback was provided by an Omega TFD thin film platinum thermistor. The

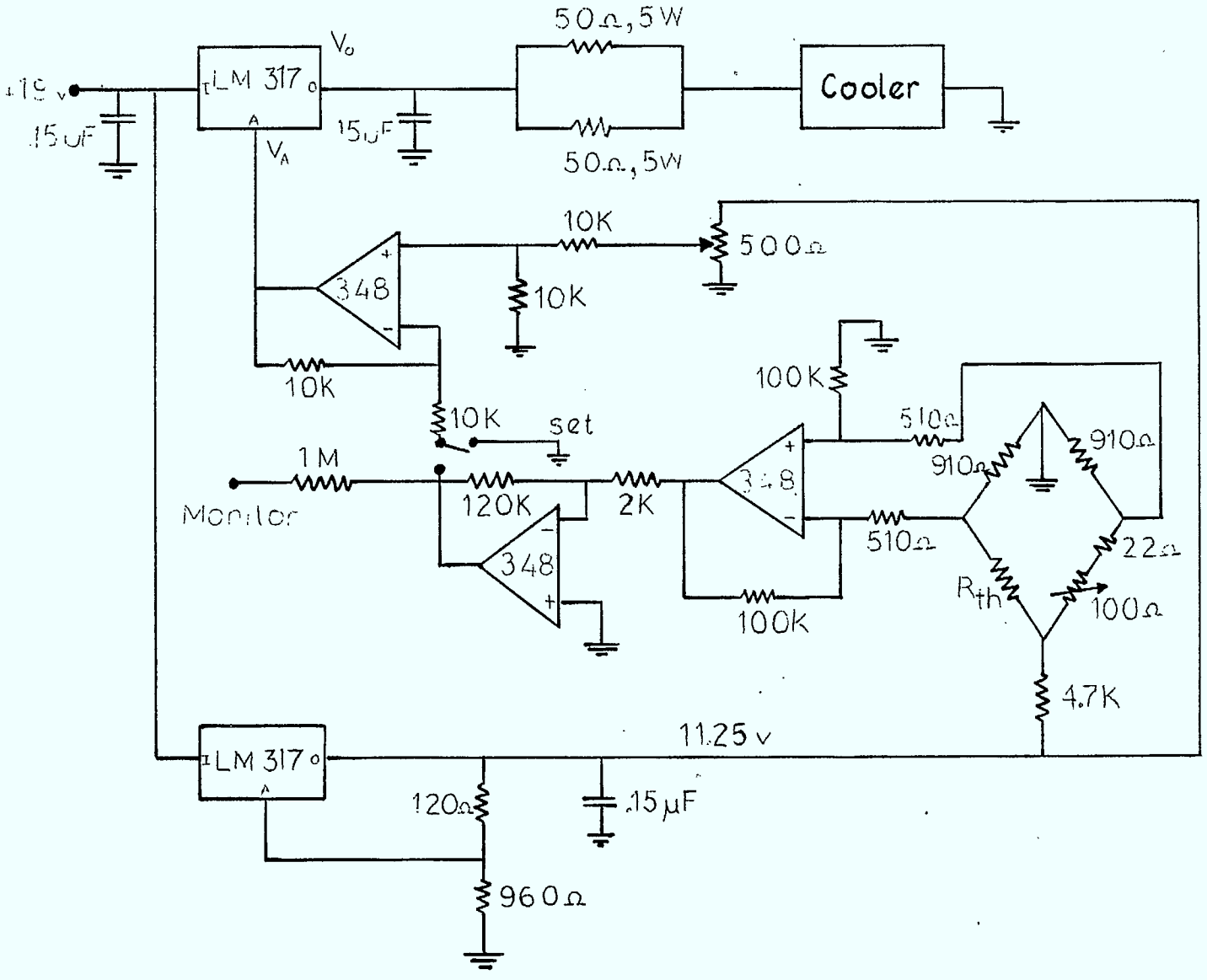


FIGURE 7.2.3: Temperature Control Circuit

thermistor was placed in a bridge, where the temperature set was determined by the 100 Ω pot. The output signal from the differential amplifier was then amplified and sent to another differential amplifier where it was compared with a reference voltage. The output of this amplifier drives a voltage regulator which feeds the thermoelectric cooler. The cooler current needed to maintain a desired heat sink temperature is set by the 500 pot after which, a temperature lock is provided by the feedback circuit. The feedback circuit could be disabled by a switch. The laser was normally operated at a temperature near 20°C, or a few degrees below normal lab temperature.

7.3 Optics

A discussion of transmitter optics requirements has already been made in Chapter 2, and 6. The transmitter optics employed in this study were limited to collimating optics for the transmitter laser. A grin-rod lens of .23 pitch, AR coated, was used to collimate the HLP 1400 laser. The output was approximately circular with a beam diameter of about 1.8 mm. A Melles Griot GLC 002 laser diode collimating lens was used with the HL8311E laser. This lens formed a beam strip of abut 8mm x 3 mm oriented perpendicular to the junction plane. This lens mated with a Melles Griot GPA 001 anamorphic prism pair which expanded the beam by a factor of two in the direction parallel to the junction. The elements were AR coated, and designed for diffraction limited operation at .8 - .85 μm. The lens/laser alignment was performed using a Linetool precision X-Y-Z positioner.

8.0 RECEIVER

8.1 System Description

The functional elements of the testbed receiver were given in Figure 2.1.1 and briefly discussed in Chapter 6. In summary, these elements include the detector, the local oscillator, bias and cooler electronics, a beam combiner, and local oscillator collection optics. Also included as part of the receiver were a custom laser mount and an environmental protection chamber for the local oscillator. An AFC loop was later developed to improve the frequency stability of the IF. This loop consisted of a low pass filter, a power splitter, dual balanced mixer and a delay line. Components were also procurred for the construction of laser external cavities. The purpose of the external cavity is to narrow the laser linewidths. Such a configuration would have added further refinement to the system.

8.2 Detector Devices

In order to maximize the possible performances of the heterodyne receiver, three detectors have been purchased and fully tested. These devices are silicon photodiodes which have high efficiency in detecting near-infrared radiation (peak efficiency $\sim$ 900 nm). Each detector has an optimum receiver power and frequency range of operation. The detectors corresponded either to a high sensitivity device having a sensitive area of about 1 mm^2 and a time response in the nanosecond range (RCA) or to a device having a faster time response ($\sim$ 50 psec) on a reduced area size and sensitivity (Antel). Two types of photodiodes have been used in this work: a pin diode and two avalanche photodetectors (APD). The avalanche photodiode differs

from the pin by the multiplication process taking place at the junction which results in more generated electrons for a given number of incident photons. This multiplication factor, called the photodiode gain, is dependent on the bias voltage applied to the detector. The gain can be varied from 1 (APD equivalent to a pin diode) up to 150. High gain values usually require voltage bias too close to the breakdown value to safely operate the detector. The pin and the APD made by Antel Optronics Inc. are very fast devices, having bandwidths of several gigahertz. However, this wide frequency response reduces both the detector sensitivity and its area. The APD built by RCA has a preamplifier module integrated into the detector. This increases the detector sensitivity by allowing an effective diode load impedance higher than 50 Ω with a bandwidth of 200 MHz. The two following subsections give the theoretical definition and the experimental measurements of the main detector parameters.

8.2.1 Theoretical Background

This part gives the relationship between the incident power on the detector and the signal and noise output measurements. The gain appears explicitly in the equations which are thus valid for both a pin and an APD detection device.

RESPONSIVITY

The detector responsivity which is wavelength dependant is related to the photodiode efficiency by the equation:

$$R_o = \frac{q\, n(\lambda)}{h\nu} \qquad 8.2.1$$

where $n(\lambda)$ is the photon-electron conversion efficiency (<1)

q and h and ν have their usual significance.

The R_o value quoted by the manufacturer usually includes radiation loss through the protective window or the focussing lens eventually mounted in front of the diode.

SIGNAL

The signal current (I_s) flowing through the diode is related to the detector received power (P_R) by the equation:

$$I_s = MR_o P_R \qquad 8.2.2$$

where M is the APD gain (M = 1 for a pin diode).

This linear relationship is observed up to a certain power level at which the detector starts to saturate. This saturation level is given as a current value which remains the same for any detector gain values. The corresponding saturation power is obtained from the Equation 8.2.2. The output voltage is given as:

$$V_{OUT} = I_s R_d = M R_o P_R R_d \qquad 8.2.3$$

where R_d is the diode load impedance.

Any additional load connected at the detector output has to be taken into account since the photodetector acts as a current source. However, the RCA photodiode has an integrated preamplifier module and so acts as a voltage source having its responsivity quoted in Volts/Watts according to the equation:

$$V_{OUT} = R_o' P_R \qquad 8.2.4$$

with $R_o' = M R_o R_d$.

NOISE POWER DENSITY SPECTRUM

The noise power density variation versus frequency is given by the following equation for the RCA photodiode coupled to its preamplifier:

$$P_n = R_d (S_o + S_2^{\ 2}) f^2(\omega)\, d\omega \qquad 8.2.5$$

The different terms in this equation have been detailed in Sections 3.1.1 (Equation 3.1.1) and 3.2.2 (Equations 3.2.5 and 3.2.6). The noise power density appears as a complex sum of noise sources: the dark current which is always present even with no light on the detector, the thermal noise, and the shot noise which is induced by the incident radiation (background + signal). Due to the preamplifier module integrated to the photodiode, the frequency response differs from a RC one. The bandwidth of the detector which is defined as the frequency value at which the signal drops by a 3 dB factor is related to the detector time response by:

$$\nu_{3dB} = 0.1024/\tau_d \qquad 8.2.6$$

The noise power density equation is much simpler in the case of a photodiode detector directly connected to its load resistance (Antel detector):

$$P_n = [R_d\ \ 2q\{i_{ds} + (i_{db} + P_R R_o) M^2 E\} + 4kT] x \frac{1}{1+\omega^2\tau_d^2}\, d\omega \qquad 8.2.7$$

The relation between the bandwidth and the detector time response is:

$$\nu_{3dB} = 0.1592/\tau_d \qquad 8.2.8$$

This case corresponds to a RC circuit where the cut-off frequency is set by the diode capacitance and load impedance values. The rise time can be related to the time response τ_d by:

$$\tau_R(10\%-90\%) = 2.197\ \tau_d \qquad 8.2.9$$

8.2.2 Experimental Measurements of Some Detector Parameters

Detector parameter values such as the responsivity, time response, etc., quoted by the manufacturer are usually typical values. Therefore more precise values have to be obtained by experimentally measuring these characteristics. In this section, it is assumed that the gain value of the APD is known. The gain experimental determination is described in detail in Section 8.2.3. Each type of measurement is performed repeatedly in order to decrease the percentage of error. The results are shown in Table 8.2.1 with other non-experimentally measured parameters.

PHOTODIODE LOAD RESISTANCE

The value of the load resistance connected in series with the photodiode has been obtained from simultaneous measurements

TABLE 8.2.1: Detector Characteristics

	RCA APD C309506	ANTEL APD AR-S5	ANTEL PIN AR-S2
Responsivity (A/W) at 830 nm	0.72 A/W	0.25 A/W	0.30 A/W
Diode Load Resistance (Ω)	1000	50	50
Time Response (ns)	0.7	0.045	0.02 *
Rise Time (10→90%) (ns)	.2	0.100	0.045 *
Bandwidth (GHz)	0.146	3.5	>5 *
Gain	1→280	1→100	1
Saturation Current (mA)	0.100	1.0	1.0
Bias Voltage (V)	234 (for gain = 60) +/-12 preamplifier	-134 (for gain = 60)	-15
Sensitive Area (mm^2)	0.25*	0.03 *	0.01 *

* As Given by Manufacturer

of the current flowing through the detector and of the output voltage (Equation 8.2.3) while shining the detector with a HeNe laser. Both the current and the voltage have been measured on a digital multimeter (Fluke 75). The measurements have been repeated for various incident power levels.

RESPONSIVITY AND SATURATION LEVEL

To do a responsivity measurement, the absolute amount of power incident on the photodetector has to be known. We have used either a HeNe laser or a semiconductor laser. Both have been calibrated in power with a Scientech power meter (Model 360203) after passing through a lens which focussed the radiation on the detector sensitive area. By recording the output voltage for different amounts of input power (using calibrated neutral density filters) the responsivity and the saturation level has been obtained for a cw radiation. Responsivity measurements have also been performed with the semiconductor laser pulsed in the 10 MHz - 1 GHz range. As expected the responsivity stays constant up to a frequency close to the detector bandwidth (see next subsection). In the case of the two Antel photodiodes no drop is observed since their bandwidth is wider than 1 GHz.

BANDWIDTH AND TIME RESPONSE

As it has been discussed in subsection 8.2.1, the detector 3 dB bandwidth frequency and the time response can be related to each other very easily. It is thus sufficient to measure only one of these two characteristics. The time response can be obtained from the measurement of the time evolution of an incident pulsed signal but this is only possible if the input pulse rise time is faster than that of the detector. The

experiment has been performed on the RCA APD. The radiation source was a light emitting diode (LED:C86011E from RCA). Its pulse shape is presented in Figure 8.2.1(a). It has been measured with a fast detector coupled to an oscilloscope (Techtronix 7834 Plug in 7A29 and 7892A). The output signal of the RCA avalanche photodiode appears in part b of the same figure. The experiment has been done for various bias voltages applied on the preamplifier module. This bias is expected to change the detector time response which is essentially determined by the preamplifier module. A value of τ_d is obtained by fitting the experimental results to the theoretical time evolution equation of the signal pulse given in Section 3.1.1. To do so an analytical function describing the LED incident pulse is used. The fit is shown in Figure 8.2.2 for two different preamplifier bias voltages. The limited accuracy given by the oscilloscope is responsible for the discrepancies observed on this graph. However, the curve slope is sensitive enough to a τ_d change to give us a reasonable time response measurement.

The detector bandwidth can be measured in the frequency domain by recording the shot noise power spectrum. The noise has been measured on a spectrum analyser after amplification. The radiation source was a HeNe laser. Depending on the frequency range, different spectrum analysers and amplifiers have been used: 0 - 500 MHz 40 dB Amplica amplifier 401 USL coupled with a spectrum analyser HP 8553B and 2 GHZ - 4 GHz 33 dB Miteq amplifier AF03 coupled with a spectrum analyser HP 8569A. A portion of the noise spectrum of the Antel avalanche photodiode is presented in Figure 8.2.3. The 3 dB frequency which defines

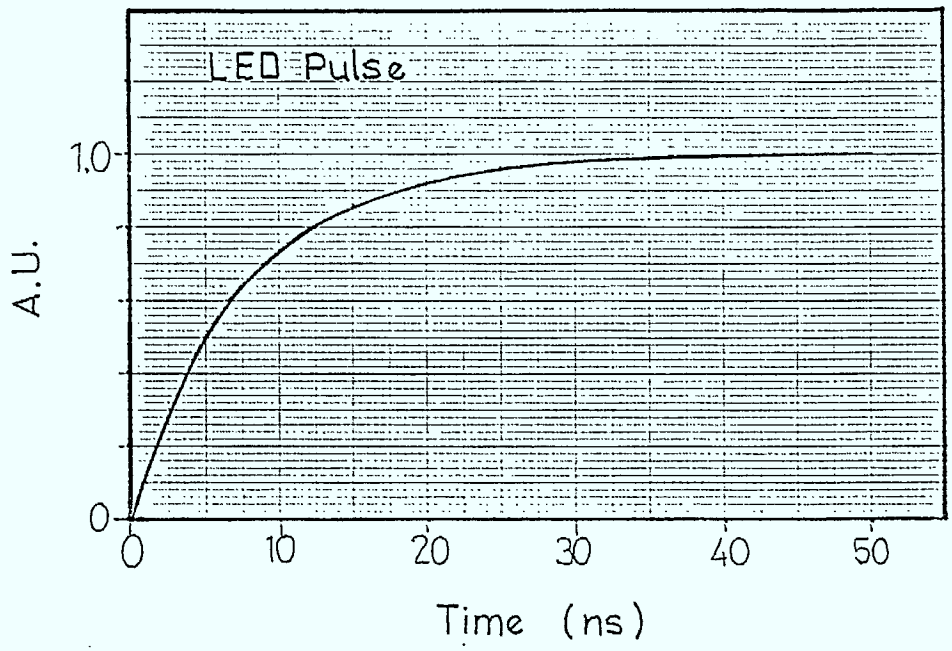


FIGURE 8.2.1a: LED Pulse

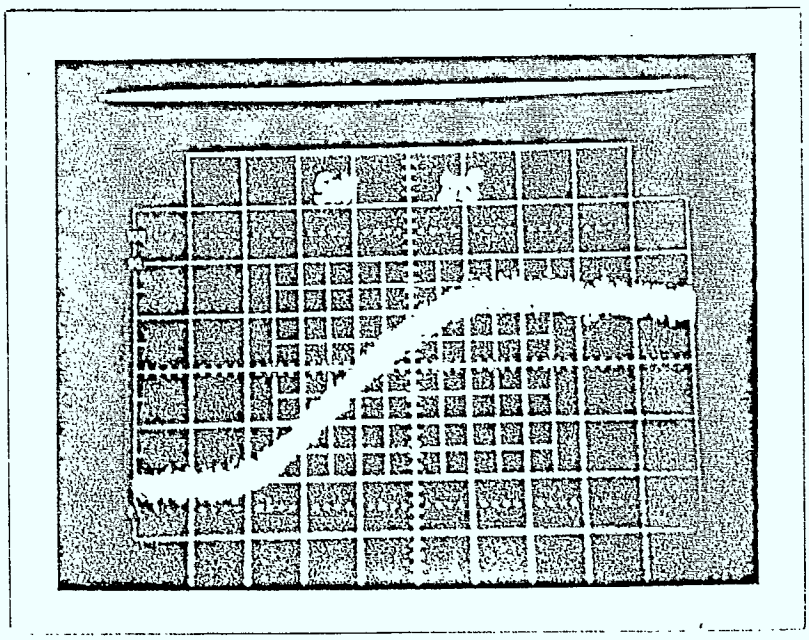


FIGURE 8.2.1b: Detector Output Pulse Time Evolution

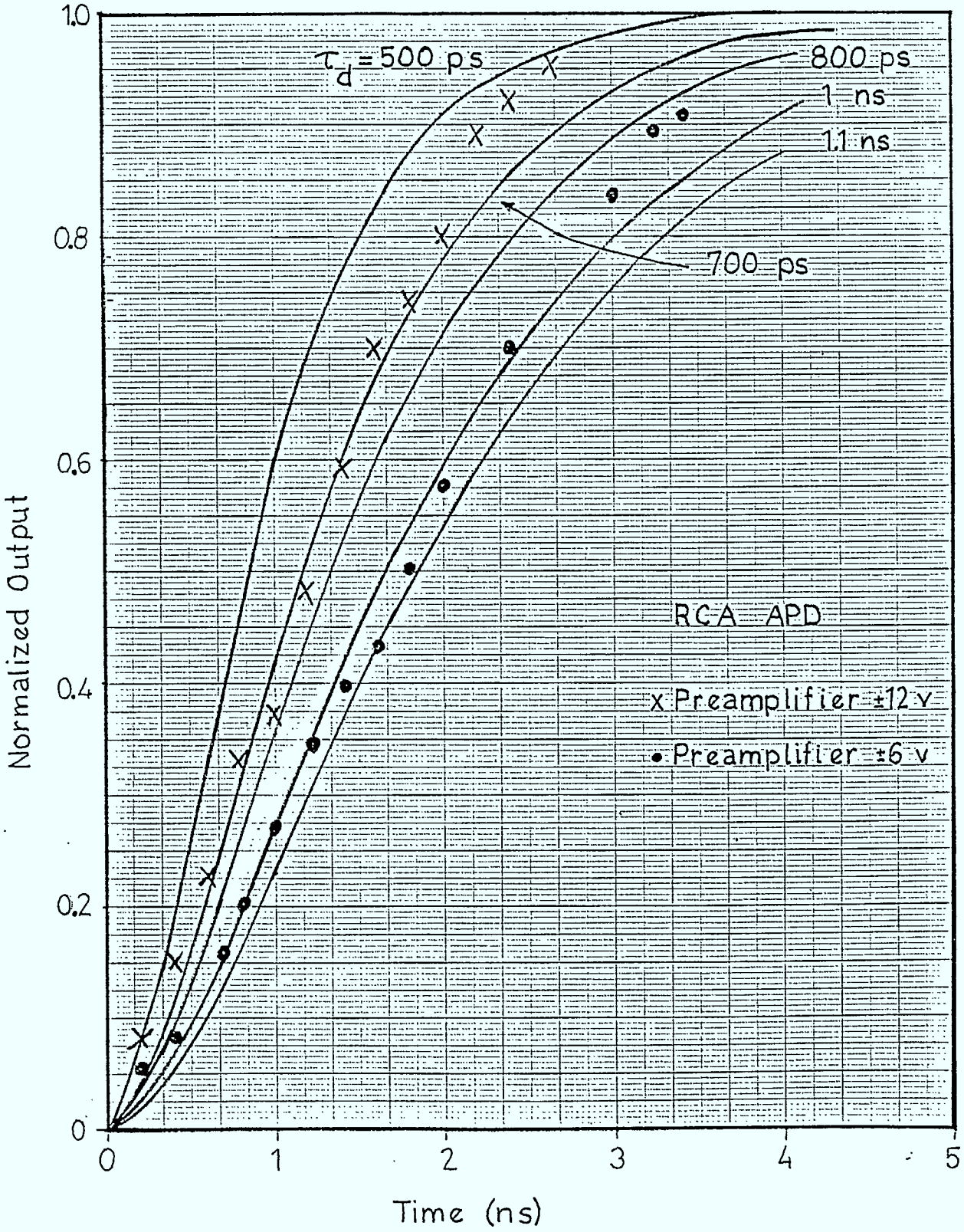
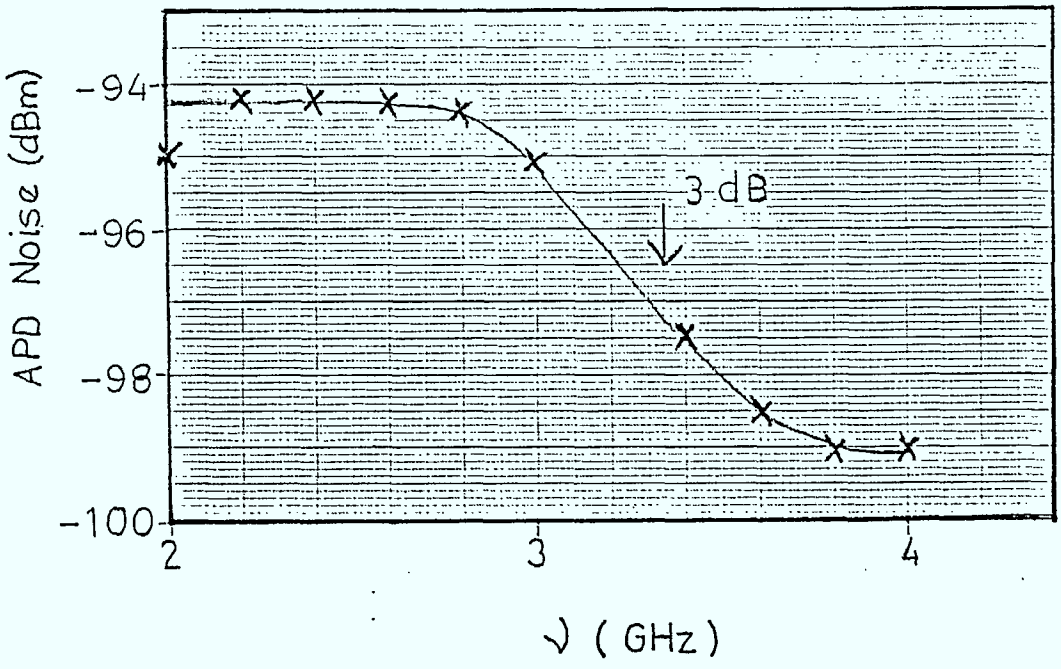


FIGURE 8.2.2: RCA APD Time Response Measurement



BW (3 dB) $\approx$ 3.5 GHz

$\tau_d \approx$ 45 psec

$\tau_r \approx$ 100 psec

FIGURE 8.2.3: Antel APD Bandwidth Measurement

the detector bandwidth is indicated on this graph with its τ_d and τ_R corresponding values.

8.2.3 Gain Measurement Technique

The gain value of an avalanche photodiode is strongly related to the signal to noise ratio (S/N) optimization for the receiver system. Its absolute value and its variation as a function of the detector bias voltage have to be accurately known but are usually not provided by the manufacturer and thus have to be experimentally measured. As discussed in Section 8.2.1, the signal current is proportional to the product of the detector responsivity and the gain. By measuring the signal variation as a function of the bias voltage for a fixed amount of received power, the relative variation of the gain is obtained since the responsivity before gain is not a function of the bias applied on the APD. However, its absolute value cannot be obtained with this type of experiment unless the responsivity is known. We thus searched for another way of measuring the avalanche gain. Looking back to the theoretical Section 3.2.2, it is found that the noise power density is a function of the gain square through the shot noise term (assuming 1 Hz BW measurement):

$$P_{\text{shot noise}} = 2\, R_d\, R_o\, M^2 E P_R\, f^2(\omega) \qquad 8.2.10$$

where $f(\omega)$ is either the RCA frequency response (Equation 3.1.1)

or the usual RC frequency response.

If the experimental conditions are such that the dominant noise source is the shot noise (dark current and thermal noise terms are negligible) and if the measurements are performed at a much

lower frequency than the 3 dB bandwidth value ($f(\omega) = 1$), the signal to noise ratio becomes a simple function of the detector gain only, given by:

$$S/N = \frac{M\, R_o\, P_R}{2q\, R_o\, P_R\, M^2 E} = \frac{1}{2qME} \qquad 8.2.11$$

where E is, as previously described, a function of the gain:

$$E = M\, k_e + (1 - k_e)(2 - \frac{1}{M}) \qquad 8.2.12$$

with $k_e = 0.02$ for a silicon photodiode.

One of the most interesting features of this technique is that the measurement does not require the knowledge of the received power. The gain value accuracy depends only cn the instrumentation used. This attractive way of making gain measurement requires one condition; that the received power level is high enough to have a shot noise limited receiver but lower than the detector saturation level. The signal and the noise have been simultaneously measured as a function of the bias voltage applied to the detector. The experimental set up has already been described (Section 8.2.2). The shot noise has been induced by either HeNe radiation or the LED operating in a pulse mode. The results are presented for the two avalanche photodiodes in Figures 8.2.4 and 8.2.5. The discrepancy seen in Figure 8.2.4 between the values obtained with the cw and the pulsed measurements is a result of the spectrum analyser scale accuracy which is of the order of +/- 1 dB.

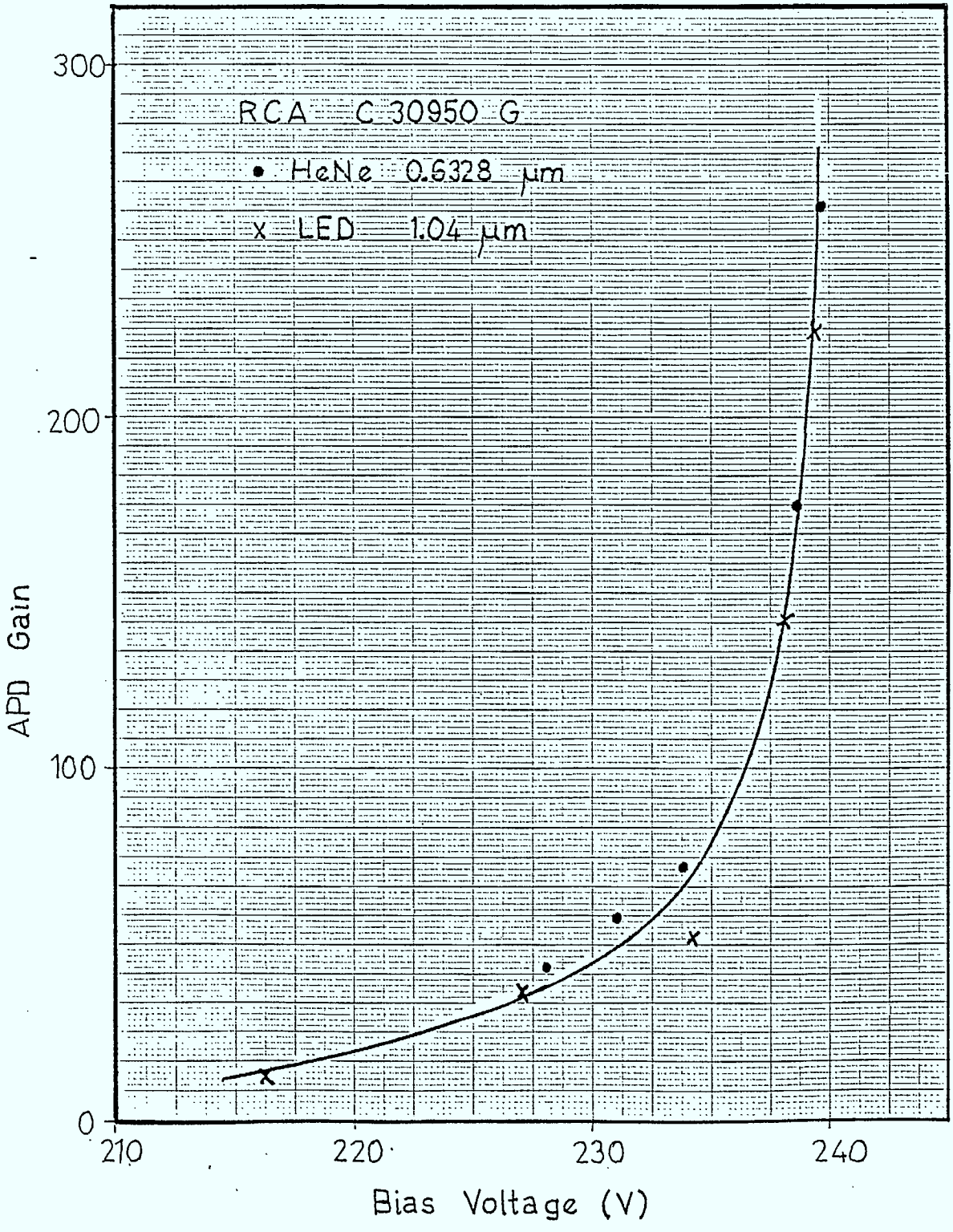


FIGURE 8.2.4: RCA APD Gain as a Function of the Bias Voltage

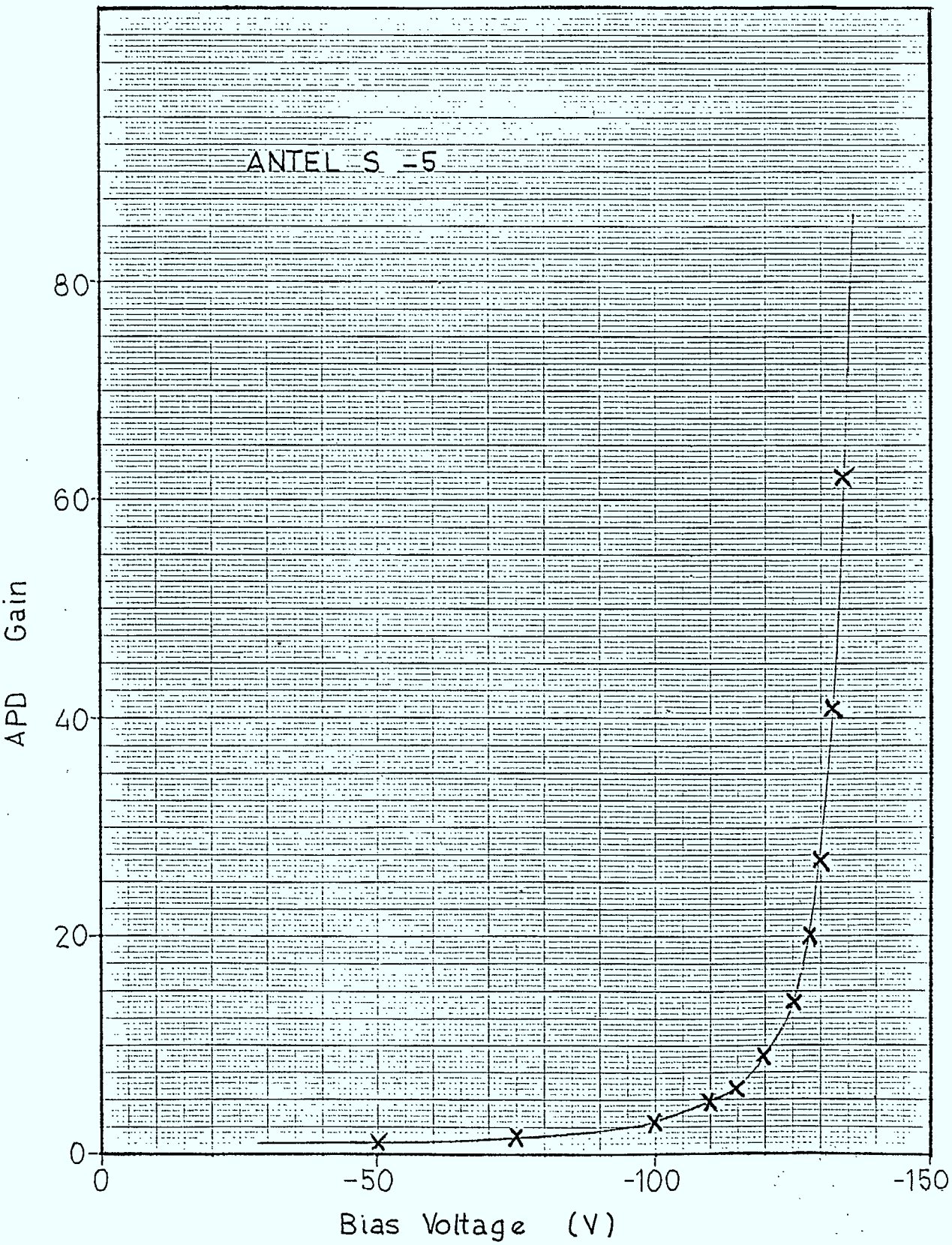


FIGURE 8.2.5: Antel APD Gain as a Function of the Bias Voltage

8.3 Local Oscillator

The laser that was used as the local oscillator was a Hitachi HLP 1400 similar to that used for the transmitter. Therefore, much of the description in Section 7.1 also applies to the local oscillator. Typical device specifications are also identical.

The most significant difference between the local oscillator and the transmitter laser is in the manner of operation. Whereas the transmitter laser is modulated at high speed, the local oscillator is operated strictly in a cw fashion with fine adjustments to its drive current made only for the purpose of frequency tracking the transmitter signal. As such, no laser drive board comparable to that used for the transmitter was required. The laser was however, connected to a microstripline board which was then connected to the dc drive circuit through SMA cable, but this was done simply for ease of connection.

The local oscillator was placed in an environmental protection chamber. This chamber was double walled with insulation between the plexiglass and aluminum walls. Such a structure was used to isolate the LO from ambient temperature variations. The housing was sealed from the top thus protecting the laser from airborne contaminants. In addition desecant was placed inside the housing in an effort to reduce the relative humidity experienced by the laser.

All issues regarding spectral purity, linewith, freqency stability and noise apply to the LO in the same manner as they do for the transmitter laser.

Coarse wavelength matching of the two lasers was accomplished by temperature tuning.

8.4 Electronics

8.4.1 Local Oscillator Bias Circuit

The local oscillator bias circuit used in this receiver is shown in Figure 8.4.1. This circuit is based on the one used to bias the HLP 1400 transmitter laser. However, this circuit is modified in a manner such that it will automatically adjust the output voltage in accordance with the error signal from the AFC loop. The error signal is compared with the current set at the differential amplifier, the output of which is used to drive the voltage adjust pin on the LM337 negative voltage regulator. A 5V zener diode has been added for diode protection. A signal is taken across the 12Ω resistor to drive an X-Y recorder when measuring light-current curves. All components with the exception of the laser diode and the ampmeter are contained on the bias board. The current range for this drive circuit is 1 - 90 mA.

8.4.2 Detector Bias Circuit

The three detectors used in this study were all biased by laboratory high voltage low current supplies. A Power Designs Pacific Inc. 310 was used to bias the Antel AR-S5 APD. A Kepco APH 1000M was used to bias the RCA C30950G APD. An HP 6234A general dual supply was used to provide the low voltage bias to the Antel AR-S2 PIN detector. In addition, an HP 6234A was also used to bias the Miteq amplifier. In all cases precautions were taken to limit the bias current below the damage threshold levels quoted by the manufacturers.

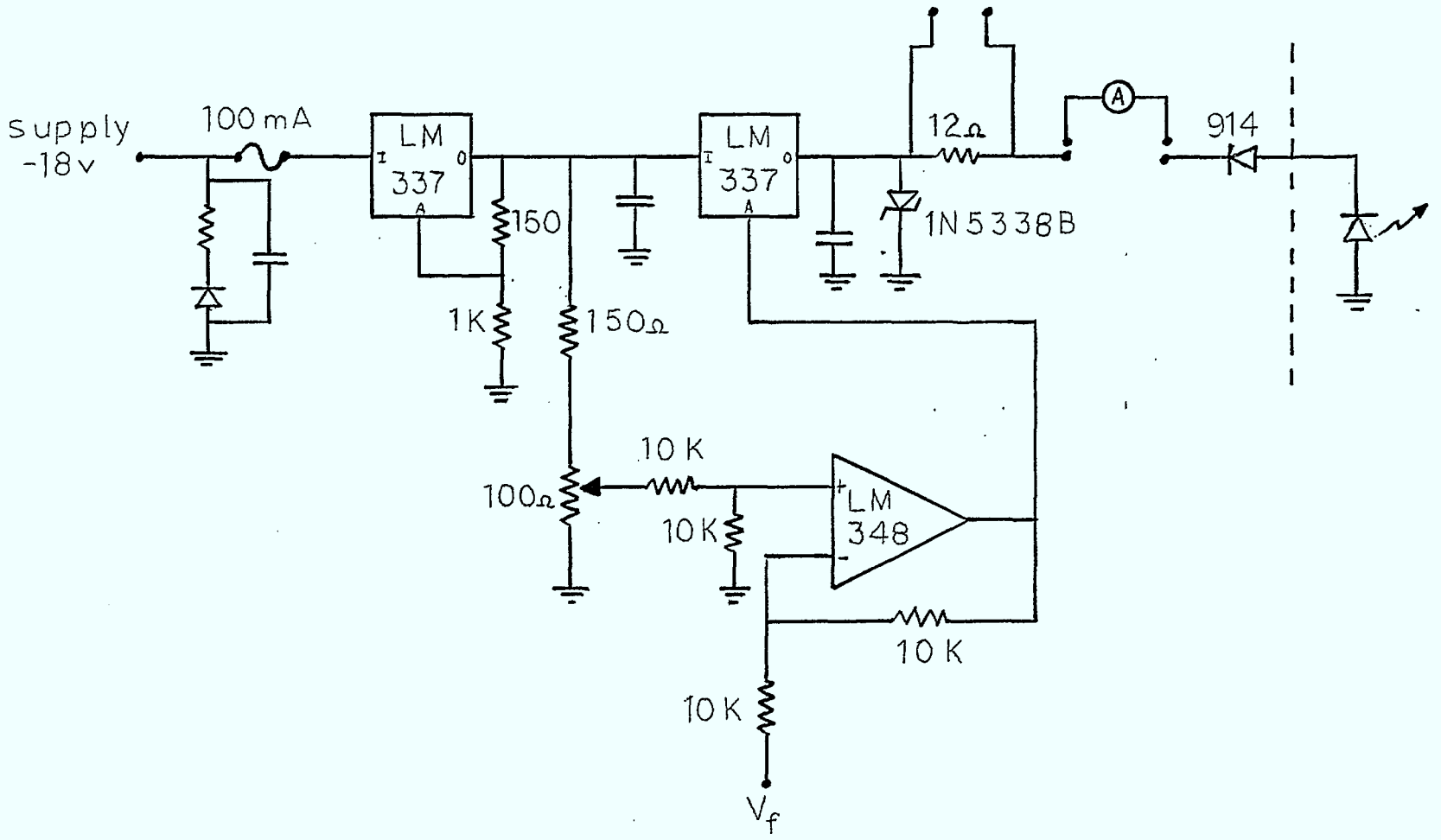


FIGURE 8.4.1: Local Oscillator Bias Circuit

8.4.3 Cooler

The cooler drive circuit used for the local oscillator is identical to that used in the transmitter with the following minor exceptions. The 1 M resistor at the monitor output was replaced by a voltage follower using one of the op amps on the LM 348 chip. A Melecor CP1.4-31-10L Thermoelectric cooler was used instead of the CP1.4-71-10L used in the transmitter. Aside from these differences, the discussion in Section 7.2.4 is also applicable here.

8.5 Optics

8.5.1 Local Oscillator Collection Optics

A grin-rod lens identical to that used for the HLP 1400 transmitter laser was used for collection and collimation of the local oscillator light. The part number is given in Table 6.1. The lens was mounted on a Linetool three-axis precision positioner. The beam spot was approximately circular with a diameter of 1.8 mm.

8.5.2 Beam Combiner

The beam combiner used in this testbed is a conventional beamsplitter cube. This beamsplitter cube was made of BK 7 glass with a broadband coating. A typical reflection curve shows a reflection coefficient of approximately 58% at 830 nm. Absorption is specified at about 5%. A disadvantage of a beamsplitter cube is the occurance of reflections off of two faces instead of one, causing a possible increase in reflections seen by the laser, and a slight increase in optical loss.

The beamsplitter cube was fixed on a prism table which was

fastened to a rotating stage. This stage was then mounted on a precision Z-axis translator. Such an assembly facilitated the optical alignment of the system.

Using the relation $\theta \lesssim \lambda/4d$ where d is the diameter of the detector, λ is the lasing wavelength, and θ is the misalignment between the two beams, the alignment of the local oscillator and the transmitter beam should be better than or equal ot 0.06° or 3.6 min. of arc.

8.6 Improvements

8.6.1 AFC Loop

As a next step in upgrading the heterodyne testbed, an AFC loop was designed. The components in this loop are shown in Figure 6.1.1. The loop is made up of a bandpass filter (in our applicaton a low pass 0 - 4 GHz filter sufficed), a power splitter, a dual balanced mixer, a delay line and a low pass output filter. Part numbers are given in Table 6.1. This signal would then normally pass through an adjustable gain amplifier in order to match the laser current - Δf slope with the feedback signal.

The mixer/delay line combination composed a frequency discriminator. The output of the discriminator is determined by the following relation:

$$v_e = \frac{A}{2} \cos\left(\frac{2\pi\partial}{v}\right) \cdot f \qquad 8.6.1$$

where A is the signal input amplitude, ∂ is the delay in the two paths, v is the speed of the wave in the guide (2×10^8 m/s in this case) and f is the signal frequency. Coaxial cable was used

as the delay line. It is important that amplitude variations be eliminated so that changes in the error signal are only due to changes in the IF frequency. Figure 8.6.1 gives the measured normalized discriminator response. The discriminator has a useable capture bandwidth of about 600 MHz centered at 3 GHz. The vertical scale, or slope in mV/MHz is determined by the strength of the input signal. AM noise in the input to the discriminator will result in a mismatching of the discriminator $\Delta V/\Delta f$ response and the local oscillator $\Delta f/\Delta I$ response. Frequency lock is lost, or at least severely degraded in such a case. It is important that discriminator noise not add excessively to drive current noise which may result in laser line broadening. For efficient discriminator operation the IF should be kept as narrow as possible.

The discriminator has been developed to track the frequency of the f_1 component. It is important therefore that this frequency component be sufficiently strong. The power contained in this component is very code dependent. For example, if a long string of zeros were sent, then the time averaged power may drop to level such that frequency lock may be lost. It is also important that the f_1 sidebands be out of the bandpass of the discriminator otherwise they may appear as false signals.

Frequency tracking can only be monitored between $\sim$ 2.7 GHz and 3.3 GHz. This corresponds to wavelength offsets near 0.1 Å. Therefore the discriminator will not be able to correct for any mode hops by the lasers.

An AGC amplifier may be needed to maintain a constant IF signal strength into the discriminator. Amplification and

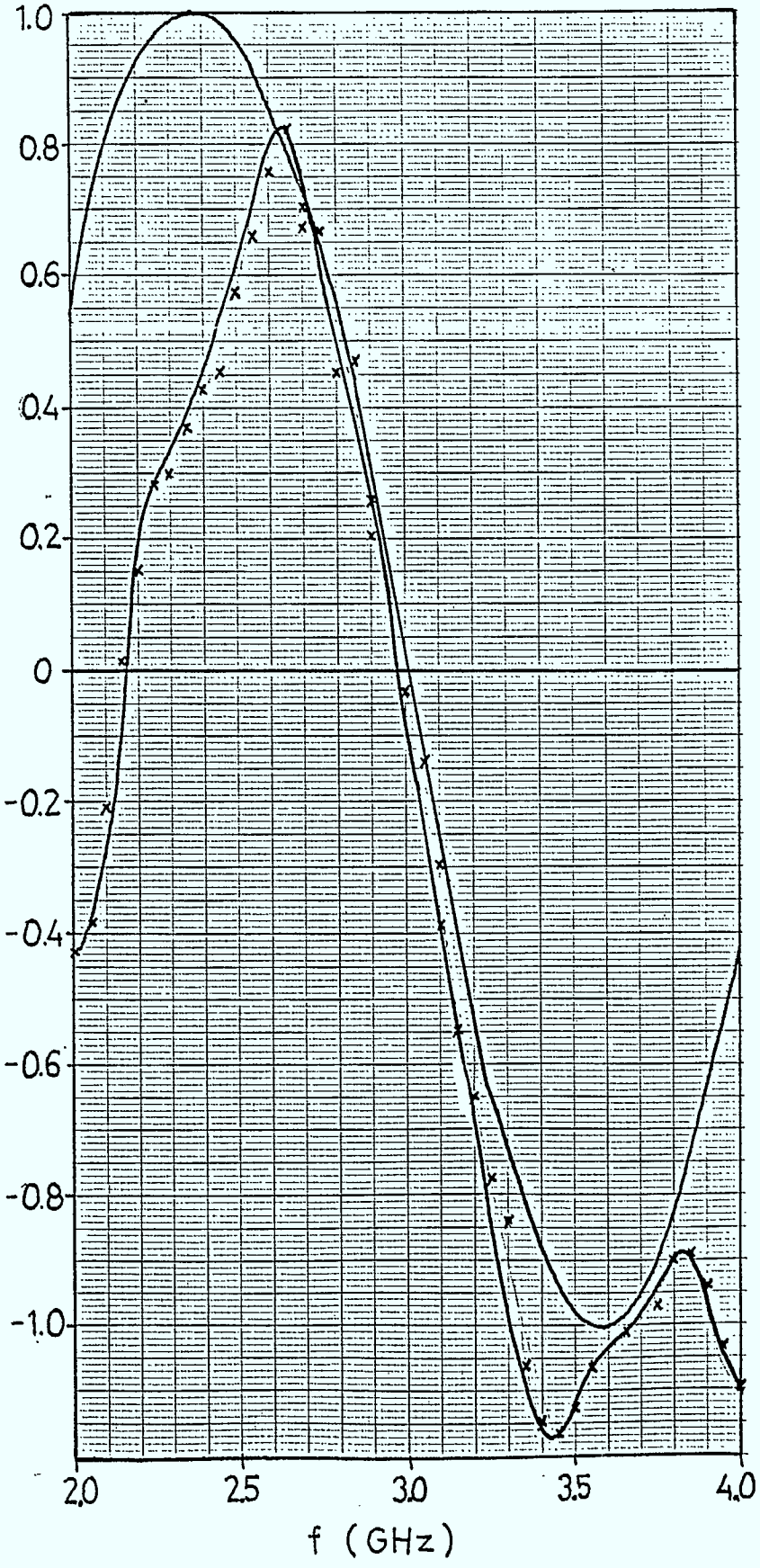


FIGURE 8.6.1: Frequency Discriminator Response

clipping may not be feasible if the clipping procedure induces excessive signal distortion.

As noted in earlier discussions, IF frequency stability is extremely important in coherent detection, and so it is in these systems that AFC operation is most critical. FSK single filter envelope detection is amongst the systems most tolerable of IF drift and linewidth.

8.6.2 External Cavity

Components were purchased to enable experimentation of laser line narrowing in an external cavity configuration. As previously mentioned, linewidths well below 1 MHz have been reported using such techniques. The external cavity was designed for tunability by using selective grating feedback, and by adjustment of the cavity length. The cavity was intended for use with the HLP 1400 laser used in the local oscillator. However, due to time and cost constraints this improvement was not fully implemented.

Pertinent measurements with this configuration would include the extent of linewidth narrowing and its sensitivity to cavity length and reflectivity, tunability, effect on FM modulation response and ease of implementation with regards to stability, complexity and use. Sophisticated cavity configurations have used automatic feedback to piezoelectric elements attached to the grating, or mirror, which can provide a fine adjustment in cavity length.

Operation of the HL8311E was not possible in this configuration because of its single facet access.

9.0 PERFORMANCE MEASUREMENTS

A series of experiments have been performed for both direct and heterodyne detection in order to measure system performance. The lasers were operated in cw and pulsed mode. Extensive studies on the laser excess noise were done as a function of various parameters such as laser current and power, the receiver frequency range and laser modulation. The results were compared to theory which, was then used to extrapolate system performance.

9.1 Direct Detection Sensitivity

The direct detection performance is described throughly the signal to noise ratio (S/N) which can be achieved under various conditions. This ratio depends on the detector characteristics that determine the amount of shot noise, and on the laser excess noise which, for semiconductor lasers, makes a significant contribution in certain frequency ranges. As a result the system operates either as a shot noise limited receiver or as a laser excess noise limited receiver. The transition between these two modes of operation is a complex function of the laser and the detector characteristics. This section describes the theoretical and experimental approaches towards this system behaviour.

9.1.1 Laser Excess Noise Equation

The fluctuations of the carrier and photon populations in the semiconductor laser result in modulation of the laser output power. The mean square noise fluctuation percentage can be written as:

$$\Delta^2_L = \left(\frac{\Delta P_L}{P_L}\right)^2 \qquad 9.1.1$$

where P_L is the laser total output power (usually different from the received power P_R).

The noise power density spectrum $(\Delta P_L)^2$ is given by:[9.1] (for 1 Hz bandwidth measurement)

$$(\Delta P_L)^2 = 2\,h\,\nu_L\,P_L\,\gamma\left[\frac{\left(\frac{1}{\sigma^2\tau^2}+\frac{\omega^2}{\sigma^2}\right)+\frac{\omega_o^2}{\beta\sigma\tau}(1-\beta)}{\left(\omega_o^2+\frac{1}{\sigma_1\tau_1}-\omega^2\right)^2+\omega^2\left(\frac{1}{\sigma_1}+\frac{1}{\tau_1}\right)^2}\right] \qquad 9.1.2$$

where -
- γ is the reciprocal value of the degree of inversion which is dependant on the laser current
- σ is the photon lifetime
- σ_1 is the effective time constant of the cavity in the presence of gain
- τ is the carrier lifetime
- τ_1 is the damping time constant
- ω_o is the resonance frequency at which the maximum noise enhancement is observed
- ω, h, ν_L have their usual significance.

This equation shows that the AM laser excess noise spectrum has a maximum around a particular resonance frequency. This frequency which is dependant on the laser current is in the gigahertz region when the laser is driven above its current threshold. The magnification factor is such that an increase up to 15 dB is obtained on the laser excess noise around ω_0. The laser parameters involved in Equation 9.1.2 have not been measured in this work. We have used typical values given in the literature[9.1].

σ = 3 ps

τ = 2 ns

From these, values for ω_0, σ_1, and τ_1, can be obtained using the following equations:(9.1)

$$\omega_0^2 = \frac{i_L/i_{th} - 1}{\beta\sigma\tau} \qquad 9.1.3$$

where i_L/i_{th} is the ratio between the laser current and its threshold current

β is typically equal to 0.2

$$\frac{1}{\sigma_1} = \frac{mc}{\omega_0^2\ \sigma^2\tau\beta} \qquad 9.1.4$$

with m = 1 for a single laser operation

c is typically of the order of 10^{-4}.

$$\frac{1}{\tau_1} = \frac{1}{\tau} + \omega_0^2\ \sigma \qquad 9.1.5$$

As it has been mentioned above, ω_0 and γ are a function of the laser current. This is also true for σ_1 and τ_1. The only parameter for which no exact value has been found is the factor γ. However this factor appears as a multiplication factor and will not affect the noise frequency variation.

9.1.2 Total Noise Equation

The total amount of noise collected at the detector output is a sum of the laser excess AM noise and the detector dark current, thermal and shot noise . The detector noise has been discussed in Sections 3.2 and 8.2.1. The laser excess noise effectively collected at the detector output is given by:

$$P_{nL} = R_d \, (\Delta_L^2 . P_R^2 \, R_0^2 M^2) f^2(\omega) d\omega \qquad 9.1.6$$

where R_d is the load resistance of the photodiode

Δ^2_L is given by the equations 9.1.1 and 9.1.2 and is a function of ω.

P_R is the portion of the total laser power received at the detector

R_0 is the detector responsivity

M is the detector gain

$f(\omega)$ is the detector frequency response function (RCA or ANTEL type).

The total noise collected at the detector output is given by:

$$P_n = R_d \, [2q\{i_{ds} + (i_{db} + R_0 P_R) M^2 E\} + 4 \, kT + \Delta_L^2 \, P_R^2 \, R_0^2 \, M^2] f^2(\omega) d\omega \qquad 9.1.7$$

This equation assumes that no background light is detected. It is interesting to observe that the laser excess noise term varies with the square of the received power while the shot noise is proportional to it. The observation of a shot noise or laser noise limited receiver will thus depend on the amount of attenuation of the laser beam. This competitive effect will also vary with the collected frequency range since the laser excess noise shows a strong frequency resonance while the shot noise is almost constant up to the detector bandwidth frequency value.

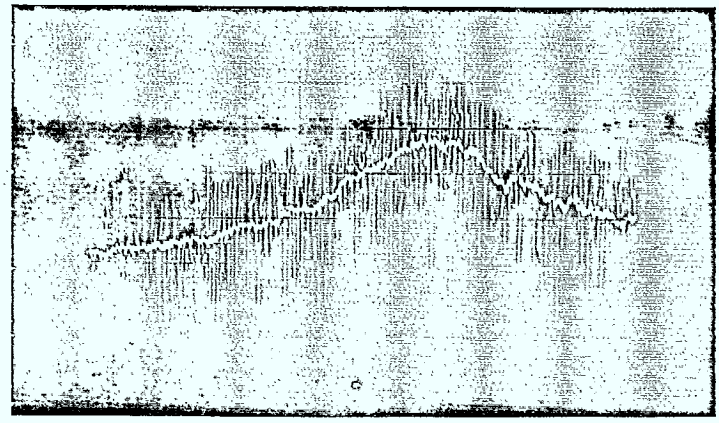
9.1.3 Laser Excess Noise Measurement

The laser AM noise has been measured on a spectrum analyser after amplification of the detector output. The amount of received laser power has been monitored simultaneously by measuring the detector signal current. The equipment used has been detailed previously (Section 8.2.2). The experiments have been done on both lasers (local oscillator and transmitter) for a wide current range and for laser cw and pulsed mode of operation. The complete set of results is not presented here. The two lasers behave in a similar way and the conclusions of these experiments are applicable for both the local oscillator and the transmitter. Two typical pictures of the total noise collected at the detector output are presented in Figure 9.1.1. They correspond respectively to the system noise with and without laser radiation incident on the photodiode. The amount of laser noise is obtained from the difference between these two noise levels, after subtraction of the shot noise contribution which can be calculated from the signal measurement. The graph in Figure 9.1.2 presents the laser noise for different laser currents. As described by the theory, the resonance frequency increases with the laser current while the noise amount follows an inverse current variation. The variation of the resonance frequency versus the laser current is compared to the theoretical one in Figure 9.1.3. A good agreement is obtained between the two curves. The next graph (Figure 9.1.4) shows the theoretical fit of one resonance peak. The maximum value is normalized to that of the experiment. The last figure of this section (Figure 9.1.5) shows the noise spectrum measured in the case of a pulse

Frequency Range: 0 → 1 GHz
Bandwidth: 1 MHz
Laser Current: 60 mA
Received Power: -13 dBm

-70 dBm-

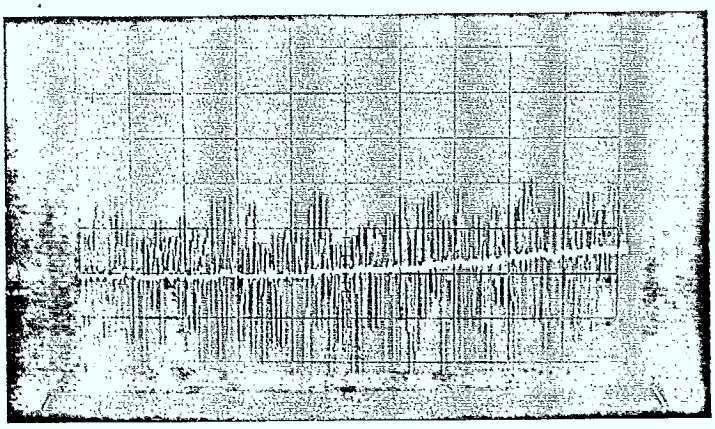
-90 dBm-



Part a: System Noise with Laser Radiation

-70dBm-

-90 dBm-



Part b: System Noise Without Laser Radiation

FIGURE 9.1.1: System Noise Spectrum Measurement

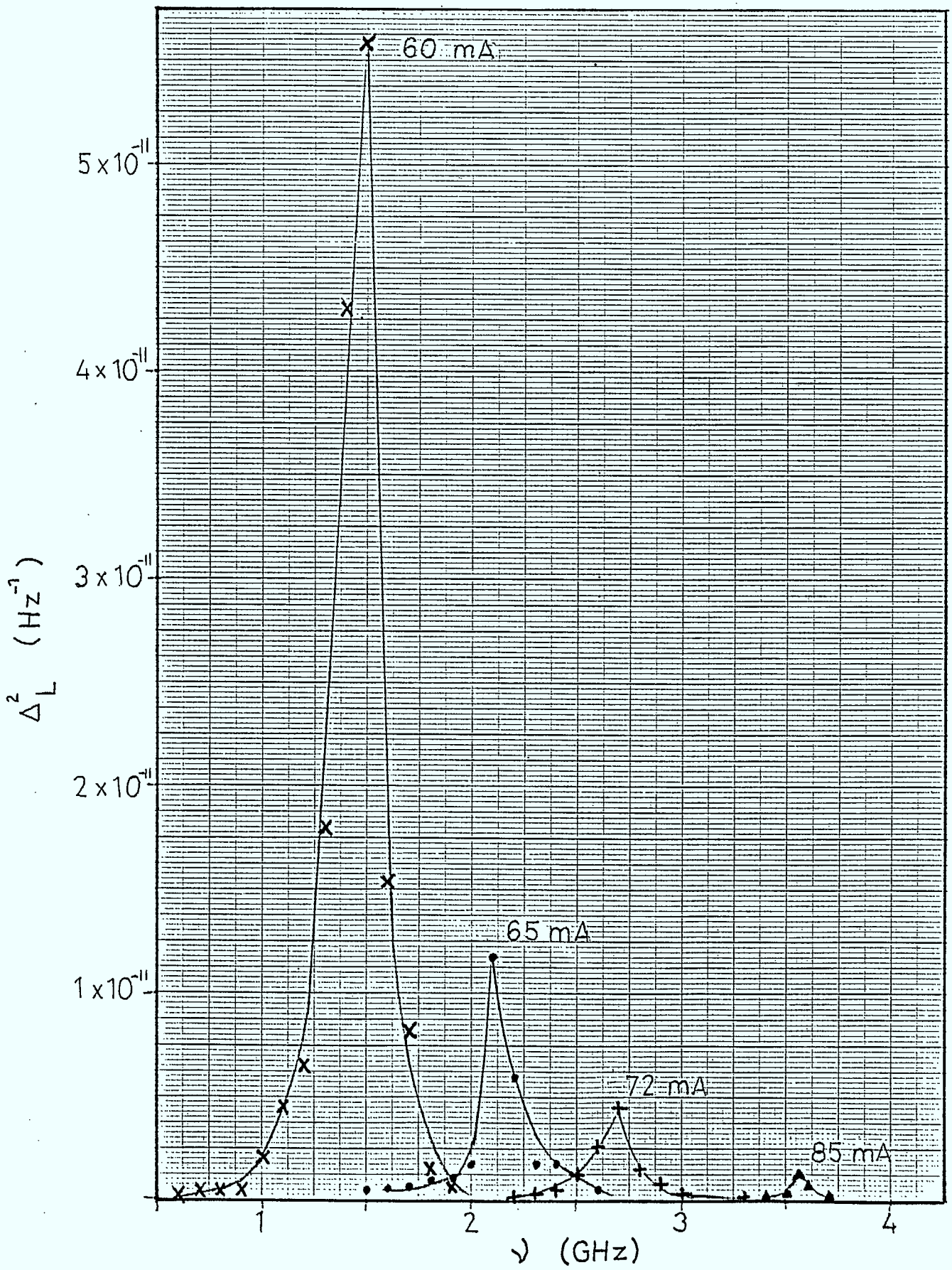


FIGURE 9.1.2: Laser AM Excess Noise

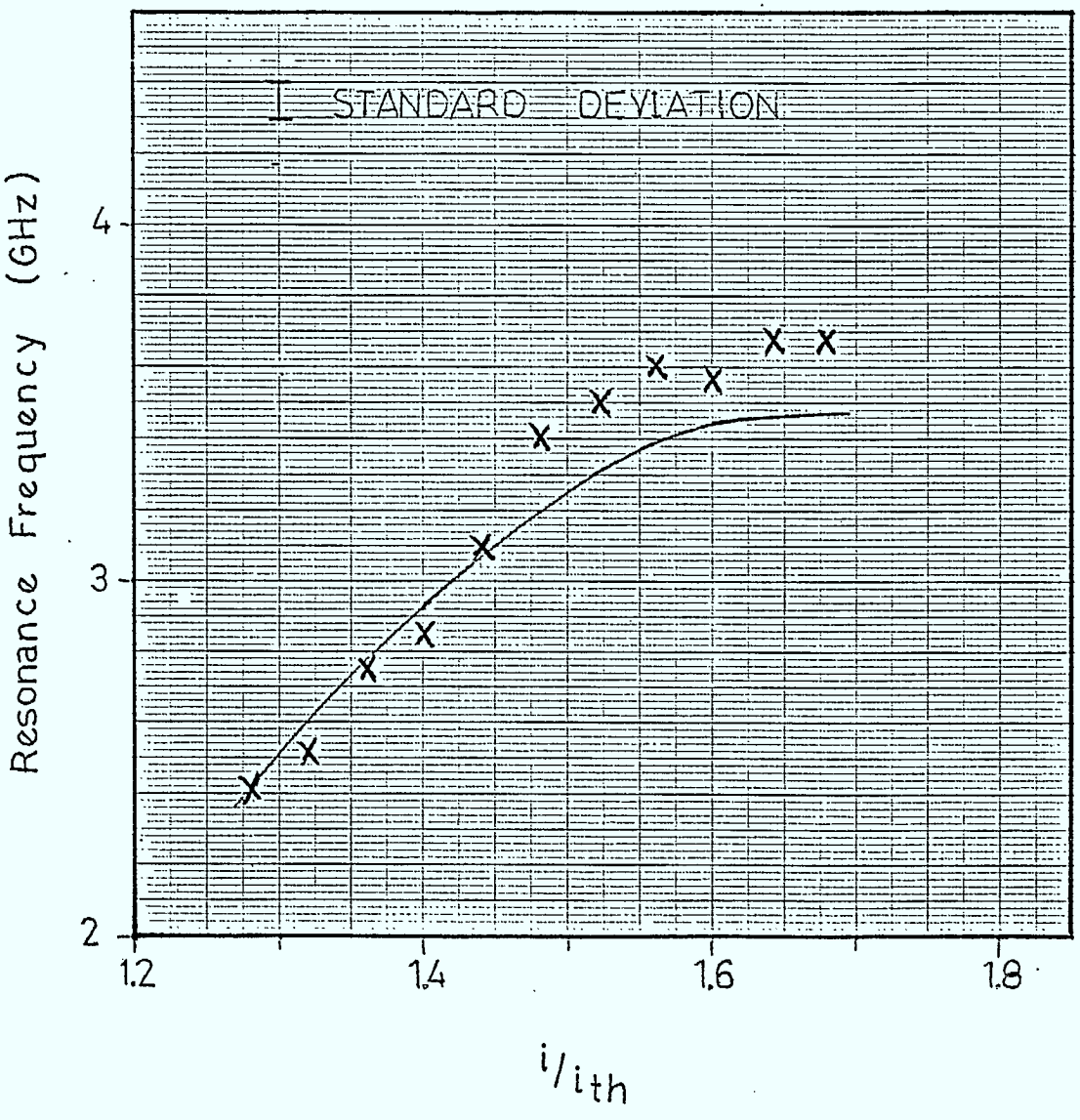


FIGURE 9.1.3: Resonance Frequency Versus Laser Bias Current

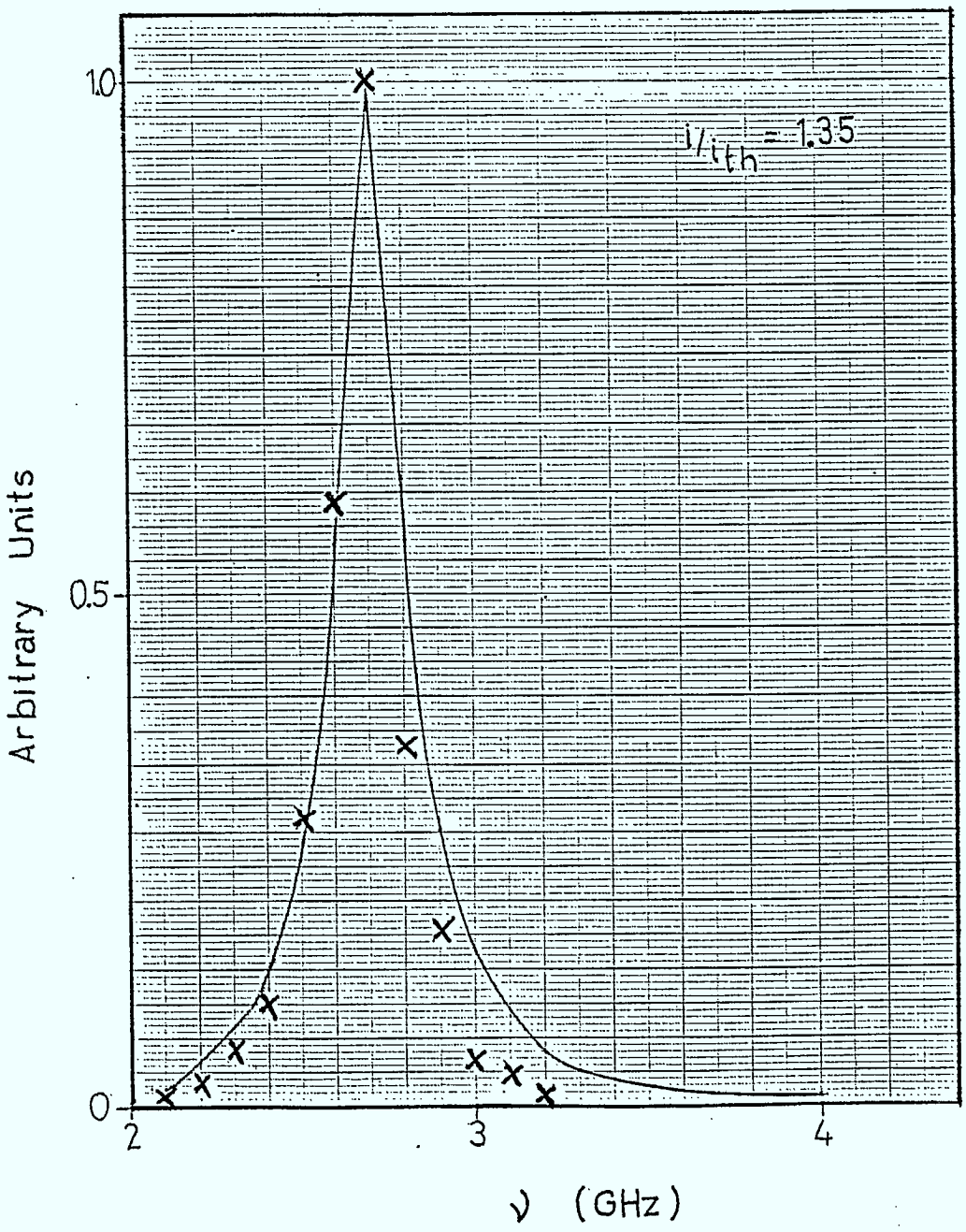


FIGURE 9.1.4: Theoretical Fit to Measured Laser Excess Noise

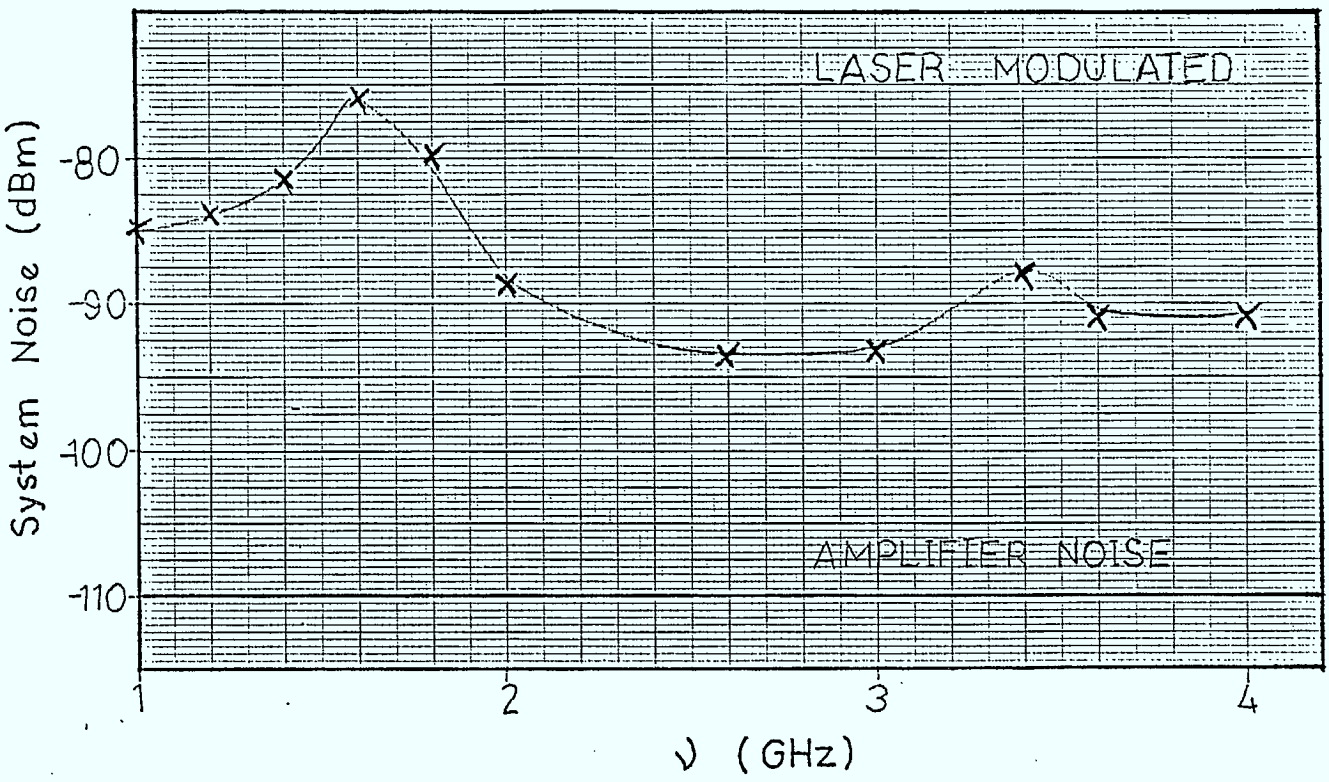
operation of the laser and compares it to the noise spectrum obtained for a cw laser radiation. A square pulse (100 MHz, 50% duty cycle) of 20 mA is applied on the laser. The lowest current portion of the pulse drives the laser just above its threshold while the highest portion corresponds to a current value close to 80 mA. By comparing the noise spectra for cw and pulsed operation, one observes that each portion of the square pulse generates the equivalent cw laser resonance peak.

9.1.4 System Total Noise Measurement

The total noise variation versus the amount of power is presented in Figure 9.1.6 for two frequencies. As expected in a shot noise limited system, the noise, at low frequency, varies proportionally with the amount of received power. At high frequency, the laser noise dominates and the system noise becomes proportional to the square of the received power. These behaviours can have important repercussions on the envelope detection performance. This is discussed in the following section.

9.1.5 System Performance

System performance is usually evaluated in terms of the signal to noise ratio obtained under different operating conditions which in this case are the amount of received power, the APD gain, the laser current and the frequency window in which the signal is received. This section concerns the direct detection of cw laser radiation. Being confident that the laser and system noises are well described by Equations 9.1.2 and 9.1.7, as it has been shown by the good fitting obtained with the experimental results, the signal to noise ratio for different



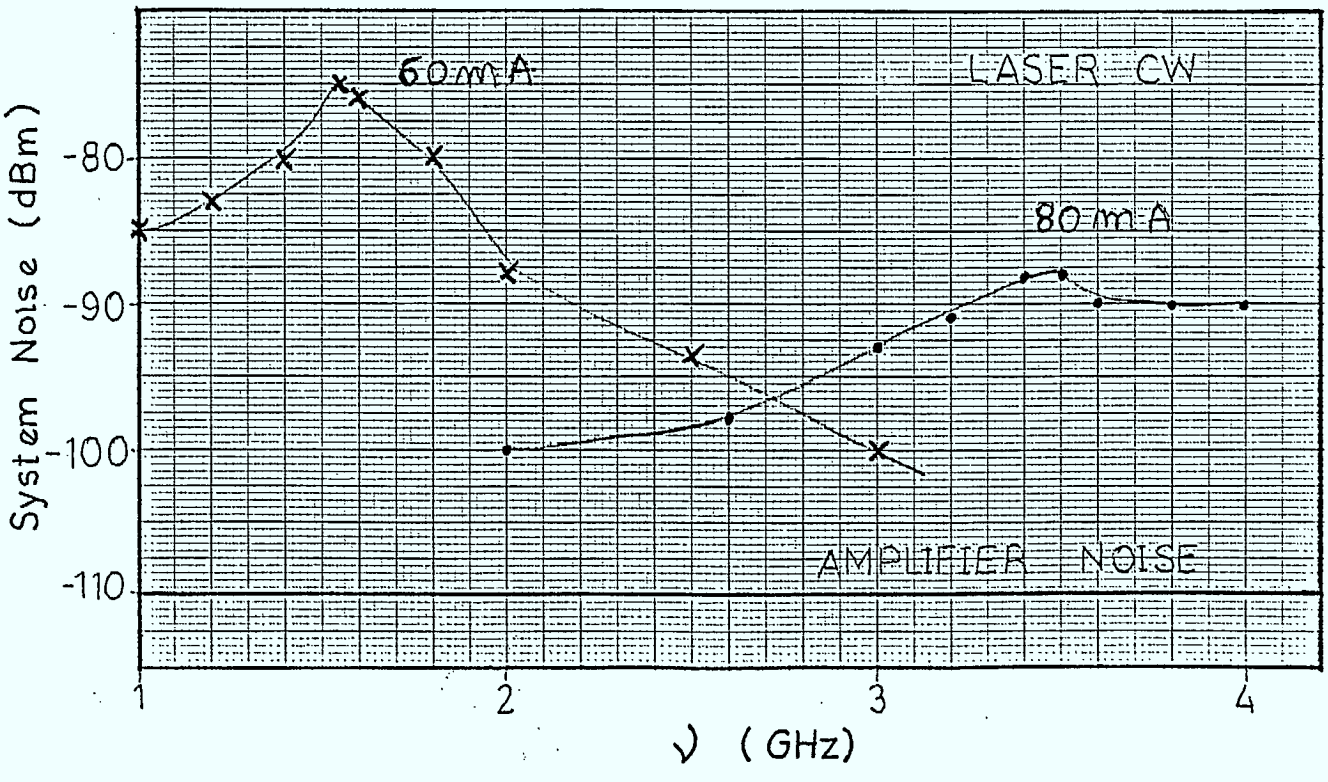


FIGURE 9.1.5: Comparison Between System Noise for Pulse and cw Laser Operation

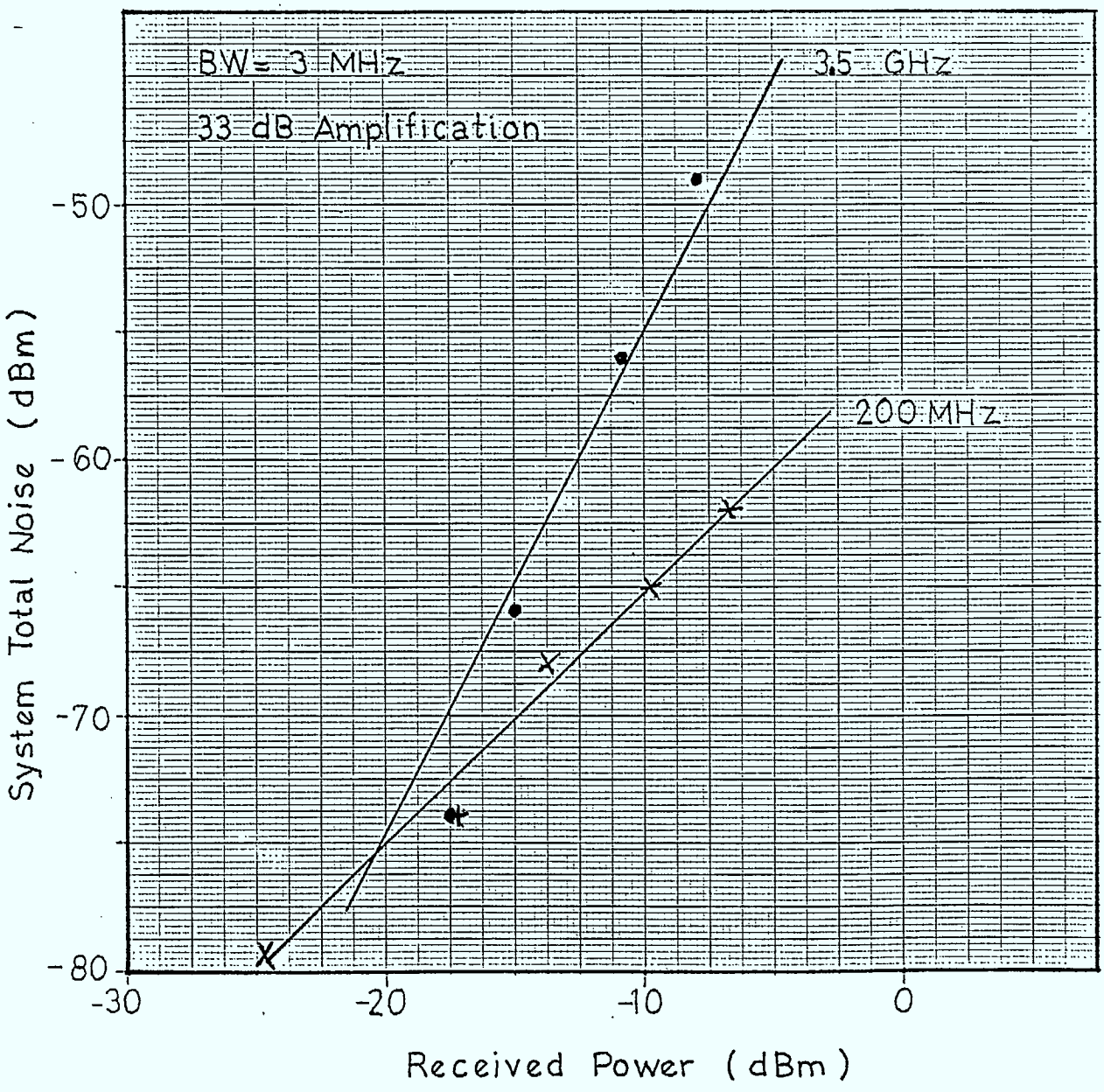


FIGURE 9.1.6: System Noise Variation as a Function of the Received Power

experimental conditions have been calculated. Figure 9.1.7 shows the S/N as a function of the received power for different APD gain values. As in all the following graphs, the calculated results are compared to a purely shot noise limited system (gas laser for example). It is seen that as soon as the system starts to be limited by the laser noise the APD gain does not influence the performance. Thus the gain should be adjusted as low as possible to ensure that the laser noise appears above the detector amplifier thermal noise level. The S/N variation as a function of the received power, for different frequencies is shown in Figures 9.1.8 and 9.1.9 at two laser current values. It is observed that the transition between a shot noise and a laser noise limited system occurs at lower frequency and at lower received power levels when the laser is operated relatively close to its threshold. Figure 9.1.10 shows the frequency and the received power combination at which the shot noise and the laser AM noise are equal. These curves are given for four laser operating currents and an APD gain of 1. They remain essentially the same for other gain values. All these performances refer to a photodetector with no preamplifier module (ANTEL type) and assume a receiver 1 Hz bandwidth. Depending on the experimental detection bandwidth, the ratio between the laser and shot noise amounts will vary. As a general rule, however, a system which involves low frequency detection (500 MHz) of low signal will be shot noise limited while the laser AM noise will dominate and deteriorate the performance when detecting high signal levels and/or when working at high frequencies. (gigahertz range)

Performances achievable for the detection of modulated

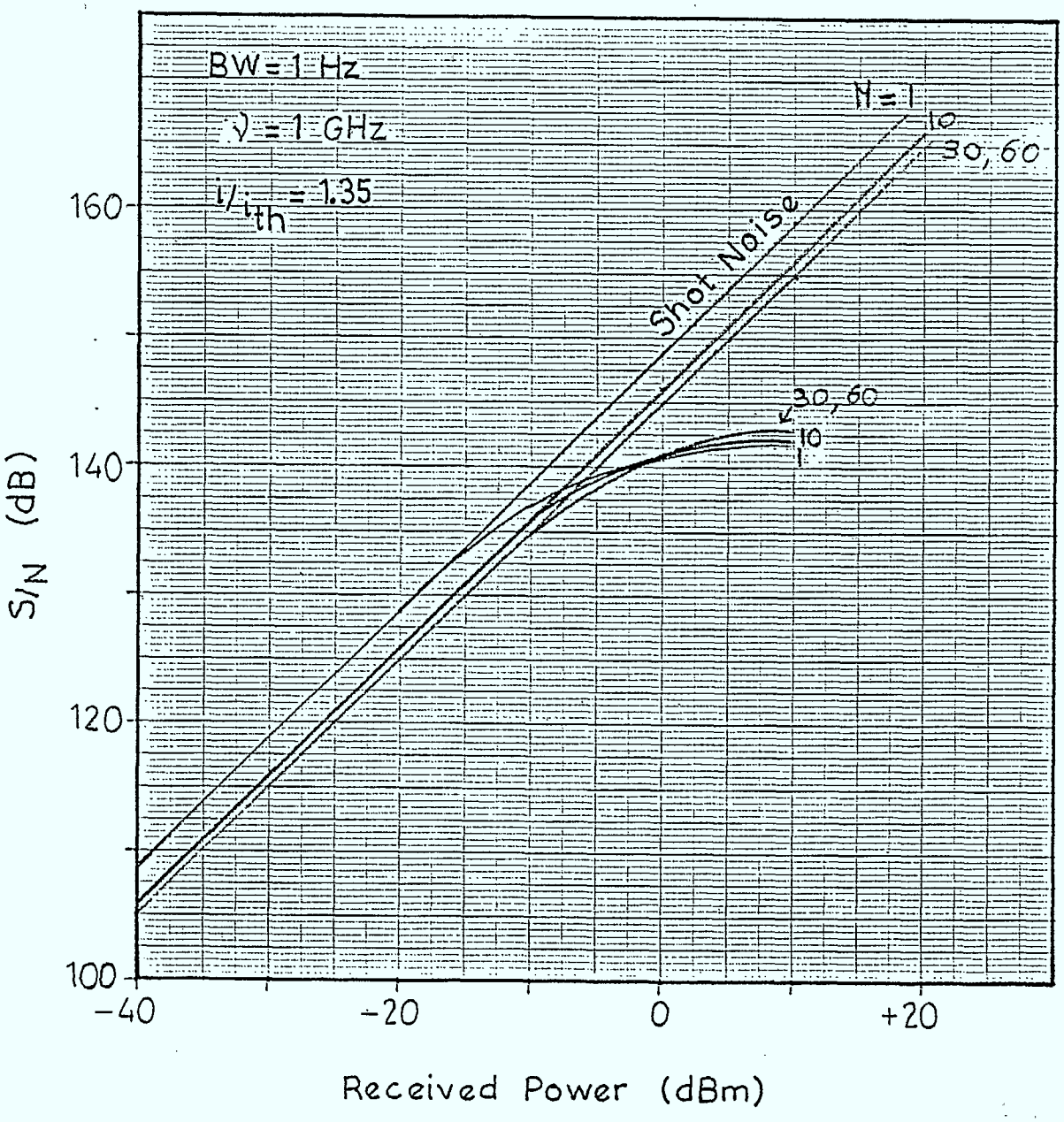


FIGURE 9.1.7: Signal to Noise Ratio as a Function of the Received Power for Various APD Gain Values

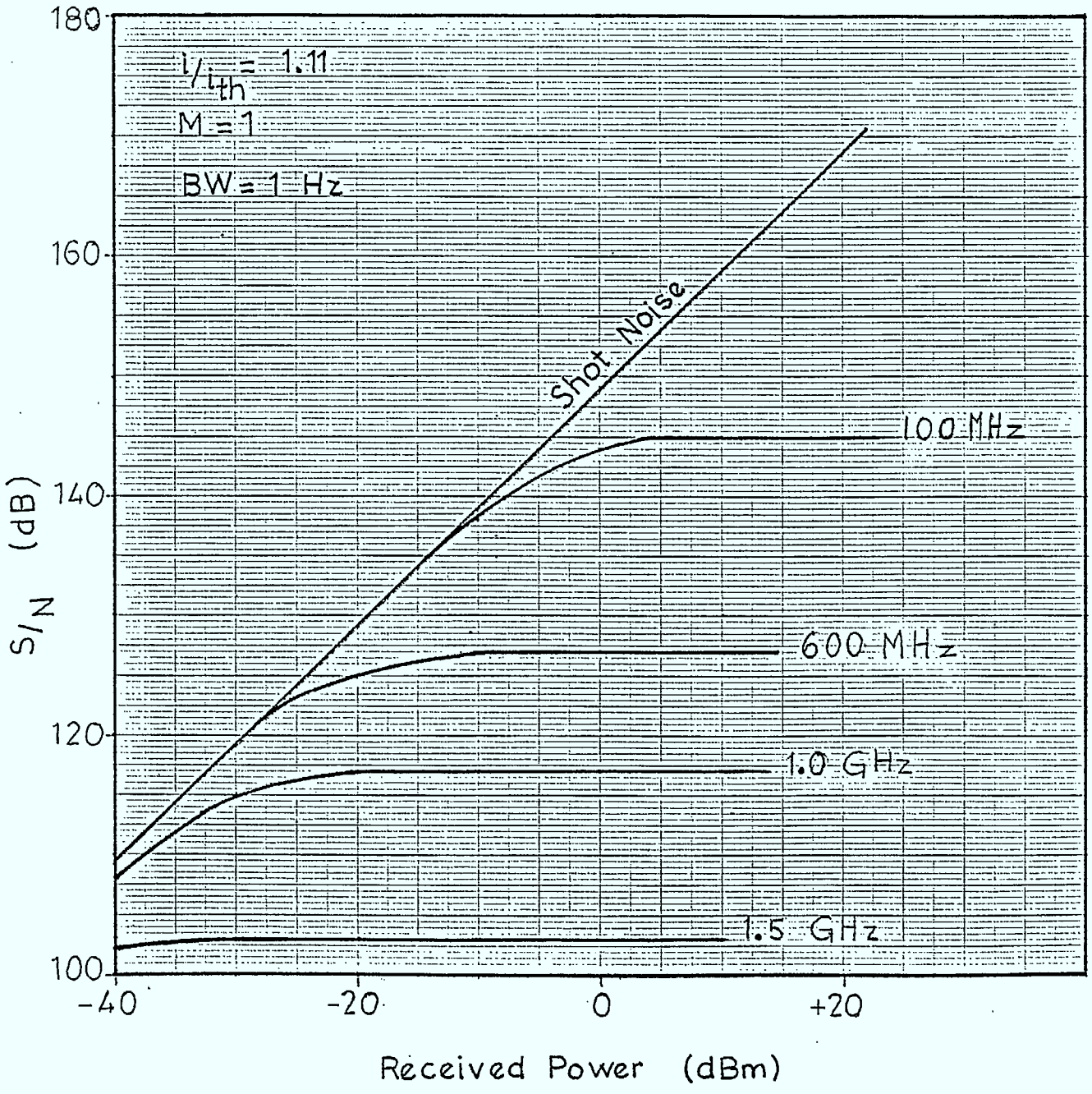


FIGURE 9.1.8: Signal to Noise Ratio as a Function of the Received Power at Various Frequencies

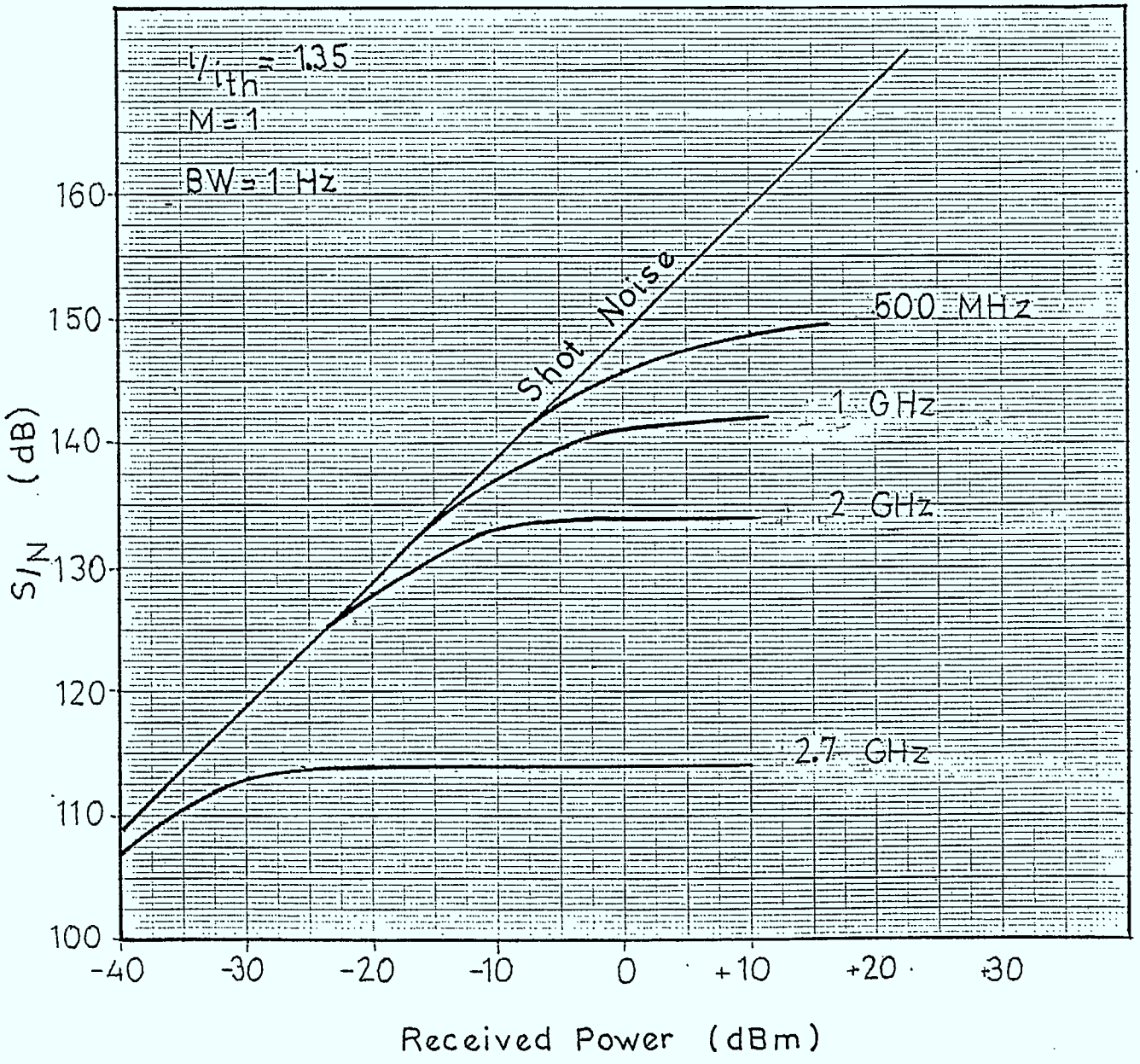


FIGURE 9.1.9: Signal to Noise Ratio as a Function of the Received Power at Various Frequencies

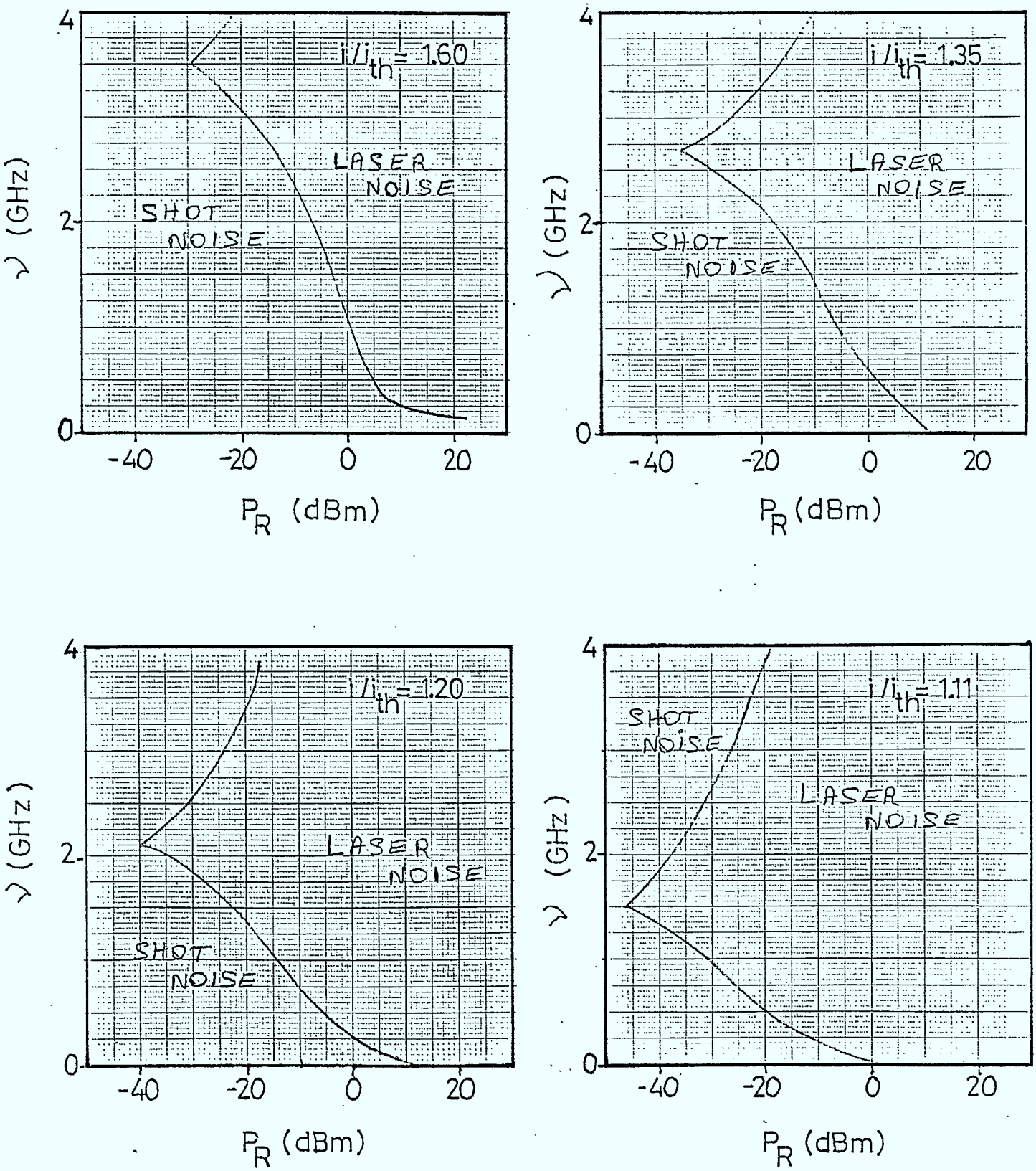


FIGURE 9.1.10: Laser Received Power and Frequency Values at Which the Laser Excess Noise is Equal to the Shot Noise

laser radiation, are much more complicated to evaluate. The experimental results presented in Figure 9.1.5 have shown that the laser excess noise contributes to both the "signal on" and the "signal off" portions of the radiation pulse. This is due to the minimum biasing current kept on the semiconductor laser to maintain its operating current above the laser threshold value. The consideration of this laser excess noise behaviour in the theory developed in Section 3 for pulse detection requires complex computer calculations which have not been done in this work. However, as a general comment based on the conclusions obtained for cw radiation detection, a system working at low frequency and involving low signal levels is expected to behave as a shot noise limited receiver. Such a system will follow the theory described in Section 3. When the laser noise starts to dominate, a degradation of the S/N is expected. This will result in a different optimization of the receiver characteristics. A decrease of the overall system performance is expected by comparison to a shot noise limited receiver, a decrease which can be relatively small if the system characteristics are carefully reoptimized.

9.2 Heterodyne Receiver Sensitivity

The system performance for heterodyne detection has been measured for cw and modulated laser operation of the transmitter. The two semiconductor lasers which are mixed together on the photodiode and the experimental testbed are described in detail in Sections 6 and 7. The detection frequency window is chosen to be 2 to 4 GHz. The detection corresponds to an envelope detection at the IF frequency. In order to minimize the laser

excess noise contribution, the lasers are driven at high current which gives a noise resonance peak frequency around or above 4 GHz. The results described here do not involve any type of stabilization feedback system except the temperature controller. A typical heterodyne cw signal is shown in Figure 9.2.1. The observed signal to noise ratio is 100 dB for a 1 Hz bandwidth detection system. In order to be sure that the two laser beams are well aligned and that the maximum available signal is actually measured, we calculate the heterodyne signal using the equation:

$$P_S^{Het} = 2R_0^2 M^2 P_T P_{LO} R_L \qquad 9.2.1$$

where R_0 and M are the photodiode responsivity and gain values

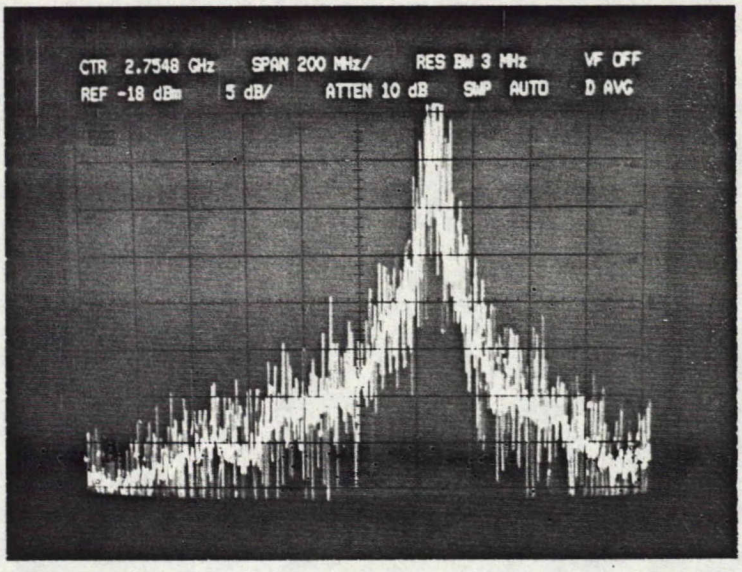
P_{LO} and P_T are the local oscillator and transmitter received laser powers

R_L is the system load resistance

Using this equation and the experimental parameters given in Figure 9.2.1, the heterodyne calculated signal is -10 dBm which is in a good agreement with the -14 dBm experimental value. This calculation takes into account the fact that the signal is spread over a wide frequency range (see Section 9.2.1).

9.2.1 Signal Frequency Width

A semiconductor laser linewidth follows typically a Lorentzian profile function. The heterodyne signal which results from the convolution of these two laser linewidths is also described by a Lorentzian equation. By fitting the



P_{LO} = - 9 dBm

$P_{Transmitter}$ = - 10 dBm

Frequency Range: 1.8 → 3.8 GHz

Bandwidth: 3 MHz

Peak Value: - 14 dBm

Amplification: 33 dB

FIGURE 9.2.1: Measured Heterodyne Signal (cw Case)

experimental signal to a Lorentzian function, an approximate value of the sum of the two laser linewidths is obtained. This fit is presented in Figure 9.2.2. Assuming that the two lasers have identical characteristics, a linewidth of 45 MHz is measured at the half maximum. This value is reasonable for the type of laser used in this work with no external cavity.

9.2.2 System Noise

The system noise collected at the detector output in the gigahertz range is essentially due to the lasers' AM and FM noises. At this stage of our work, it is difficult to discriminate between the relative FM and AM contributions since they are strongly related to each other. By comparing the measured noise level to the one obtained for similar laser power in the case of the direct detection, an increase of about 10 dB is observed. However based on other results reported in the literature,[9.2] the use of a dual detector receiver should allow a cancellation of the AM noise contribution which is claimed to be about 20 dB.

9.2.3 S/N Variation Versus the APD Gain

According to the signal equation 9.2.1, the heterodyne signal varies with the square of the photodiode gain. This is confirmed by our measurements as shown in Figure 9.2.3. It is interesting to observe that the noise presents the same behaviour and that the signal to noise ratio is thus constant. This noise power variation is expected and is identical to the laser excess noise variation observed for the diredt mode of detection, proving that this heterodyne receiver is laser noise limited in the gigahertz range. Such a system will be considerably improved

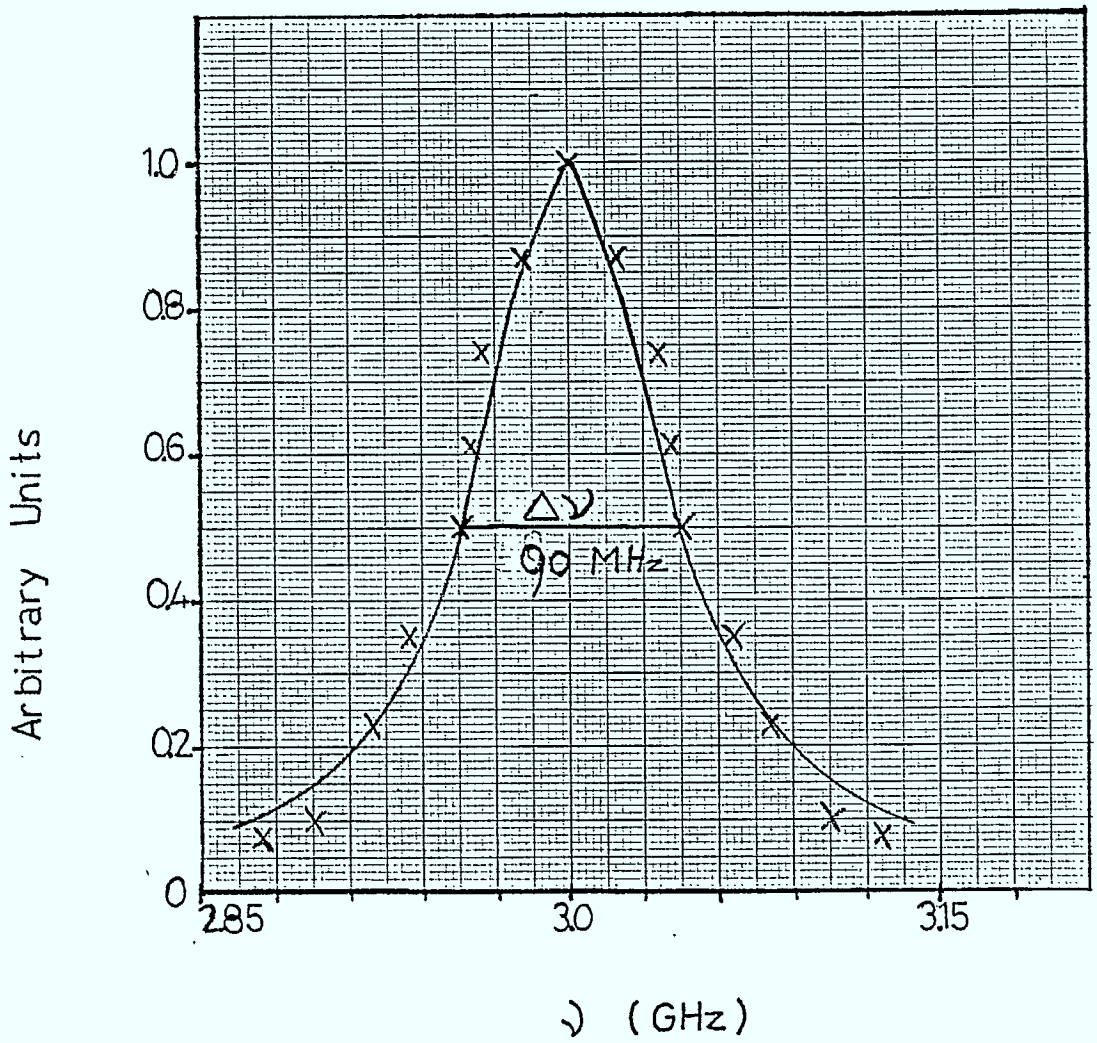


FIGURE 9.2.2: Fit of the Heterodyne Signal to a Lorenztian Profile

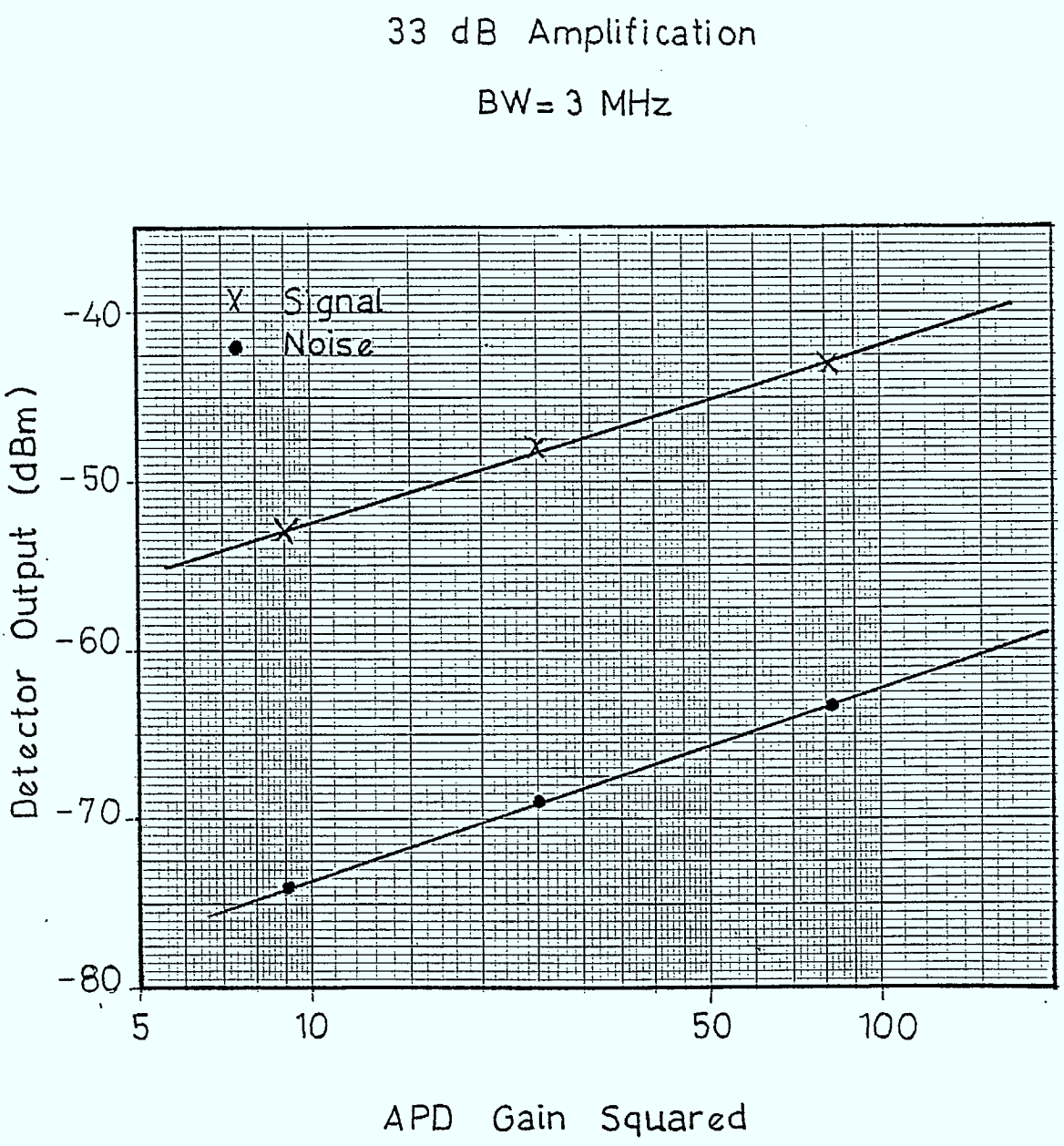


FIGURE 9.2.3: Heterodyne Signal and Noise Variations as a Function of the APD Gain

by using a dual detector scheme.

9.2.4. FSK Heterodyne Experiments

The feasibility of FSK modulation on our system has been tested up to 1 Gbit/s using a 1010 NRZ pulse. The laser frequency was modulated by applying a current modulation of a few milliamp around the laser DC bias current. The current pulse was a square pulse of 500 MHz repetition rate with a current amplitude adjustable from less than 1 milliamp up to 10 milliamp. At this stage of our work, the detection system was very simple. The signal was detected on a spectrum analyser after a suitable amplification of the detector output. The results are presented in Figure 9.2.4. The experiment corresponds to a heterodyne IF signal at 2.8 GHz when no frequency modulation is applied on the transmitter laser. When the laser frequency is modulated, the heterodyne signal displays two frequency components, f_0 and f_1, which correspond respectively to the transmission of a zero and a one signal (see also Figure 6.2.1b).

The frequency spacing between these two components constitutes a measurement of the laser frequency shift induced by the laser current modulation. This shift increases with the current modulation amplitude.

Around each of these frequency components, sidebands spaced by 500 MHz (pulse repetition rate) are observed also.

Figure 9.2.4a shows the signal observed for a +/- 2 mA current modulation. The two frequency components (f_0 and f_1) are collected inside the frequency window set on the spectrum analyser (1.75 GHz - 4.25 GHz; 500 MHz per division). From the

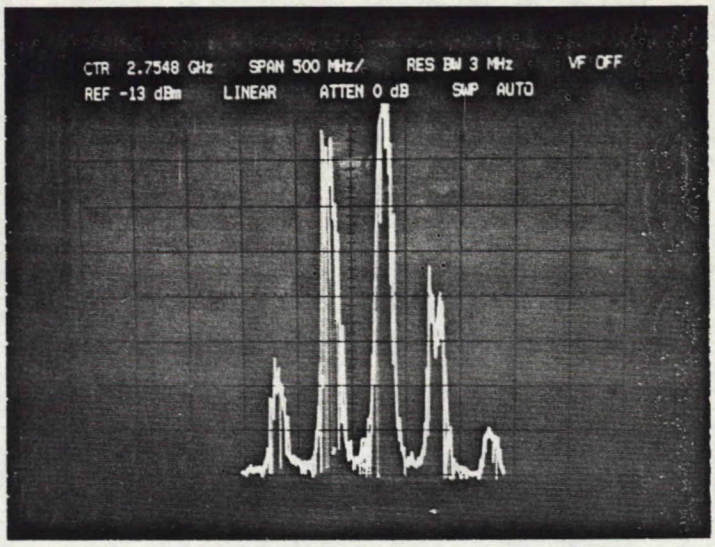
frequency spacing measured between them, a laser frequency shift of 250 MHz/mA is obtained. Higher shift frequency is observed if the amplitude of the current modulation is increased. This is shown in part b of the same figure. In this case only f_1 is detected in our spectrum analyser frequency window. f_0 appears somewhere above 4.25 GHz indicating a laser frequency shift of at least 325 MHz/mA. These experimental results are very promising, regarding the capability of our heterodyne system to operate in a FSK mode. From Figure 9.2.1a, it can be seen that no important amplitude change is observed between the two frequency components.

9.3 Comparison with Other Works

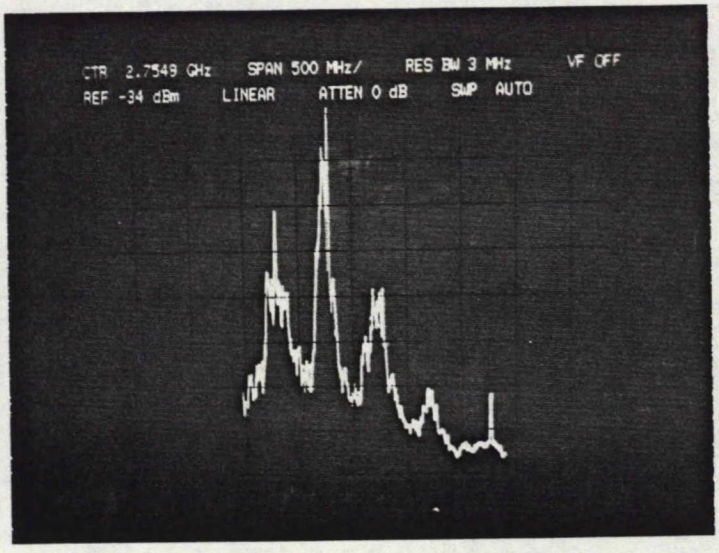
An extensive number of experimental and theoretical studies on direct detection and heterodyne communication systems using semiconductor lasers have been published in the last 10 years. These works correspond to a wide range of system design, receivers and modulation formats. Some of these state-of-the-art systems have been discussed in Section 4. In this section, the experimental results obtained with our testbed are compared to published performances of similar systems.

9.3.1 Laser AM Excess Noise

Our measurements of the laser AM excess noise can be compared to the results given by Yamomoto[9.3]. From his data (Reference 9.3 Figure 8.a) which gives the laser noise normalized to the shot noise, both referred to the laser output power, a value of the received laser noise equivalent to our factor Δ^2_L is obtained at different laser bias and compared to our measured values in Table 9.3.1.



Current Modulation: +/- 2 mA



Current Modulation: +/- 4 mA

FIGURE 9.2.4: Measured Heterodyne Signal (FSK Case)

TABLE 9.3.1: LASER AM EXCESS NOISE

	This Work	Reference 9.3
$i/i_{th} = 1.11$	$\Delta^2_L = 5.6\times10^{-11}Hz^{-1}$	$6.3\times10^{-11}Hz^{-1}$
1.20	1.2×10^{-11}	1×10^{-11}
1.33	0.45×10^{-11}	0.43×10^{-11}
1.57	0.15×10^{-11}	0.23×10^{-11}

A good agreement is obtained between the two sets of values.

9.3.2 Direct Detection

No direct detection performance results have been reported by Yamomoto and his co-workers to which ours may be compared. A study similar to ours,[9.4] but involving a different laser reports an important decrease of the S/N due to the laser noise compared to the shot noise limited receiver. In their case, a drop of 10 dB is observed at 50 MHz for a received power value of -10 dBm. By further increasing the laser output, no improvement is obtained and the S/N stays constant at 120 dB. Looking at our results presented in Figures 9.1.8 and 9.1.9 , similar behaviour is observed except that the laser noise appears at higher frequency (above 300 MHz or 1 GHz depending on the laser bias current). The final achievable S/N constant value is also higher and is about 140 dB. This comparison points out the importance of choosing carefully the type of semiconductor laser which is

going to be used in relation to the application (bit rate, received signal level, receiver type etc.).

9.3.3 Heterodyne Detection

A typical heterodyne signal obtained with our system as the one shown in the previous figure 9.2.1, can be compared with no restriction to the signal reported by Saito[9.5]. The lasers are identical and operate at similar power levels. In the work of Saito, an observed S/N of 36 dB is reported which is very close to the one measured with our set up.

Results with FSK modulation reported by the same workers[9.6] shows a laser frequency shift of 200 MHz/mA of laser current modulation at a frequency of 500 MHz. This is similar to our performance which appears to be 250 up to 325 MHz/mA.

In conclusion, this comparison between our results and performances reported in the literature for similar systems has shown that our testbed gives good performances compared to what was expected. This allows us to extrapolate that the addition of other more sophisticated subsystems (AFC loop, external cavity, dual detection scheme, etc.) should bring our system performances to the actual state-of-the-art.

9.4 Summary of Accomplishments

A number of experiments have been performed on a laser diode communication testbed, in order to study the possible use of such a system for ISL.

A first set of results concerns various detector performances. Different types of Silicon photodiodes (pin diode and avalanche diode) have been studied in relation to their use in a wideband heterodyne receiver. Detailed experimental and theoretical work has been done on the detector noise.

A second series of experiments are related to direct detection. The receiver performances and the noise contribution induced by the laser itself (laser AM excess noise) have been studied in great detail.

The heterodyne mode of operation of our system has also been investigated. For these experiments, the laser transmitter has been operated cw and in a FSK modulation mode.

These experimental results give us useful information on the performances of a laser diode communication system for direct and heterodyne detection modes. They also point out the effect of the laser excess noise on the overall system signal to noise ratio. The performances achieved in this work show the strong relation existing between a particular system requirement and the type of receiver which has to be used.

9.5 Impact of the Results

Many studies have been made for possible uses of diode lasers for ISL.[9.7-9.10] Comparison of GaAlAs laser systems with others such as the Nd:Yag and CO_2 lasers shows that although no definitive superiority exists among these laser systems,

the GaAlAs diode based systems have been found to hold appealing promise. The main advantages include small size, high transmitter efficiency, direct modulation capability, wavelength selectability and high reliability.

Laser power requirements for OOK envelope detection are shown through curves such as [9.8] Fig. 9.4.1. It is seen, for example, that an optical aperture of the order of 30 cm would require about 100 mW of laser power for a range of 35 - 40 km at 500 Mb/s. Most laser diodes cannot produce such power levels in single mode operation. Power in the 35 - 50 mW range can be expected at present. Projected power in the 100 mW level is expected in the early 1990's.[9.10] However, special techniques are employed presently which can bring up the power level of today's lasers substantially. For example Casey[9.11] has used power combining schemes where 6 single mode lasers have been combined to produce 250 mW peak power at 500 Mbit/s.

If heterodyne systems are used, the receiver sensitivity can be improved by 10 - 20 dBs.[9.12] This then drops the power requirement by the same amount or alternatively, reduces the size of the optics necessary for equivalent performance.

In our own work, we have studied the power requirement of laser systems as a function of S/N and thus error bit performance (see section 3.3 and 3.4). Also, laser and detector noise have been studied experimentally (see section 9.1).

It was shown that diode lasers produce excess noise which is not present in other lasers such as gas lasers. This can seriously affect the performance of laser systems under certain conditions. It is therefore useful to examine whether or not in

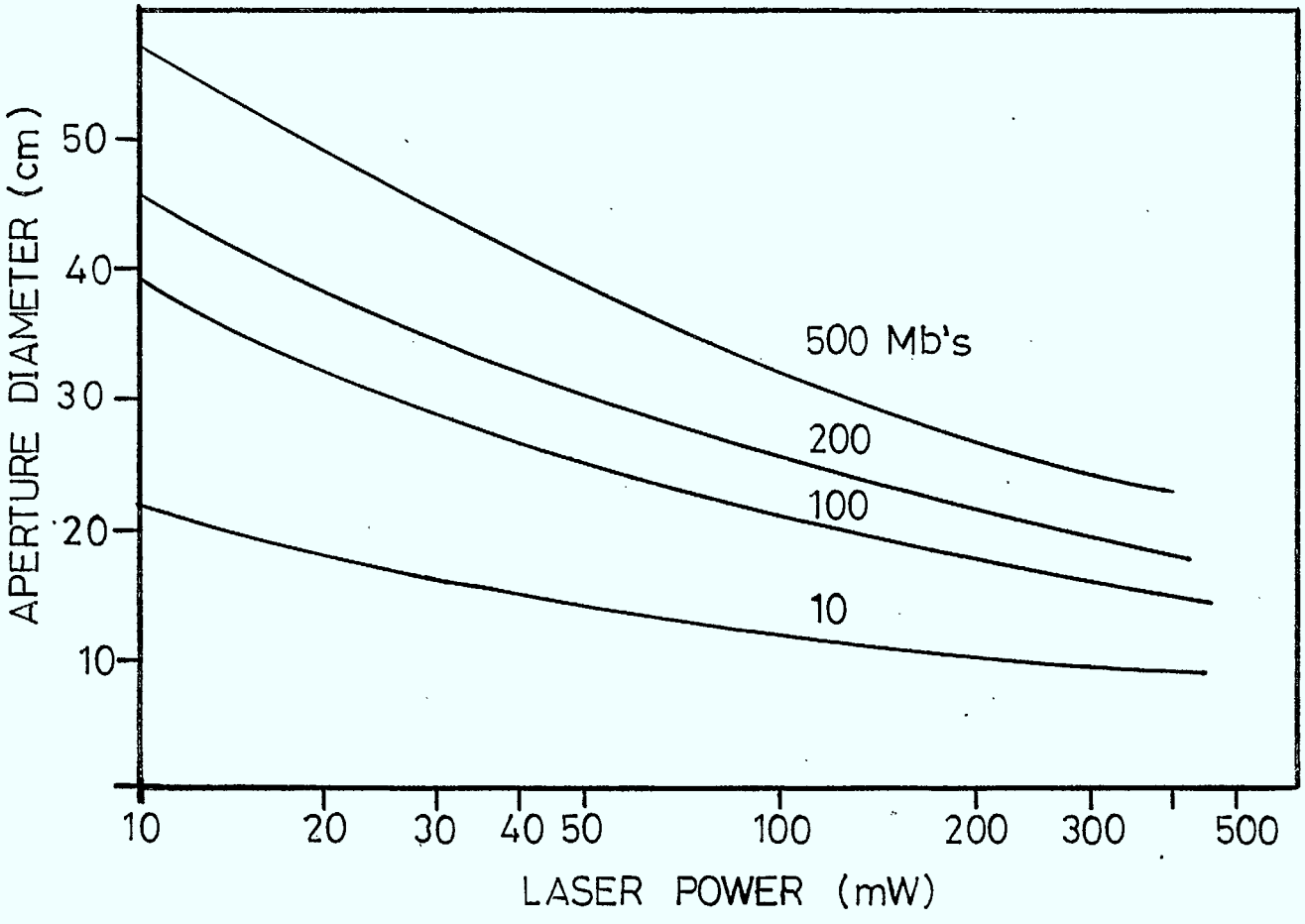


FIGURE 9.4.1: Aperture Size Dependance on ISL Distance

our own case this laser excess noise is consequential.

Let us take, as a starting point, power requirement as determined in Section 3.3.2 for an integrated pre-amp type detector. For a 1 Gbit/s system (p = 1 ns) we find from Fig. 3.3.5 that for a BER of 1 part in 10^9, the laser power required is about -35 dbm. Now, let us look at Fig. 9.1.10 for i/i_{th} = 1.60. The curve defines the frequency and power values at which the laser excess noise and shot noise are equal. It is seen that at -35 dbm, the laser excess noise is smaller than the shot noise during the "on" part of the pulse. We assume here that the "off" power of the laser is close to zero so that its contribution to excess noise is negligible. We can conclude then that for the above specifications, and for the laser we used, the excess noise should not be a problem.

These remarks pertain to direct detection. Our heterodyne measurements have not been pushed far enough at this point to allow a conclusion as to the impact of the laser FM excess noise on a communication system.

In general, our performance results reported in previous sections confirm predictions made by others regarding the potential of laser diode in ISL. The additional effort required to bring the technology closer to its ultimate aim is the subject of the next section.

10.0 RECOMMENDATIONS FOR FUTURE WORK

This study has resulted in a characterization of the applicability of optical heterodyne detection technology to ISL systems, an identification of the important system elements, a definition of the key areas of technology and the review of potential solutions to problem areas.

Based on this information, recommendations will be made regarding the direction of further work effort in the development of the key areas of technology for the purpose of providing Canada with a capability for competing in the supply of subsystems for perceived ISL links utilizing optical direct and heterodyne systems. At this point, the possible users of this technology include Intelsat, NASA, ESA and military agencies.

10.1 Recommendations

The scope of this study was very wide ranging. As a result many systems and general design issues received attention. It is now recommended however, that based on the work presented in this report, and from known efforts by the international ISL community, our own effort should be directed towards the development of improved optical heterodyne/direct systems. The particular modulation/detection format may be chosen based upon the analyses performed in this study. Although this system would have its design based primarily upon one modulation/detection format, it should still be made general enough to accommodate, as much as possible, different operational parameters such as transmission rates and modulation indices. As such, this system would still be considered to be an experimental laboratory system. In addition, the underlying technology would be

adaptable to a specific ISL requirement, be it based on a heterodyne or direct detection system.

Whereas the current study has been concerned with general measures of performance and implementation, the development of a defined system would require the identification of more specific technical problems which in turn would require the development of detailed technical solutions. In this way, specific technical expertise can be developed and a large jump can be made on the learning curve.

The current study has determined the critical subsystems involved in the different design approaches to optical envelope and heterodyne detection. The proposed systems work would involve detailed analysis of these subsystems. Subsystems that would require specific attention would be:

i) the development of sensitive, state-of-the-art receiver electronics, including low noise amplifiers, demodulation and regeneration circuitry.

ii) the development of an improved optimum bandwidth, automatic frequency control circuit with maximized capture bandwidth.

iii) the development of a fully integrated laser driver optimized for the selected modulation format yet tunable over a range of operating conditions. Design should accommodate an ECL baseband message signal. In other words, the baseband signal need not be modified to vary the laser operating condition. The baseband signal may be supplied by a transmission test set or by a pulse generator as done in the

current study.

Secondary priority should be given to optimizing the beam combining procedure. It is recommended that the detector and local oscillator be integrated into a single package for this purpose. Although it is not believed that beam expansion would be a necessary issue for this system, a certain amount of consideration may be given to collection and transmission telescopes. The inclusion of a pointing and tracking subsystem would be premature at this stage of development although it will eventually represent a critical ISL subsystem.

It is recommended that the design of this system be complete enough to permit performance measurement in sufficient detail to allow comparison to state-of-the-art in both heterodyne and direct detection systems. Performance measurement may be performed using either sensitivity vs. bit rate for a specified BER or BER vs. received optical power for a specified bit rate, as deemed appropriate.

Certain technical problems, unique to semiconductor laser heterodyne systems, have been identified in the current study. Examples include intrinsic laser noise due to resonance phenomenon, laser phase noise resulting in broad linewidths and laser instability caused by reflections into the laser cavity from external components. The severity of these problems to system performance is generally dependant upon the operating condition of the laser as discussed in this study (i.e. modulation format, detection technique, transmission frequency, bias currents, etc.). Solutions to these problems should be addressed as required in the development of the experimental

system. This should become evident in the design phase of the system.

It is believed that the development of such an experimental system is required to fully establish a Canadian systems expertise in optical heterodyne detection technology for intersatellite link application.

In addition to the two global recommendations made above the following observation is noted as a result of the systems analysis performed in this study. The performance analysis has shown that at high transmission rates (i.e. greater than 1 GHz), the performance of an optimized intensity modulated direct detection system approaches the performance of an ASK heterodyne system at a much reduced level of system complexity. However, this must be qualified by the fact that the analysis assumes optimum APD gains which are not always practical at the current level of technology. Therefore more theoretical work is required to bring out more clearly the conditions under which direct detection is definitely superior. In any case, such a situation indicates that some consideration should perhaps be given to the development of a sensitive direct detection system for operation at transmission rates exceeding 1Gbit/s. Of the direct detection formats analysed in this study, the optimum candidate would appear to be a QPPM system. Investigation could also be made into the development of laser arrays and appropriate beam combining schemes in order to increase manyfold the maximum transmittor power available in a direct detection system. Although the development of such a system is potentially very valuable, it should not preclude further effort in optical heterodyne

detection. The performance gains in heterodyne detection are significant at lower bit rates and possibly at higher bit rates in the case of future more sophisticated heterodyne and homodyne systems. In fact, a point may be reached where neither system will provide the required sensitivity at high bit rates, given the limitations on practical antenna size and actual laser output powers. This sensitivity may be obtained however through the use of a dual multiplexed heterodyne or m-ary heterodyne systems at a reduced bit rate.

In addition to the development of the system described above it is further recommended that consideration be given to research in particular difficult technical areas in semiconductor based heterodyne detection systems regardless of the actual system implementation. Reduction of the effects of intrinsic laser noise through the development of a dual balanced detector could involve one such area of research. Excess transmitter noise may be induced by poor stability in the driver electronics therefore the development of low noise transmitter and receiver electronics may be investigated. More theoretical work is also required to determine more clearly the degradation of a laser system due to excess noise.

The reduction of laser phase noise can be very important to coherent demodulation techniques and to efficient AFC loop operation. Reduction of phase noise is mandatory if PSK homodyne detection systems are ever to be realized. Reduction of laser phase noise may be accomplished using external cavity techniques or through the design of new laser structures, or through a combination of both. The design of efficient tunable stable

external cavities is a very demanding technical problem.

The investigation of laser frequency control systems other than the electronic AFC loop already discussed could be potentially important. Such a system may involve the use of optical or opto-electronic feedback techniques. Although techniques such as injection locking presently appear very difficult to implement, they may offer an interesting field of study with potentially valuable applications.

Finally, further development of external modulation techniques may be beneficial. Although the use of external modulation eliminates the direct modulation advantage of injection lasers, these techniques may be valuable at high bit rate operation if insertion loss and drive voltages can be reduced. The modulators must also be capable of handling the power densities expected in these systems.

From a device perspective, any development work resulting in an improvement in laser operating characteristics such as higher frequency resonance, lower intrinsic noise, higher power, and greater frequency stability would be very beneficial as discussed in Chapter 7. Also, the development of high bandwidth (up to 10 GHz), low noise and high sensitivity detectors is very important to ISL applications. These detectors should have high quantum efficiencies at the short wavelengths where the laser operates.

10.2 Canadian Participation

In order that Canada could have a place as a supplier of ISL technology, it is essential that there be in this country, strong on-going R & D program in this field. This program should

be relevant and should be adaptable to follow the requirements of potential users.

The key areas of technology that are relevant today include a systems and integration capability, an electronic design capability encompassing both integrated circuit design and microwave electronic design, a capability in device technology referring to high bandwidth detector design and laser development, and a capability in optical design and packaging. Of these key areas the most critical is the systems capability which involves not only systems engineering but also a fundamental understanding of the science involved with optical envelope and heterodyne detection. Expertise in communication theory and experience in supply of space qualifiable hardware are also vital areas of technology.

Canada appears to be well adapted to many of these key areas. With the commissioning of this study a basic systems capability has been developed although a continued research and development effort is required to increase this capability to the supply stage. The capability to pursue the recommendations discussed in Section 10.1 certainly exists at present and should not be lost.

As discussed in this report, operation at frequencies in the GHz region requires the use of microwave electronic components and microwave design techniques. Examples of these components were given in the system description. Many firms exist in Canada which have the potential to become suppliers of these components.

The high bandwidth detectors used in this study were

purchased from a Canadian company. Two such high speed detector suppliers currently exist in Canada with a third being a recognized supplier of reliable, low noise, state-of-the-art detectors. In addition, there is one laser diode manufacturer in Canada, however this manufacturer does not produce a laser appropriate for this application at this time.

Canada also has experience in the supply of space hardware along with the facilities for space qualifying hardware at the NRC Florida Laboratory. Canada has developed many communication satellites based upon conventional microwave technology, and is currently involved in the space station program. Much of this experience should be transferrable to optical communication system development effort.

It should also be noted that certain specific areas of research such as laser diode device research, external modulator research, high speed detector research, and research into various areas of integrated optics can be supported in Canadian universities and research institutes.

In addition to the laser ISL development, the kind of research recommended here has the further benefit in that the experience and knowledge that will be gained can be spun off into other areas such as,electro-optic microwave technology.

Canada's reputation in satellite communications is well established worldwide. In order that this expertise remains competitive in this fast moving technical field, Canada has to enter the era of Optical ISL in earnest for fear of being left behind. This can only be done through resolute further committments in R&D effort for laser satellite communications systems.

11.0 CONCLUSION

The objective of this study was to define and develop key areas of ISL technology so as to put Canadian industry in a favourable position in competing for international supply contracts on subsystems for ISLs. The study was to emphasize the development of technology for heterodyne detection using semiconductor lasers.

The objective was achieved through a general analysis of a number of systems which evaluated performance and complexity. In conjunction with this systems analysis, an experimental testbed was developed which was used to gain hands-on experience with the fundamental technology required in a free space optical heterodyne communication system. In addition, a survey of the state-of-the-art and direction of current research effort was performed.

ACKNOWLEDGEMENT

The authors wish to thank Dr. B. Felstead of C.R.C. for his continuing encouragement in this project as well as his helpful comments in connection with the preparation of this report.

REFERENCES

Section 2

2.1 Y. Yamamoto, "Receiver Performance Evaluation of Various Digital Optical Modulation-Demodulation Systems in the 0.5 - 10 μm Wavelength Region", IEEE, JQE, Vol. QE-16, No. 11, Nov. 1980.

2.2 T. Matsumoto and S. Shimada, "Advances in Coherent Optical Transmission Systems", ThJ2, CLEO '86, San Francisco, CA. June 1986.

2.3 P.S. Henry, "Coherent Lightwave Communications", TUH1, OFC '86, Atlanta, GA., Feb. 1986.

2.4 S. Yamazaki et al, "A 280 Mb/s Long-Span Optical FSK Heterodyne Transmission Experiment Utilizing a Flat FM Response Phase Tunable DFB Laser Diode", TUK2, OFC '86.

2.5 R.A. Linke et al, "Coherent Lightwave Experiments Using Amplitude and Phase Modulation at 400 Mb/s and 1 Gb/s Data Rates", WE1, OFC '86.

2.6 V.W.S. Chan, "Heterodyne Lasercom Space Demonstration", ThJ3, CLEO '86, San Francisco, CA., June, 1986.

2.7 D. Botez, "Laser Diodes are Power-Packed", IEEE Spectrum, June 1985.

2.8 J.C. Campbell et al, "High Performance Avalanche Photodiode with Separate Absorption 'Grading' and Multiplication Regions", Electronics Letters, Vol. 19, No. 20, 29 September 1983.

2.9 B.L. Kasper, "Sensitivity Limits on Direct Detection Receivers", WJ4, OFC '86.

2.10 R. Hickling et al, "1.5 Gbit/s GaAs Multiplexer and

Demultiplexer Forge Fast Fiber-Optic Links", Electronic Design, January 1986.

2.11 Honeywell HGG-2020.

2.12 J.E. Kaufmann and L.L. Jeromin, "Optical Heterodyne Intersatellite Links Using Semiconductor Lasers", 28.4, Globecom '84, Atlanta, GA., November 1984.

2.13 B.L. Kasper and C.A. Burrus, "Balanced Dual-Detector Receiver for Heterodyne Communication at Gb/s Speeds", WE2, OFC '86.

2.14 M.R. Mathews et al, "Packaged Frequency-Stable Tunable 20 KHz Linewidth 1.5 μm InGaAsP External Cavity Laser", Electronics Letters, Vol. 21, No. 3, 31 Jan. 1985.

2.15 E.J. Bachus et al, "Two Channel Heterodyne Type Fibre Optic Transmission Experiments", 9B3, ECOC '84, West Germany, September, 1984.

Section 3

3.1 R.G. Smith and S.D. Personick, "Semiconductor Devices for Optical Communications", p. 89, Springer-Verlag, N.Y., (1980).

3.2 A. Waksberg, "Optimization of a Photon Pulse Integrating Receiver Using an Avalanche Photodetector", Apl. Opt. 23, p. 3382, (1974).

3.3 P.P. Webb, R.J. McIntyre and J. Conradi, "Properties of Avalanche Photodiode", RCA Rev. 35, 234, (1974).

3.4 A. Waksberg, to be published.

Section 4

4.1 S. Haykin, Communication Systems, John Wiley & Sons, 1978.

4.2 R.G. Marshalek and G.A. Koepf, "Comparison of Optical Technologies for Intersatellite Links", Proc. of SPIE, 616, 29, (1986).

4.3 J.B. Abshire, "Performance of OOK and Low-Order PPM Modulations in Optical Communications When Using APD-Based Receivers", IEEEE Trans. on Comm., Vol. COM-32, No. 10, October 1984.

4.4 S. Kobayashi and T. Kimura, "Optical Phase Modulation in an Injection Locked AlGaAs Semiconductor Laser", IEEE JQE, Vol. QE-18, No. 10, October 1982.

4.5 T. Okoshi et al, "Computation of Bit-Error Rate of Various Heterodyne and Coherent-Type Optical Communication Schemes", J. Opt. Commun., 2, (1981) 3.

4.6 Y. Yamamoto, "Receiver Performance Evaluation of Various Digital Optical Modulation-Demodulation Systems in the 0.5 - 10 μm Wavelength Region", IEEE JQE, Vol. QE-16, No. 11, November 1980.

4.7 R.G. Smith and S.D. Personick, "Receiver Design for Optical Fiber Communication Systems", in Topics in Applied Physics; Volume 39; Semiconductor Devices for Optical Communications, Springer-Verlag, NY, 1980.

4.8 K. Kikuchi et al, "Degradation of Bit-Error Rate in Coherent Optical Communications Due to Spectral Spread of the Transmitter and the Local Oscillator", IEEE JLT, LT-2, 6, Dec. 1984.

Section 5

5.1 Spacecraft System Study - A Study to Define the Impact of Laser Communication Systems on Their Host Spacecraft.

NASA Report CR-175272.

5.2 P.W. Young, L.M. Germann, R. Nelson, "Pointing, Acquisition and Tracking Subsystem for Space-Based Laser Communications", SPIE Proceedings 616, 118, (1986).

5.3 G.A. Koepf and R. Peters, "Analysis of Burst Error Occurence on Optical Intersatellite Link (ISL) Design", SPIE Proceedings 616, 129, (1986).

5.4 J.D. Barry and G.S. Mecherle, "Communication Channel Burst Errors Induced by Gaussian Distributed Mispointing", SPIE Proceedings, 616, 137, (1986).

5.5 E. Sein, J.F. Clervoy, M. Lequine, B. Moreau and J.L. Hibon, "Acquisition and Fine-Pointing Control for a 400 Mbps Link Between a Low-Earth Orbiter and a and a Geostationary Satellite", SPIE Proceedings, 616, 141, (1986).

5.6 K.J. Held and J.D. Barry, "Precision Optical Pointing and Tracking from Spacecraft with Vibrational Noise", SPIE Proceedings, 616, 160, (1986).

5.7 P.W. Scott and P.W. Young, "Impact of Temporal Fluctuations of Signal to Noise Ratio (Burst Error) on Free-Space Laser Communications System Design", SPIE Proceedings, 616, 174, (1986).

5.8 J.D. Barry and G.S. Mecherle, "Beam Pointing Error as a Significant Design Parameter for Satellite Borne, Free-Space Optical Communication Systems", Optical Engineering, 24 (6), 1049, (1985).

5.9 V.A. Vilnrotter, "The Effects of Pointing Errors on the

Performance of Optical Communication Systems", The TDA Progress Report 42-63, 136, Jet Propulsion Laboratory, Pasadena, California, (1981).

5.10 J.D.Barry, G.S. Mecherle, R.A. Dye, A.J. Einhorn and T.A. Nussmeier, "Design of an Optical Communication Crosslink for Space Application Using GaAs/GaAlAs Laser Diodes", SPIE Conference Record. Military Applications of Lasers, 19-21, (1983).

5.11 J.D. Barry, A.J. Einhorn, R.A. Dye and T.A. Nussmeier, "Free Space Optical Communications Using GaAs/GaAlAs Laser Diodes", CLEO Conference Record, Baltimore, M.D., 152, (1983).

5.12 J.D. Barry, "Design and System Requirements Imposed by the Selection of GaAs/GaAlAs Single Mode Laser Diodes for Free Space Optical Communications", IEEE J. Quantum Elect., QE 20 (5), 478, (1984).

5.13 J.M. Lopez and K. Yong, "Acquisition, Tracking and Fine Pointing Control of Space-Based Laser Communication Systems", SPIE Proceedings, 295, 100, (1981).

5.14 L.G. Kazovsky, "Theory of Tracking Accuracy of Laser Systems", Optical Engineering, 22, 339, (1983).

Section 6

6.1 Y. Yamamoto and T. Kimura, "Coherent Optical Fiber Transmission Systems", IEEE JQE, Vol. QE-17, No. 6, June 1981.

Section 7

7.1 M. Nakamura et al, "Longitudinal-Mode Behaviours of Mode-Stabilized $Al_xGa_{1-x}As$ Injection Lasers",

J. Appl. Phys., 49 (9), September 1978.

7.2 S. Kobayashi et al, "Direct Frequency Modulation in AlGaAs Semiconductor Lasers", IEEE JQE, Vol. QE-18, No. 4, April 1982.

7.3 C.H. Henry, "Theory of the Linewidth of Semiconductor Lasers", IEEE JQE, Vol. QE-18, No. 2, Feb. 1982.

7.4 K. Kikuchi and T. Okoshi, "Estimation of Linewidth Enhancement Factor of AlGaAs Lasers by Correlation Measurement Between FM and AM Noises", IEEE JQE, Vol. QE-21, No. 6, June 1985.

7.5 K. Furuya, "Dependence of Linewidth Enhancement Factor on Waveguide Structure in Semiconductor Lasers", Electronics Letters, 21, (5), 1985.

7.6 K. Peterman, "Noise and Distortion Characteristics of Semiconductor Lasers in Optical Fiber Communication Systems", IEEE JQE, Vol. QE-18, April 1982.

7.7 O. Hirota and Y. Suematsu, "Noise Properties of Injection Lasers Due to Reflected Waves", IEEE JQE, Vol. QE-15, March 1979.

7.8 T. Kanada and K. Nawato, "Injection Laser Characteristics Due to Reflected Optical Power", IEEE JQE, Vol. QE-15, March 1979.

7.9 R. Long and K. Kobayashi, "External Optical Feedback Effects on Semiconductor Injection Laser Properties", IEEE JQE, Vol. QE-16, March 1980.

Section 9

9.1 G.H.B. Thompson, Physics of Semiconductor Laser Devices, John Wiley and Sons Ltd., (1980).

9.2 S. Machida and Y. Yamamoto, "Quantum-Limited Operation of Balanced Mixing Homodyne and Heterodyne Receivers", IEEE J. Quantum Elect., QE22(5), 617, (1986).

9.3 Y. Yamamoto, "AM and FM Quantum Noise In Semiconductor Lasers, Part II: Comparison of Theoretical and Experimental Results for AlGaAs Lasers", IEEE J. Quantum Elect., QE19(1), 47, (1983).

9.4 H.B. Kim, R. Maciejko, J. Conradi, "Effect of Laser Noise on Analogue Fibre Optic System", Electronics Lett. 16 (24), 919, (1980).

9.5 S. Saito and Y. Yamamoto, "Direct Observation of Lorentzian Line Shape of Semiconductor Laser and Linewidth Reduction with External Grating Feedback", Electronics Lett., 17 (9), 325, (1981).

9.6 S. Saito, Y. Yamamoto and T. Kimura, "Optical FSK Heterodyne Detection Experiments Using Semiconductor Laser Transmitter and Local Oscillator", IEEE J. Quantum Elect., QE17(6), 935, (1981).

9.7 J.E. Kaufmann and L.L. Jeromin, "Optical Heterodyne ISL Using Semiconductor Lasers", MIT, (1984).

9.8 Y.S. Lee and R.E. Eaves, "Implementation Issues of Intersatellite Links for Future Intelsat Requirements", Communications Satellite Corp., (1983).

9.9 J.D. Barry, "Design and System Requirements Imposed by the Selection of GaAs/GaAlAs Simple Mode Laser Diode for Free Space Optical Communications", IEEE, J. Quantum Elect., QE20, 478 (1984).

9.10 R.G. Marshalek, "Comparison of Optical Technologies

for ISL", Proc. of SPIE, 616, 29, (1986).

9.11 W.L. Casez, "Design of a Wideband Free-Space Laser Communication Transmitter", Proc. of SPIE, 616, 92, (1986).

9.12 V.W.S. Chan, "Coherent Optical Space Communications Systems Architecture and Technologies Issues", Proc. of SPIE, San Diego, (1981).

DUE DATE

~~SEP 2 1 1987~~			
SEP 2 3 1987			
MAR 2 2 1988			

201-6503

Printed in USA