DESIGN CONSIDERATIONS FOR AN OPTICAL PSK HOMODYNE RECEIVER.

A thesis submitted to

the Faculty of Engineering of the

University of Glasgow for the Degree of

Doctor of Philosophy

by

Michael Alistair Grant

October 1986

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Discovery consists of seeing what everybody has seen,
and thinking what nobody has thought.

Unknown.

TABLE OF CONTENTS

SUMMARY

This thesis presents studies concerned with the design of an optical homodyne PSK receiver.

Fundamental receiver sensitivities for a number of optical detection schemes are presented. The principal of performance degradation due to finite phase error in the detection process is established. The major noise source in semiconductor lasers is considered, lineshape anomalies, which may affect coherent communication systems, outlined, and methods for laser phase noise reduction discussed.

An experimental technique for the measurement of laser phase noise is implemented, and results from free running lasers presented. Techniques for linewidth reduction are investigated and results given.

An analysis of the optical Costas loop, an optimum solution of PSK detection, is presented. Performance in the presence of non-negligible phase noise is considered. Equivalent noise bandwidth integrals are evaluated for a number of loop filter functions, and a design procedure for a second order loop incorporating an active lead-lag filter derived. Consideration is then given to system operation in the presence of significant loop propagation delay. A model of the system is developed and results presented for beat linewidths typical of a number of common laser sources.

A number of components necessary for the implementation of the Costas loop are developed. These include a wideband detector/amplifier module, and a phase modulator. Consideration is also given to the design of a 90^{o} optical hybrid and an active filter network.

Experimentation using the semiconductor laser modules

developed in the project is reported. It was predicted from theory that the beat linewidths obtained would not, with the available system components, enable phase-lock to be acquired. It was not possible to confirm this experimentally due to a fire in the Department.

It was concluded that the reliable acquisition of phase-lock in an optical Costas loop will require further development of the laser sources. Acceptable loop performance will be achieved with an additional improvement of approximately one order of magnitude in system beat linewidths.

CHAPTER 1

INTRODUCTION

1.1. Historical Background.

The current generation of optical fibre communication systems employ direct intensity modulation of some optical source at the transmitter, either a light emitting diode (LED) or laser diode (LD), and power detection at the receiver, with an avalanche or PIN photodiode. This system is commonly known as intensity modulation/ direct detection, or IM/DD. This method of information transfer is analogous to the spark gap transmitters developed at the birth of radio communications around the turn of the century. The development of tuned circuits and associated theory coupled with the invention of the vacuum valve led rapidly to the use of frequency selective transmitters and receivers, and to the 'coherent' modulation and demodulation of amplitude, frequency, and phase modulated carriers. The use of these techniques resulted in an improvement of several orders of magnitude in receiver sensitivity and frequency selectivity, and, hence, more efficient use of the available radio frequency spectrum.

Similar advantages are to be gained from the use of coherent techniques in the field of optical fibre communications. It has been estimated that the useful spectral 'window' in currently available single mode silica fibre at 1.5μm is ~50,000GHz[1]. Coherent techniques would allow 1GBit channels to be spaced perhaps 2GHz apart in a frequency division multiplex system. In addition, receiver sensitivity improvements could allow an increase of up to four times in the distance between repeaters in a fibre optic link. This makes, for example, 'island hopping' of submarine cables feasible over a larger number of routes worldwide, reducing dramatically cable maintenance costs. Such

systems are still, however, several years in the future. Reliable operation of optical coherent receivers has yet to be proven in the laboratory, and the theoretical analysis of these receivers, and of certain artifacts unique to optical detection, has only recently advanced to a competent level. The work presented here seeks to contribute to this understanding.

1.2. The Development of Coherent Optical Systems.

With the invention in the 1950's of the laser, a narrow linewidth (temporally coherent) optical source, the development of coherent technology for optical communication systems became feasible. Atmospheric transmission experiments began in the 1960's with HeNe and CO_2 gas lasers [2,3,4,5,6]. However, these experiments were not extended to fibre optic systems until recently, when single mode CW semiconductor lasers became more readily available. The first successful demonstration of an optical heterodyne experiment applicable to fibre optic systems was reported in 1980[7,8]. Frequency shift keyed (FSK) modulation was employed, and AlGaAs semiconductor laser diodes were used as the transmitter and local oscillator. Also in 1980, Yamamoto[9] quantified for the first time the performance limits for coherent detection schemes in the optical regime.

The performance improvements over IM/DD systems predicted by Yamamoto have been the driving force behind the considerable body of research which has since been reported. This research effort can be divided into three broad areas:

1) Theoretical analysis of coherent detection systems, including the transmitters, fibre waveguide, and receivers.

2) Research into and development of special devices required for coherent detection systems, such as frequency stabilised and

spectrally pure lasers[10,11,12,13], polarization state controllers, and modulators.

3) Systems experiments, including bit error rate (BER) measurements on various modulation formats.

The effort in this Department was initiated in 1980. In 1983, R.C.Steele reported, for the first time, phase locking of semiconductor lasers[14]. This was done with two 825nm GaAlAs devices. A number of other papers around this time reported the results of BER measurements for various ASK, FSK, and PSK demodulation systems[15,16,17,18]. Sensitivity figures only 1dB away from the shot noise limit[15], and 19dB better than direct detection systems[18] were being achieved. Thus the theory derived by Yamamoto was confirmed experimentally. However, these measurements were done under ideal conditions, using either injection locked lasers or a self heterodyne process. No attempt had been made to phase lock two sources in a similar manner to Steele. There was, therefore, considerable scope for work on systems, and in particular phase locking, experiments. Most challenging was seen to be the case of the homodyne PSK receiver.

To improve on Steele's loop, development of the laser sources to improve stability and spectral purity was required. It was hoped that sufficient advances could be made to bring the sources within the scope of the loop theory available at the time[9,19]. This required that the effects of phase noise on the loop be negligible. Several new techniques for line narrowing were investigated. However, as outlined in chapter 4, the linewidth reduction and stability improvements achieved did not meet the theoretical criterion. A phase-locked loop model was therefore required which incorporated phase noise effects and optimised the loop parameters to take account of them.

In 1979, Armor and Robinson[20] reported on studies conducted into the coherent combination of CO_2 lasers for phase array applications. They derived an expression for the optimum equivalent loop bandwidth (minimum phase error variance) in this relatively noiseless situation. This was found to be a function of the heterodyne receiver shot noise, and various phase perturbations produced as a result of acoustic disturbance of experimental elements (mirrors, etc), and also of quantum phase fluctuations of the two laser sources. The application of this result to fibre optic coherent systems, which utilised large linewidth (noisy) semiconductor lasers, was not realised until almost six years later.

Recently, a considerable amount of work has been published based on Armor's model[21-27]. All of the workers base their analysis on PSK systems. These systems are most affected by phase noise. It was concluded that PSK homodyne detection requires a bit rate to linewidth ratio of $\sim 5 \times 10^{-4}$ for less than a 1dB performance penalty. The model presented in chapter 5 is a synthesis of these reports and work conducted in parallel in this project. Included for completeness is the evaluation of a number of equivalent noise bandwidth integrals not previously considered.

An assumption of a narrow beat linewidth (~ tens of kHz) was made throughout the above analyses. In this context, beat linewidth is the linewidth of the optical signal resulting from the mixing of two individual optical emissions, as occurs in a coherent optical receiver. The magnitude of the beat linewidth is the sum of the individual linewidths of the transmit and local oscillator lasers. At the present time, only one semiconductor laser source[11] has been reported as having a linewidth of the order of 10kHz. Therefore, beat linewidths of 1MHz or greater are

still likely to be encountered in experimental phase-locked coherent communication systems. Current theoretical models cannot fully describe the performance of phase-locked loops operating under these conditions. A modified loop model was therefore developed to fill this knowledge gap.

To maintain the variance of the loop phase error to within acceptable levels under the above conditions, a large value of loop bandwidth will be required. This in turn will mean that the effects of propagation delay in the loop will be significant. A theoretical model which evaluates the system phase error variance accounting for these effects is described in chapter 6. This analysis is relevant to all types of optical PLL's operating in this regime. However, only for the case of a Costas loop does chapter 6 represent a complete set of design criteria. In other cases, effects such as data to phase-lock crosstalk may have to be considered[22].

1.3. Outline of Thesis.

The work presented in this thesis falls into four major divisions. Chapters 2 and 3 present the necessary background material to this work. Chapter 2 contains the derivation of equations describing the theoretical minimum receiver sensitivities for a number of optical detection schemes operating under ideal conditions. This is known as shot noise limited detection, and yields the underlying reason for pursuing this field of research. In addition, the principle of performance degradation resulting from phase noise in the detection process is established.

Chapter 3 examines the major sources of spectral broadening in semiconductor lasers. Initially, the free running laser is

studied, and an expression for the fundamental 3dB linewidth, or full-width half-maximum (FWHM), derived. Anomalies in the expected spectral lineshape of the emission are discussed. Techniques for the reduction of the 3dB linewidth are also considered in this chapter. These include several passive techniques as well as discussion of innovative laser structures.

Experimental measurements of laser linewidths are made in chapter 4 using a technique first described by Okoshi et al[28]. This self heterodyne experiment allowed various line narrowing techniques to be evaluated. Results from two types of line narrowing systems are reported, and the correlation of experiment with theory examined.

The development of a loop model which will satisfactorily describe the operation of a Costas loop with the beat linewidths obtained from the sources reported in chapter 4 is described in chapters 5 and 6. Chapter 5 contains an analytical model which describes the effects of phase noise (non-negligible linewidth) on loop performance. However, this analysis falls down when large values of beat linewidth (>~500kHz) are considered, and large values of loop bandwidth are required. The effects of propagation delay in the loop will then bring about a degradation in performance. A model which incorporates propagation delay and which quantifies these effects is described in Chapter 6.

Chapter 7 details experimental steps taken to forward the implementation of an optical Costas loop. Several loop components were designed and constructed. Measurements were made of beat spectra between two of the modules described in chapter 4. Using the model of chapter 6 and result of loop propagation delay measurements, conclusions were drawn relating to the likelihood of success of this experiment.

Conclusions and recommendations for further work are outlined in chapter 8.

References.

1. D.W.Smith, T.G.Hodgkinson, D.J.Malyon, P.Healy 'Demonstration of a Tunable Heterodyne Receiver in a Two Channel Optical FDM Experiment' IEE Colloq. no.1985/30, Savoy Place, London, March 1985.

2. F.E.Goodwin 'A 3.39μm Infrared Optical Heterodyne Communication System' IEEE J. Quant. Elect. vol QE-3(11) 1967. pp524-531.

3. T.A.Nussmeier, F.E.Goodwin, J.E.Zavin 'A 10.6μm Terrestial Communication Link' IEEE J. Quant. Elect. vol QE-10(2), 1974. pp230-235

4. W.R.Leeb, H.K.Philipp, A.L.Scholtz, E.Bonek 'Frequency Synchronization and Phase Locking of CO_2 Lasers' Appl. Phys. Lett. 41(7) 1982. pp592-594.

5. H.K.Philipp, A.L.Scholtz, E.Bonek, W.R.Leeb 'Costas Loop Experiment for a 10.6μm Communications Receiver' IEEE Trans. Commun. vol COM-31 no.8 Aug. 1983. pp1000-1002.

6. H.K.Philipp, A.L.Scholtz, W.R.Leeb 'Homodyning of Subnanowatt 140MBit/s Data at λ = 10.6μm' ECOC'84, no. 9B6, Stuttgart, Sept. 1984.

7. S.Saito, Y.Yamamoto, T.Kimura 'Optical heterodyne Detection of Directly Frequency Modulated Semiconductor Laser Signals' Elect. Lett. vol 16(22) 1980. pp826-827.

8. S.Saito, Y.Yamamoto, T.Kimura 'Optical FSK Heterodyne Detection Experiments using Semiconductor Laser Transmitter and Local Oscillator' IEEE j. Quant. Elect. vol QE-17(6) 1971. pp935-941.

9. Y.Yamamoto 'Receiver Performance Evaluation of Various Digital Optical Modulation-Demodulation Systems in the 0.5-10μm Region' IEEE J. Quant. Elect. vol QE-16(11) 1980. pp1251-1259.

10. F.Favre, D.LeGuen 'Effect of Semiconductor Laser Phase Noise on BER Performance in an Optical DPSK Heterodyne-Type Experiment' Elect. Lett. vol 18(22) 1982. pp964-965.

11. R.Wyatt, W.J.Devlin '10kHz Linewidth 1.5μm InGaAsP External Cavity Laser with 55nm Tuning Range' Elect. Lett. 19() 1983. pp110-112.

12. T.P.Lee, C.A.Burrus, D.P.Wilt 'Measured Spectral Linewidth of Variable Gap Cleaved Coupled Cavity Lasers' Elect. Lett. vol 21(2) 1985. pp53-54.

13. T.Fujita et al 'Narrow Spectral Linewidth Characteristics of Monolithic Integrated Passive Cavity InGaAsP/InP Semiconductor Lasers' Elect. Lett. vol 21(9) 1985. pp374-376.

14. R.C.Steele 'Optical Phase-Locked Loop using Semiconductor Laser Diodes' Elect. Lett. 19(2) 1983. pp69-71.

15. K.Kikuchi, T.Okoshi, J.Kitano 'Measurement of Bit-Error-Rate of Heterodyne Type Optical Communication System - A Simulation Experiment' IEEE J. Quant. Elect. vol QE-17(2), 1981. pp2266-2267.

16. T.G.Hodgkinson, R.Wyatt, D.W.Smith 'Experimental Assessment of a 140MBit/s Coherent Optical Receiver at 1.52μm' Elect. Lett. vol 18(12) 1982. pp523-525.

17. S.Saito, Y.Yamamoto, T.Kimura 'S/N and Error Rate Evaluation for an Optical FSK-Heterodyne Detection System using Semiconductor Lasers' IEEE J. Quant. Elect. vol QE-19(2) 1983. pp180-193.

18. D.J.Malyon et al 'PSK Homodyne Receiver Sensitivity Measurements at 1.5μm' Elect. Lett. vol 19 1983. pp144-146.

19. T.Okoshi et al 'Computation of Bit-Error-Rate of Various Heterodyne and Coherent Type Optical Communication Schemes' J. Opt. Commun. vol 2(3) 1981. pp89-96.

20. J.B.Armor, S.R.Robinson 'Phase-Lock Control Considerations for Coherently Combined Lasers' Appl. Opt. vol 18(18) 1979. pp3165-3175.

21. L.G.Kazovsky 'Decision-Driven Phase-Locked Loop for Optical Homodyne Receivers: Performance Analysis and Laser Linewidth Requirements' J. Lightwave Technol. vol LT-3(6) 1985. pp1238-1247.

22. L.G.Kazovsky 'Balanced Phase-Locked Loops for Optical Homodyne Receivers: Performance Analysis, Design Considerations, and Laser Linewidth Requirements' J. Lightwave. Technol. vol LT-4(2) 1986. pp182-194.

23. L.G.Kazovsky 'Performance Analysis and Laser Linewidth Requirements for Optical PSK Heterodyne Communication Systems' IEEE J. Lightwave. Technol. vol LT-4(4) 1986. pp415-425.

24. T.G.Hodgkinson 'Phase-Locked Loop Analysis for Pilot Carrier Coherent Optical Receivers' Elect. Lett. vol 21 1985. pp1202-1203.

25. T.G.Hodgkinson 'Costas Loop Analysis for Optical Receivers' Elect. Lett. vol 22(7) 1986. pp394-396.

26. B.Glance 'Performance of Homodyne Detection of Binary PSK Optical Signals' J. Lightwave Technol. vol LT-4(2) 1986. pp228-235.

27. B.Glance 'Minimum Required Power for Carrier Recovery at Optical Frequencies' IEEE J. Lightwave Technol. vol LT-4(3) 1986. pp249-255.

28. T.Okoshi, K.Kikuchi, A.Nakayama 'Novel Method for high Resolution Measurement of Laser Output Spectrum' Elect. Lett. 16(16) 1980. pp630-631.

CHAPTER 2.

A PERFORMANCE COMPARISON OF OPTICAL DETECTION SCHEMES.

2.1. Introduction.

In any research project it is necessary to derive some datum level against which experimental or theoretical results can be measured. This chapter will perform such a function by defining a maximum achievable performance for a number of binary optical detection schemes. Theoretical limits will be derived and a comparison of the results made. Consideration will also be given to the practical realisation of the calculated performance. The detection systems discussed will be[1,2,3,4,5]:

a) Intensity Modulated Direct Detection (IM/DD) systems,

b) Amplitude Shift Keyed (ASK) coherent detection systems,

c) Frequency Shift Keyed (FSK) coherent detection systems, and

d) Phase Shift Keyed (PSK) coherent detection systems.

The effect of laser phase instabilities on coherent system performance[6,7,8,9,10] will also be discussed and conclusions reached as to the modulation format most likely to give the best performance.

2.2. Noise in Optical Receivers.

The basic components of any optical receiver are shown in figure 2.1. They are the photodetector (an avalanche photodiode (APD), or PIN diode), a front-end amplifier and equalisation circuitry, and a decision circuit to recover the transmitted data. This essentially completes an IM/DD receiver. A coherent receiver will require, in addition, some form of feedback path (dashed lines in figure 2.1) to control the frequency or phase of the

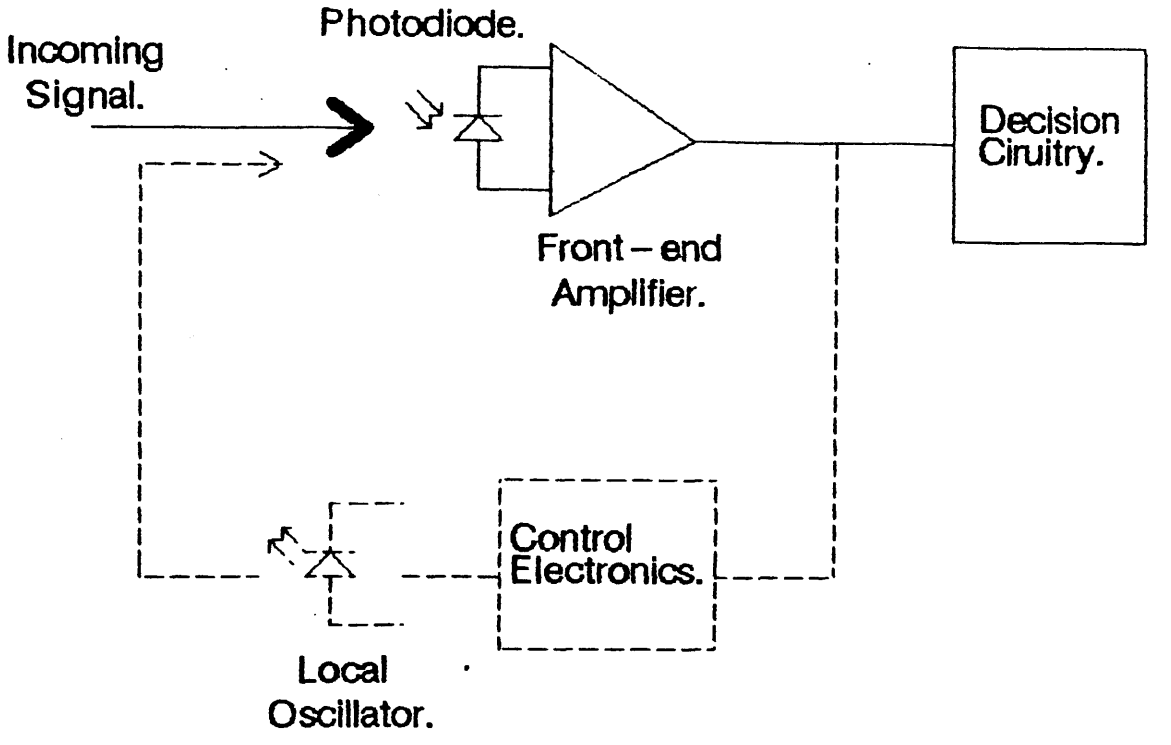


Figure 2.1. A Generalised Optical Receiver.

local oscillator, the reference signal against which the received optical signal is compared.

2.2.1. Shot Noise.

All optical receivers will be subject to shot noise due to the quantum nature of light. Noise sources of this type will include the received signal power, any received background emissions, the dark current internal to the diode, and, in the case of coherent receivers, shot noise due to the incident local oscillator power. Each of these sources will produce some equivalent mean square photocurrent at the output of the diode:-

$$\overline{i^2}_s = 2.e.(\eta.e/h.v).P_s.\langle M\rangle^{2+X}.B \qquad \text{----[2.1]}$$

$$\overline{i^2}_b = 2.e.(\eta.e/h.v).P_b.\langle M\rangle^{2+X}.B \qquad \text{----[2.2]}$$

$$\overline{i^2}_d = 2.e.(I_m\langle M\rangle^{2+X} + I_n).B \qquad \text{----[2.3]}$$

and $$\overline{i^2}_{LO} = 2.e.(\eta.e/h.v).P_{LO}.\langle M\rangle^{2+X}.B \qquad \text{----[2.4]}$$

where s denotes signal power, b, background light, d, dark current, and LO, local oscillator power and where

B is the bandwidth of the diode/front-end amplifier combination,

η is the quantum efficiency of the diode,

h is Planks constant,

e is the charge on an electron,

v is the frequency of the incident radiation,

M is the gain of the APD (= 1 for PIN),

X is the excess noise factor of the APD,

P is the received power of the appropriate source,

I_m is multiplied dark current,

and I_n is non multiplied dark current.

Dark current results from the spontaneous emission of hole-electron pairs in the device which travel to the device terminals constituting a current under no light conditions. If the device is an APD then, depending on where the emissions occur, they may or may not undergo a multiplication process. In a PIN diode, $I_m = 0$.

It should be noted that no attempt is made in this analysis to account for the effect of pulse shape, intersymbol interference, or any other effect associated with a non-ideal IM/DD receiver. The reader is referred to references 11 and 12 for a fuller treatment of these subjects.

2.2.2. Thermal Noise.

Associated with the front-end amplifier and the equivalent load resistance will be a significant noise contribution due to thermal or Johnson noise. If a simple load resistor is used to bias the diode, then the equivalent noise power will be:-

$$\overline{i^2}_r = 4.k.T.B/R_L \qquad \text{----[2.5]}$$

where k is Boltzmanns constant, R_L is the load resistance, and T is the temperature of the device. If R_L is chosen so as not to be the dominant thermal noise source then the noise contribution from the front-end amplifier must be considered. A likely choice of amplifier is a high impedance FET. The equivalent noise power of this device is given by the expression[1]:-

$$\overline{i^2}_c = [(4.k.T/R_L) \cdot \{1+\Gamma/g_m.R_L\} + 2.e.I_{GATE}] + 4.k.T.\Gamma.(2.\pi.C_T)^2.B^3 / g_m \quad \text{----[2.6]}$$

where, in addition to those terms already defined:-

g_m = FET transconductance,

Γ = a numerical factor given by the FET material,

I_{GATE} = gate leakage current of the FET,

and C_T = total load capacitance of the amplifier.

Using equations [2.1] to [2.6] the theoretical performance of the optical receivers under consideration can now be derived.

2.3. Theoretical Performance of an IM/DD System.

2.3.1. Probability of Error.

As stated in section 2.2.1, the solid line components of figure 2.1 constitute a basic direct detection receiver. The transmission format in such a system is 'power on' denoting one state and 'power off' denoting the second. As such, the receiver is subject to two noise conditions. If 'power on' is defined as the mark or '1' state, and 'power off' a space or '0' state, then when a mark is received the total variance of the input noise current is the sum of the variances of the individual noise sources[13]:-

$$\sigma^2_{D1} = \overline{i^2}_s + \overline{i^2}_d + \overline{i^2}_b + \overline{i^2}_c \quad \text{----[2.7]}$$

where D denotes direct detection. Similarly, for a received space:-

$$\sigma^2_{D0} = \overline{i^2}_d + \overline{i^2}_b + \overline{i^2}_c \quad \text{----[2.8]}$$

It is implicitly assumed above that the noise processes

described in [2.2] are governed by gaussian statistics. This is not strictly accurate since photodetection is a discrete process governed by Poisson statistics. However, the gaussian analysis is a much simpler approach, and, if the number of observed events is large, the associated errors are very small.

If the corresponding signal current for a received mark is given by (assuming $\langle M \rangle = 1$ for simplicity):-

$$S_D = \eta.e.P_s \ / \ h.v \qquad \text{----[2.9]}$$

then the probability density functions for the reception of the two states can be calculated[13]:-

$$P_0(x) = \frac{1}{(2.\pi.\sigma^2_{D0})^{0.5}} \exp -(x^2/2.\sigma^2_{D0}) \qquad \text{----[2.10]}$$

and

$$P_1(x) = \frac{1}{(2.\pi.\sigma^2_{D1})^{0.5}} \exp -((S_D-x)^2/2.\sigma^2_{D1}) \qquad \text{----[2.11]}$$

These functions are plotted in figure 2.2. D is the decision level derived from the overlap of the two functions.

If $P_{e0}(x)$ is the probability that a mark is received when a space is transmitted and $P_{e1}(x)$ is the corresponding mark error, then the total probability of error is given by:-

$$P_D = P_{e0}(x).P_0 + P_{e1}(x).P_1 \qquad \text{----[2.12]}$$

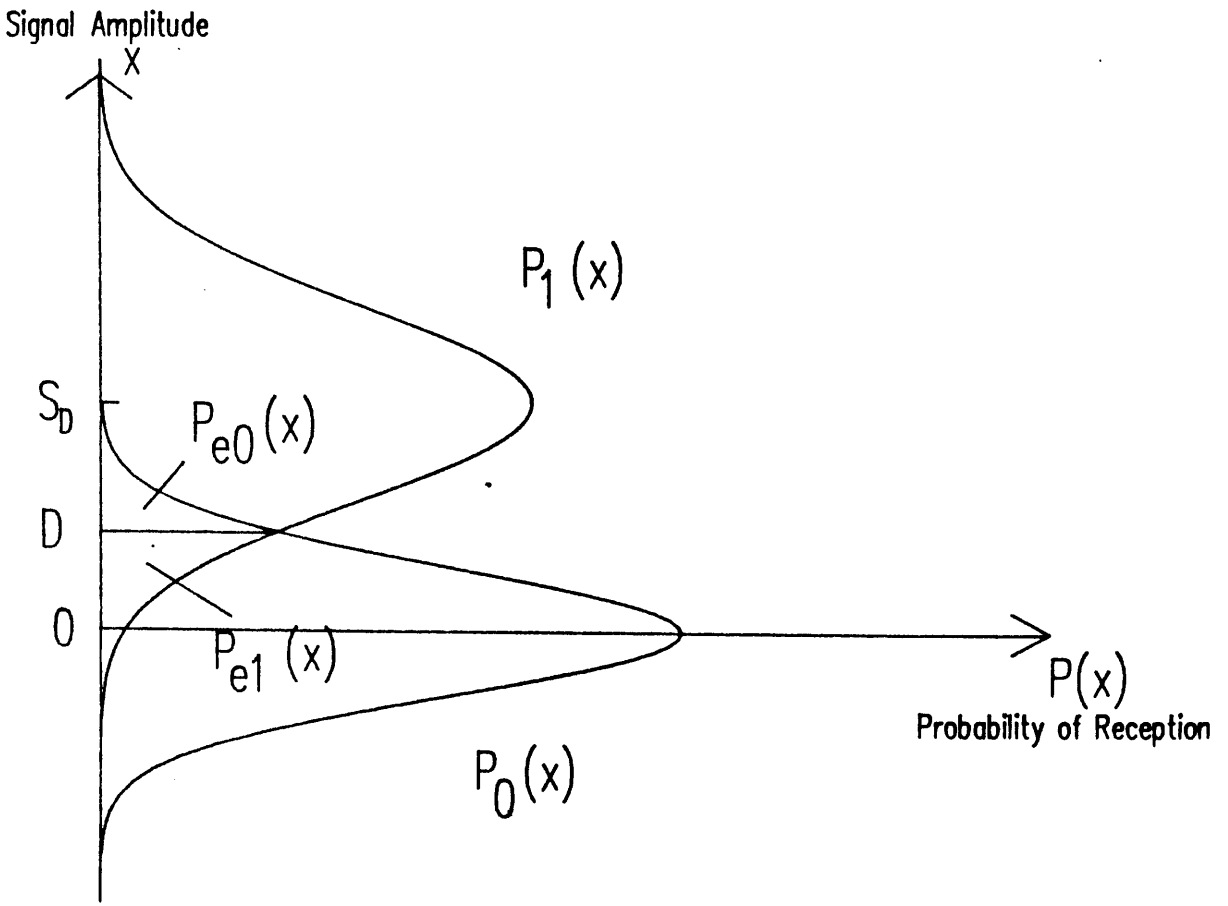


Figure 2.2.
Probability Density Functions
of Direct Detection.

where P_0 and P_1 are the probabilities that a space and mark respectively were initially transmitted. For simplicity, it is assumed that on average an equal number of ones and zeros were sent. Thus, $P_0 = P_1 = 1/2$. From figure 2.2 it can be seen that $P_{e0}(x)$ is the area of the PDF, $P_0(x)$, which falls between D and ∞:-

$$P_{e0}(x) = \frac{1}{(2.\pi.\sigma^2_{D0})^{0.5}} \int_D^{\infty} \exp -(x^2/2.\sigma^2_{D0})\, dx \qquad \text{----[2.13]}$$

Similarly, $P_{e1}(x)$ is the area of $P_1(x)$ between $-\infty$ and D:-

$$P_{e1}(x) = \frac{1}{(2.\pi.\sigma^2_{D1})^{0.5}} \int_{-\infty}^{D} \exp -((S_D-x)^2/2.\sigma^2_{D1})\, dx \qquad \text{----[2.14]}$$

Substitution of [2.13] and [2.14] into [2.12] yields the total probability of error or bit error rate (BER) of the system:-

$$P_D = \frac{1}{2}\left[\frac{1}{(2.\pi.\sigma^2_{D0})^{0.5}} \int_D^{\infty} \exp -(x^2/2.\sigma^2_{D1})\, dx \right.$$

$$\left. + \frac{1}{(2.\pi.\sigma^2_{D1})^{0.5}} \int_{-\infty}^{D} \exp -((S_D-x)^2/2.\sigma^2_{D1})\, dx \right] \qquad \text{----[2.15]}$$

Introducing the change of variable, $t=(S_D-x)/\sigma_{D1}$ in $P_{e1}(x)$ and $t=x/\sigma_{D0}$ in $P_{e0}(x)$ results in:-

$$P_D = \frac{1}{2}\left[\frac{1}{(2.\pi)^{0.5}}\int_{D/\sigma_{D0}}^{\infty} \exp -t^2/2 \, dt\right.$$

$$\left. + \frac{1}{(2.\pi)^{0.5}}\int_{(S_D-D)/\sigma_{D1}}^{\infty} \exp -t^2/2 \, dt\right] \qquad \text{----[2.16]}$$

Assuming an equal probability of erroneous mark or space, then the limits in both integrals can be equated to give an optimum decision level:-

$$D/\sigma_{D0} = (S_D - D) \; / \; \sigma_{D1}$$

$$\Rightarrow \qquad D_{opt} = \frac{\sigma_{D0} \, S_D}{\sigma_{D0} + \sigma_{D1}} \qquad \text{----[2.17]}$$

and hence, substituting [2.17] into [2.16] yields a total probability of error, or BER, of :-

$$BER = \frac{1}{(2.\pi)^{0.5}}\int_{\frac{S_D}{(\sigma_{D0}+\sigma_{D1})}}^{\infty} \exp -t^2/2 \; dt \qquad \text{----[2.18]}$$

The form of the integral in [2.18] is well defined[13] and is known as the complementary error function of Q (erfc(Q)), where Q is the finite lower limit of the integral ($Q = S_D/(\sigma_{D0}+\sigma_{D1})$ in this case).

It should be recognised that $(S_D/(\sigma_{D0}+\sigma_{D1}))^2$ is the detector input signal to noise ratio (SNR). Thus, for a given input SNR the system bit error rate is easily calculated:-

$$BER = erfc[(S/N)^{0.5}] \qquad \text{----[2.19]}$$

This function is plotted in figure 2.3. From this graph it can be seen that the required received signal to noise ratio (SNR) for a BER of 10^{-9} is 15.56dB (Q = 6).

2.3.2. The IM/DD Shot Noise Limit.

Consider the case of a noiseless receiver, ie $\sigma_{D0} = 0$. Under those circumstances an expression for the theoretical minimum received power per bit for a prescribed error rate is obtained:-

$$Q = S_D / \sigma_{D1} \qquad \text{----[2.20]}$$

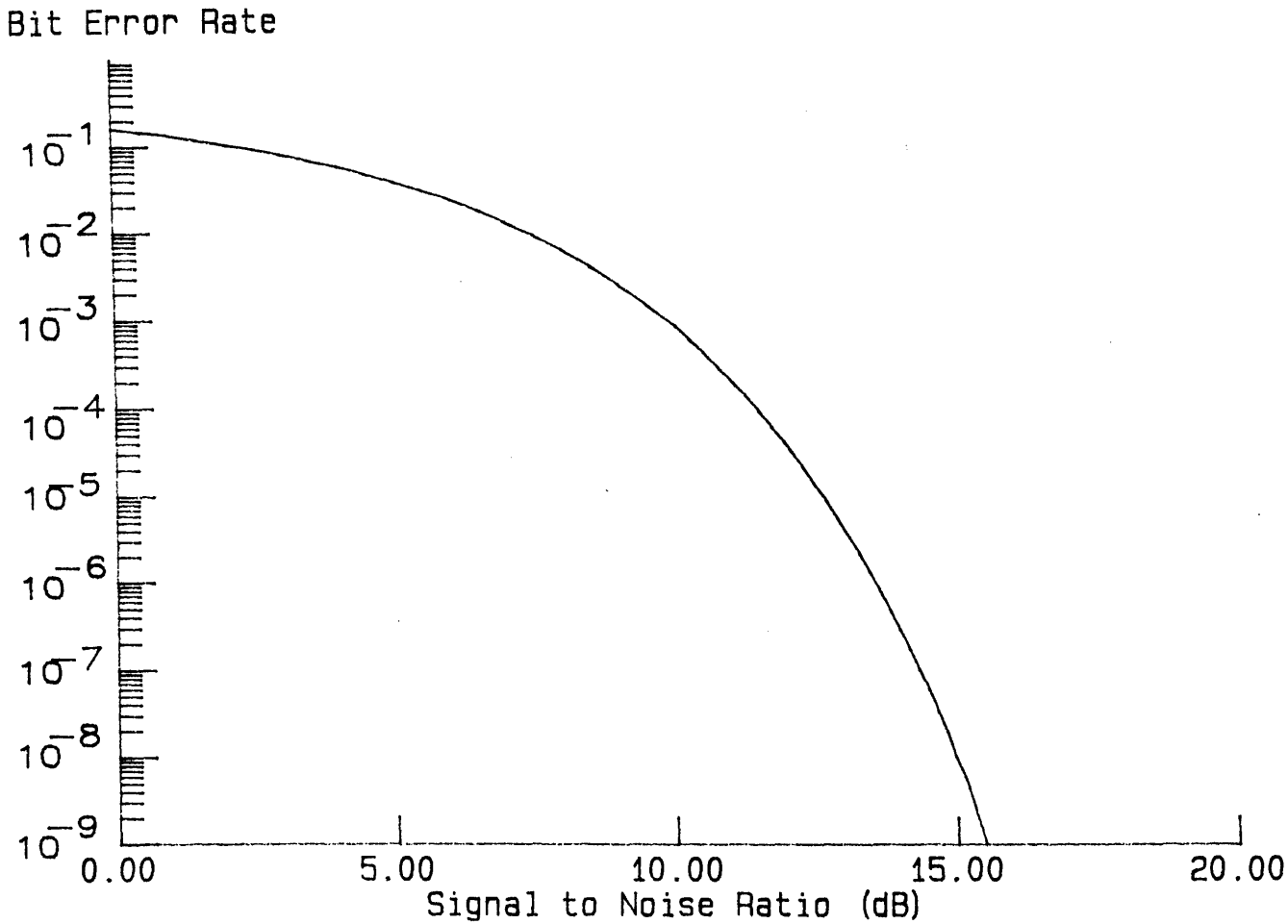
Substituting [2.9] and [2.7] (and hence [2.1]) into [2.21] yields:-

$$Q = \frac{\eta.e.P_s / h.v}{(e^2.B.\eta.P_s / h.v)^{0.5}}$$

$$\Rightarrow \quad Q^2 = \eta.P_s / (h.v.B) \qquad \text{----[2.21]}$$

The term $\eta.P_s / h.v$ is the number of photons incident on the diode per second. In an ideal receiver, the input bandwidth, B, can be minimised to pass power in frequency components only up to the bit rate, f_B. Therefore, the minimum required number of photons per bit can be found. This figure is 36 photons per bit for a 10^{-9} BER. If a system bit rate and receiver responsivity ($R=\eta.e/(h.v)$) is then defined, this number will correspond to an average received power. A typical system responsivity for a silicon PIN diode operating at 830nm is R=0.85. Also a bit rate of 565Mbit/s will be specified. This results in a minimum received

BER versus Signal to Noise Ratio
Figure 2.3



power of -54.2dBm required to achieve the specified performance in this system.

2.3.3. Discussion on Direct Detection Systems.

It can be seen that the sensitivity of a direct detection system depends on a number of parameters. In a optical fibre system, background light is normally insignificant, and depending on the wavelength of operation and hence the type of detector used, the detector dark current can usually be ignored as well. This leaves the thermal noise component of the receiver as the dominant noise source. Combining this with the fact that the quantum efficiency of the detector is always less than 1 (0.8 - 0.9 for silicon devices, and 0.6 - 0.8 for germanium), results in a system that can be some 10-15dB less sensitive than the shot noise limit.

2.4. Theoretical Performance of Coherent Detection Systems.

2.4.1. General Considerations.

Returning to figure 2.1, the broken line portion of the illustration represents the additional components required for the general implementation of a coherent detector. These are;

a) some form of feedback control electronics providing frequency or phase information to the local oscillator from the beat signal impinging on the detector,

b) the local oscillator, which provides a reference signal with which the receiver compares the incoming signal, and,

c) some form of mixer to combine the local oscillator signal with the transmit, or information bearing, signal.

The output of the mixer can be represented as the sum of the complex amplitude of the two input signals.

Let:-

$$E_T(t) = Re\{(2P_T)^{0.5}.\ exp(i(\omega_T t + \emptyset_T\}$$

$$= (2P_T)^{0.5}\ cos(\omega_T t + \emptyset_T) \qquad \text{----[2.22]}$$

be the magnitude of the received optical signal and:-

$$E_{LO}(t) = Re\{(2P_{LO})^{0.5}.\ exp(i(\omega_{LO} t + \emptyset_{LO})\ \}$$

$$= (2P_{LO})^{0.5}\ cos(\omega_{LO} t + \emptyset_{LO}) \qquad \text{----[2.23]}$$

the magnitude of the local oscillator signal. ω is the mean angular frequency and $\emptyset$ the absolute phase of the respective waveforms. At this point, no consideration of any phase instabilities in these signals is being made, i.e. the source linewidths are assumed to be zero. Since the photodiode (PIN or APD) in the receiver is a power detector the photocurrent resulting from the mixing of E_T and E_{LO} is of the form:-

$$i_s = R.[\ E_T(t) + E_{LO}(t)\ ]^2$$

$$= R.[\ P_T + P_{LO} + 4(P_T.P_{LO})^{0.5}\ cos(\omega_T t + \emptyset_T)cos(\omega_{LO} t + \emptyset_{LO})]$$

$$\text{----[2.24]}$$

where R is the responsivity of the diode. Expanding cosines and ignoring the 2ω terms, since no optical response is possible at those frequencies, yields:-

$$i_s(t) = R.[P_T + P_{LO} + 2(P_T.P_{LO})^{0.5} \cos((\omega_T-\omega_{LO})t + (\phi_T-\phi_{LO}))] \quad \text{----[2.25]}$$

The term $(\omega_T-\omega_{LO})$ is the difference or intermediate frequency term of the beat signal denoted ω_{IF}. $\phi_T - \phi_{LO}$, denoted ϕ_E, represents the phase error between the reference signal and the input. In a frequency locked loop, only ω_{IF} is used to provide feedback information. In a phase locked loop ϕ_E is utilized to provide a control signal. If stable lock is achieved (in frequency or phase) with $\omega_{IF} \neq 0$, then the coherent detection process is known as heterodyne detection. If $\omega_{IF} = 0$, then the detection is a homodyne process. The following analysis will assume heterodyne detection. The performance variances of homodyne detection will be considered in section 2.4.5..

2.4.2. ASK (On-Off-Keyed) Coherent Detection.

As indicated below in equation [2.26], there is an additional noise source in coherent detection systems not present in direct detection systems. This is the quantum shot noise introduced by the local oscillator. From equation [2.4]:-

$$\overline{i^2}_{LO} = 2.e.R.P_{LO}.B_H \quad \text{----[2.4]}$$

where B_H is the IF bandwidth of the coherent heterodyne receiver. It should again be noted that, at this point, laser phase noise is excluded from the discussion. Incorporating this in [2.7] and [2.8] the total variance of the input noise current for on-off-keyed (OOK) systems is obtained:-

$$\sigma^2_{H1} = \overline{i^2}_T + \overline{i^2}_{LO} + \overline{i^2}_d + \overline{i^2}_b + \overline{i^2}_c \quad \text{----[2.26]}$$

and:-

$$\sigma^2_{H0} = \overline{i^2}_{LO} + \overline{i^2}_d + \overline{i^2}_b + \overline{i^2}_c \qquad \text{----[2.27]}$$

However, due to the proximity of the local oscillator to the receiver, $P_{LO} >> P_T$. Moreover, P_{LO} can be made large enough to dominate all other noise terms. Thus, since $\overline{i^2}_{LO}$ is present in both [2.26] and [2.27], the noise variance for mark and space detections are the same:-

$$\sigma^2_{H1} = \sigma^2_{H0} = \sigma^2_H = \overline{i^2}_{LO} \qquad \text{----[2.28]}$$

The signal amplitude in an ASK scheme, S_H, is given by the magnitude of the AC component of [2.25]:-

$$S_H = 2.R(P_T.P_{LO})^{0.5} \qquad \text{----[2.29]}$$

Thus it can be seen by comparing [2.28] with [2.29] that P_{LO} appears in both the signal and noise terms. This obviously has implications for the signal to noise ratio of the system in as much as by increasing P_{LO}, $\overline{i^2}_{LO}$ becomes the dominant noise source, as was stated, while at the same time the signal power, S_H, is also increased. Such a situation realises the maximum input SNR for the receiver, and is known as shot noise limited detection. This state can be reached for any modulation format employing a coherent receiver and is the reason why coherent detection systems may achieve quantum limited detection when direct detection systems do not.

From [2.28] and [2.29] the PDF's for zero and one transmissions can be calculated (Figure 2.4). Since $\sigma^2_{H1} = \sigma^2_{H0}$,

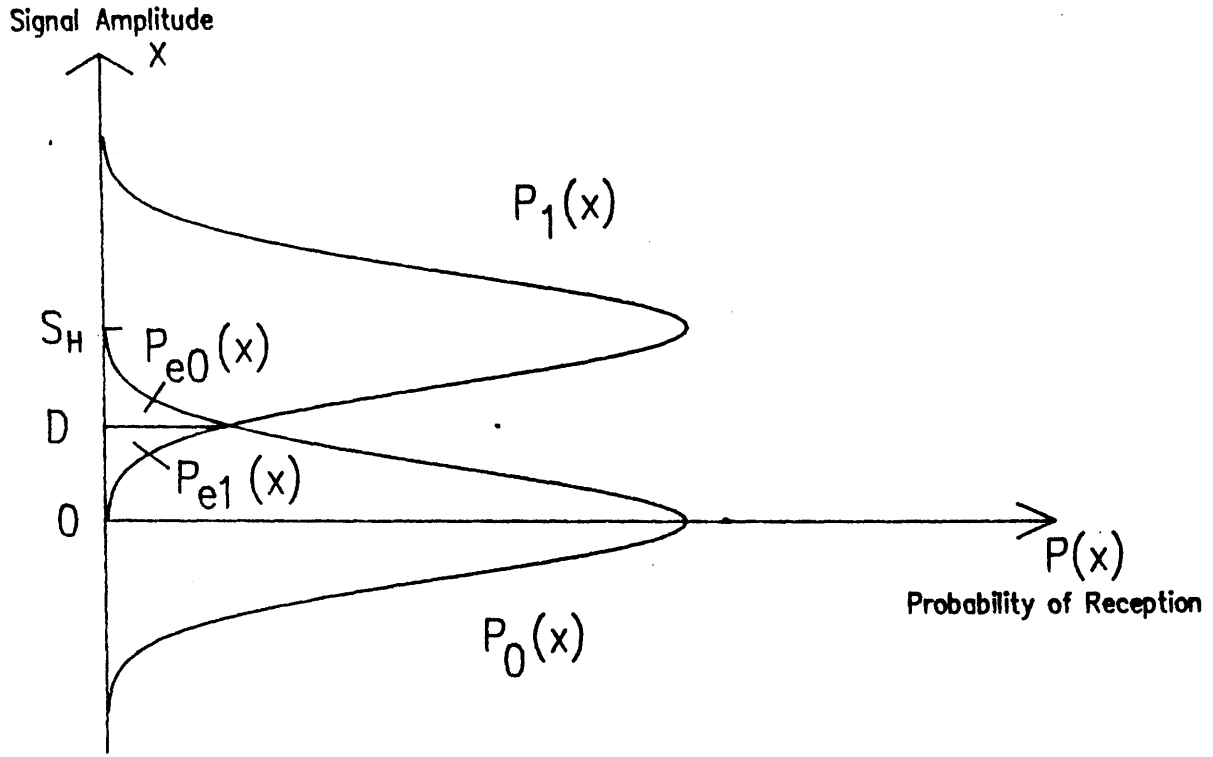


Figure 2.4.
Probability Density Functions
of ASK Coherent Detection.

the optimum detection level will be half the mark signal power level:-

$$D_{opt} = S_H/2 \qquad \text{----[2.30]}$$

Using an identical argument to that described in section 2.3.1., an expression for the total probability of error can be derived. Utilising [2.28], [2.29], and [2.30] in [2.10] through to [2.18]:-

$$BER_{ASK} = \frac{1}{2}\left[\text{erfc}\left[\frac{S_H - D_{opt}}{\sigma_H}\right] + \text{erfc}\left[\frac{D_{opt}}{\sigma_H}\right]\right]$$

$$= \text{erfc}\left(\frac{S_H}{2\sigma_H}\right) \qquad \text{----[2.31]}$$

The minimum receiver sensitivity of a coherent ASK system is obtained by substituting [2.28] and [2.29] into [2.31] and equating to Q, as in section 2.3.2.:-

$$Q^2 = \frac{R.P_T}{2.e.B_H} \qquad \text{----[2.32]}$$

The IF bandwidth of an ideal heterodyne receiver will be twice the system bit rate. Thus, if the same system parameters are used (R = 0.85, f_B = 565MBit/s), the minimum received power/bit for a 10^{-9} BER is -51.2dBm.

Depending on the wavelength of operation, bit rate of the system, and quality of the direct detection system, the received optical power for an ASK heterodyne direct detection system can be in practice some 7-12dB less than that required in a direct

detection system for the same performance. This is due purely to the ability of the coherent receiver to attain shot noise limited detection.

2.4.3. FSK Coherent Detection.

FSK, unlike ASK, is a constant carrier mode, ie there is a carrier transmitted with both mark and space bits. Information is transmitted by keying between two frequencies, f_1 and f_2 for mark and space respectively, and detection is implemented through the use of two receivers tuned to f_1 and f_2, producing in turn output voltages, v_1 and v_2 (Figure 2.5). Thus, by summing these two outputs, a decision can be made on the received bit. The total probability of error thus requires consideration of the outputs of both receivers in one bit period.

Considering the reception of a mark bit and hence the presence of a signal power, S_H, given by [2.29], at the output of receiver 1, then the PDF is[1]:-

$$P(v_1) = \frac{1}{(2.\pi.\sigma^2_{v1})^{0.5}} \exp -((S_H-v_1)^2/2.\sigma^2_{v1}) \qquad ----[2.33]$$

while the PDF of receiver 2 is:-

$$P(v_2) = \frac{1}{(2.\pi.\sigma^2_{v2})^{0.5}} \exp -((S_H-v_2)^2/2.\sigma^2_{v2}) \qquad ----[2.34]$$

These signals are then input to the summing device. An error will therefore occur if $v_2 > v_1$. The probability of error when a mark is transmitted is then:-

$$P_e(v_1) = \text{Prob}(v_1 - v_2 < 0)$$

$$= \frac{1}{(2.\pi.(\sigma^2_{v1}+\sigma^2_{v2}))^{0.5}} \int_{-\infty}^{0} \exp -((w-S_H)^2/2.(\sigma^2_{v1}+\sigma^2_{v2}) \, dw$$

----[2.35]

Since, as described in 2.4.2., the local oscillator signal will be the dominant noise term at the input to both receivers, the noise variance for a received mark will equal that for a received space, ie $\sigma^2_{v1} = \sigma^2_{v2} = \sigma^2_H$. Thus [2.35] becomes:-

$$P_{e1}(w) = \frac{1}{2.(\pi)^{0.5}.\sigma_H} \int_{-\infty}^{0} \exp -(w-S_H)^2/4.\sigma^2_H \, dw \quad ----[2.36]$$

Using the change of variable $u = (w-S_H)/2^{0.5}\sigma_H$:-

$$P_{e1}(w) = \frac{1}{(2\pi)^{0.5}} \int_{\frac{S_H}{2^{0.5}\sigma_H}}^{\infty} \exp -(u^2/2) \, du$$

$$= \text{erfc}\ (S_H/2^{0.5}\sigma_H) \quad ----[2.37]$$

An identical argument yields the same result for $P_{e0}(w)$. Therefore, by using [2.12] and assuming that $P_0 = P_1 = 1/2$, the total probability of error can be obtained:-

$$BER_{FSK} = \text{erfc}\ (S_H/2^{0.5}\sigma_H) \quad ----[2.38]$$

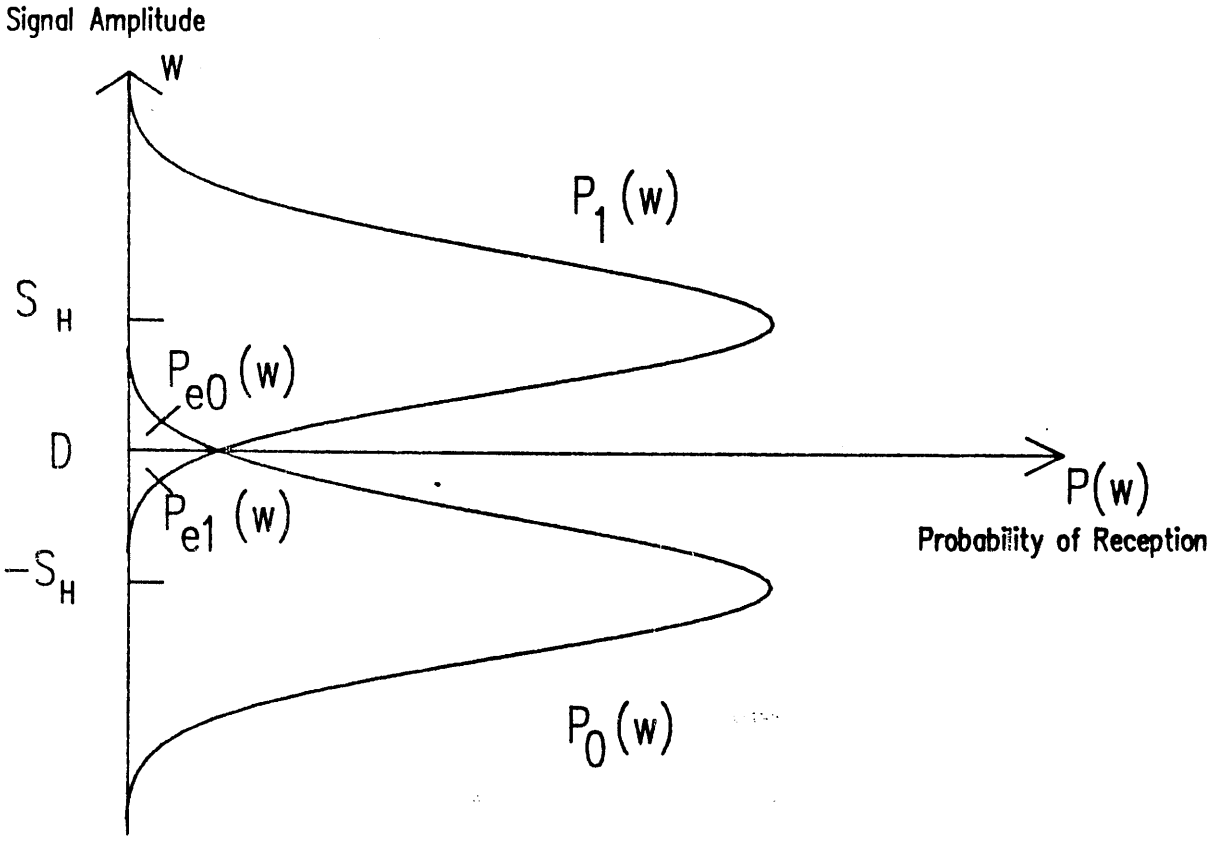


Figure 2.6.

Probability Density Functions

of FSK Coherent Detection.

Comparison of [2.38] with [2.32] shows a 3dB improvement in the sensitivity of an FSK heterodyne system over an ASK heterodyne system. Thus, for the parameters previously specified, the minimum received power in this case is -54.2dBm.

However, it should be noted that because in ASK power is transmitted only on a mark, the average power transmitted in an ASK coherent detection scheme for a specified bit stream is half that required for an FSK scheme. Thus, when comparing average transmitted powers, the two formats are found to be equally sensitive.

2.4.4. PSK Coherent Detection.

Phase shift keyed (PSK) transmission is also a constant carrier mode. However, instead of a change in carrier frequency, a change in carrier phase is used to transmit information. This can be represented as:-

$$E_T(t) = (2P_T)^{0.5} \sin (\omega_T t + \Delta\theta p(t)) \qquad \text{----[2.39]}$$

where ω_T is the carrier frequency (in radians), $\Delta\theta$ is the peak phase deviation, and p(t) is some binary switching function with possible states of ± 1. It is convenient to define a modulation index for [2.39] as:-

$$m = \cos \Delta\theta \qquad \text{----[2.40]}$$

Substituting this into [2.39] and using a standard trigonometric expansion, an expression in terms of carrier and modulation sideband signals can be obtained:-

$$E_T(t) = P_T^{0.5}\,[\sin\omega_T t.\cos[p(t)\cos^{-1}m] + \cos\omega_T t.\sin[p(t)\cos^{-1}m]]$$

$$= P_T^{0.5}\,[m.\sin\omega_T t + p(t)(1-m^2)^{0.5}\cos\omega_T t] \qquad ----[2.41]$$

From [2.41] it can be seen that the average power in the carrier component of the waveform is $m^2P_T/2$, and that in the modulation sidebands is $(1-m^2)P_T/2$. Thus, if $\Delta\theta = \pi/2$ $(m=0)$, all the available signal power is in the modulation sidebands. This special case of PSK is known as phase reversal keying or PRK, and, because sideband power is maximised, it is an optimum modulation format. The lack of a transmitted carrier can, however, be a major problem in demodulating the signal. Some method of regenerating a carrier component at the receiver is required in order to demodulate the signal with minimum performance degradation. Inevitably, this adds to receiver complexity (chapter 5).

It can be seen from [2.41] that the amplitude of the received modulation signal is modified by a factor of $p(t)\sin\Delta\theta$. Thus, the received signal level , S_H, becomes:-

$$S_H = S_H \sin\Delta\theta \qquad ----[2.42]$$

for a mark bit and:-

$$-S_H = -S_H \sin\Delta\theta \qquad ----[2.43]$$

for a space, $S_H = 2R(P_T.P_{LO})^{0.5}$ as before. Thus, the phase shift keyed signal is an antipodal signal with PDF's of the form shown in figure 2.7. This, as expected, indicates that maximum signal power and minimal PDF overlap is achieved at $\Delta\theta = \pi/2$.

Again, following a similar argument to section 2.4.2., the

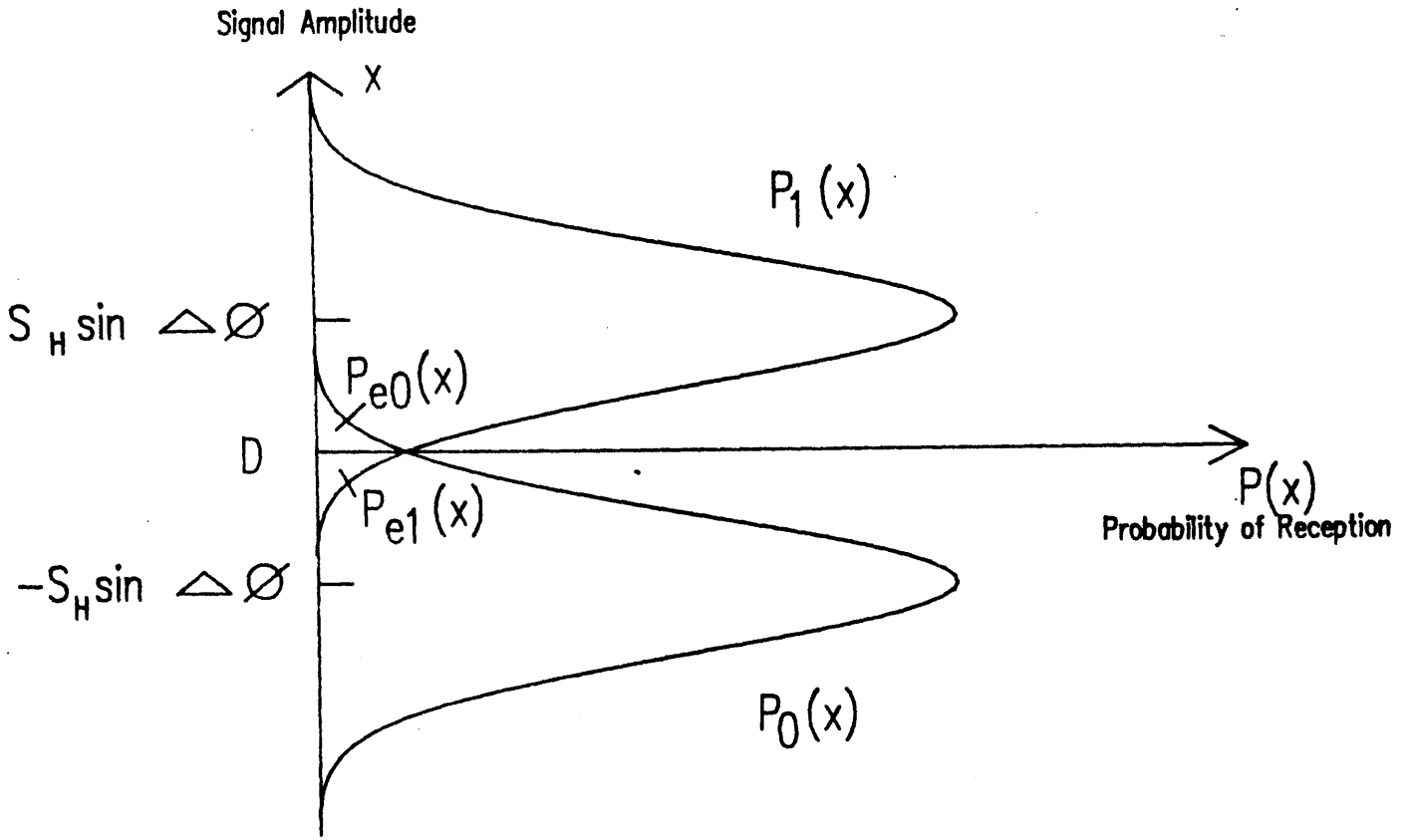


Figure 2.7.
Probability Density Functions
of PSK Coherent Detection.

probability of error can be derived. Taking the case of the mark bit:-

$$P_{e1}(x) = \frac{1}{(2\pi)^{0.5}.\sigma_H} \int_{-\infty}^{0} \exp -(S_H \sin\Delta\theta - x)^2/2\sigma^2_H \, dx$$

$$= \text{erfc}\left[\frac{S_H \sin\Delta\theta}{\sigma_H}\right] \qquad \text{----[2.44]}$$

and since $P_{e1}(x) = P_{e0}(x)$, and also assuming $P_0 = P_1 = 1/2$, then from [2.12], the total probability of error is :-

$$\text{BER} = \text{erfc}\left[\frac{S_H(1-m^2)}{\sigma_H}\right] \qquad \text{----[2.45]}$$

In a PRK system (m=0), this gives a 3dB sensitivity improvement over FSK heterodyne (cf. [2.38]). Hence, for the parameters previously discussed, an ideal heterodyne PRK receiver would require -57.2dBm of received power to achieve a 10^{-9} BER. However, if a modulation index of less than $\pi/2$ is used, this sensitivity will be reduced by $10 \log_{10}(1-m^2)$dB. Thus, a 1dB penalty will be incurred if $\Delta\theta$ is reduced from 90^o to 63^o, and the required power at the receiver input will be -56.2dBm.

2.4.5. Homodyne Coherent Detection.

The analysis so far has assumed that a heterodyne detection system (Figure 2.8) has been used. In order to avoid the effects of spectral overlap in this detection process, the Nyquist criterion requires that the bandwidth of the detector/front-end amplifier combination is equal to or greater than twice the

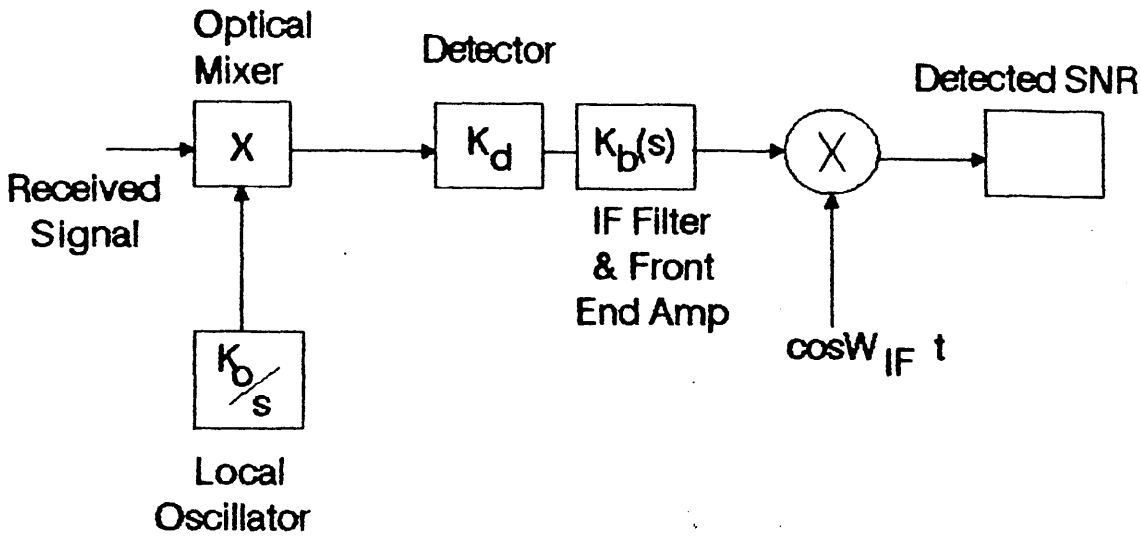


Figure 2.8. Heterodyne Detection System.

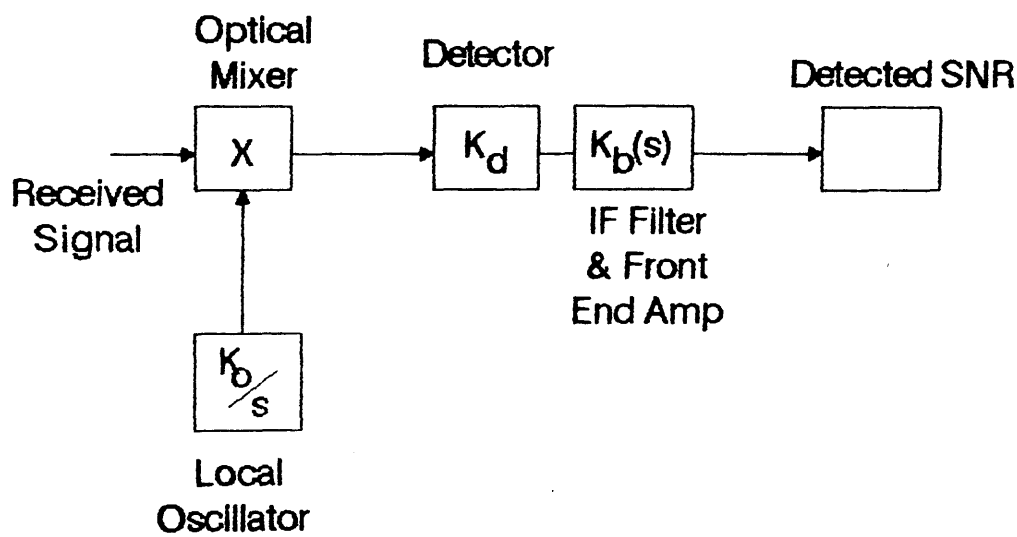


Figure 2.9. Homodyne Detection System.

maximum frequency present in the modulation (baseband) bandwidth[13]:-

$$B_H = 2.B_B \qquad \text{----[2.46]}$$

where B_B is the modulation bandwidth and B_H is the heterodyne receiver bandwidth centred on ω_{IF}. Since the local oscillator shot noise power is proportional to receiver bandwidth, a reduction in B_H would reduce σ^2_H and increase the receiver sensitivity. Such a reduction is possible in ASK and PSK systems by employing the system of figure 2.9. Here the local oscillator frequency is the same as the received signal ($\omega_T = \omega_{LO}$) converting the received data directly to baseband for detection, thus requiring a receiver bandwidth of only B_B. This produces a 3dB improvement in received SNR for homodyne systems against heterodyne systems. Unfortunately, as will be shown in chapter 5, this also requires an improvement in local oscillator and transmit source stability, which may well lead to a more complex and costly system.

2.5. BER Degradation In Coherent Systems Due To Laser Phase Noise.

The coherent systems considered in section 2.4 have been ideal in the sense that only fundamental noise sources, namely quantum and thermal perturbations in the detection process, have been included in the discussion. This has led to the set of theoretically achievable results listed in figure 2.10 (section 2.6). However, if it is now assumed that laser sources with non-negligible linewidths are to be used in such systems then some degradation in performance will be observed. The quantitative evaluation of this degradation is the subject of this section.

Spectral spread in a laser can arise from phase instabilities

both internal and external to the device itself[3,4,6]. The source of these instabilities will be discussed in chapter 3. It is sufficient to state that here that, in general, gas lasers are some four orders of magnitude more stable than currently available unmodified semiconductor lasers. However, for reasons to be discussed in chapter 3, it is the latter group of devices with which this project will concern itself. Their linewidths can be of the order of tens of megahertz. This results in substantial phase uncertainty at the receiver, and if not controlled in some manner, the effect on the system BER could be catastrophic. Depending on the type of modulation scheme used, various control techniques can be employed. Since, in section 2.4, it has been shown that a PSK format is the most efficient modulation scheme, the following argument will be confined to such a system. Moreover, the use of a phase locked or data synchronising control loop as distinct from a frequency locked loop will also be assumed.

The effect of phase noise can be analysed in the classical PLL sense as phase jitter in carrier recovery[4] and as such an analysis leading to a modified set of BER curves can be derived from standard rf techniques[8]. This phase jitter is incorporated in the photodetector output [2.26] as a phase error between the IF carrier and the reference LO signal:-

$$i_s(t) = S_H \cos[\phi_E(t) + \Delta\phi_n(t)] + x(t) \qquad \text{----[2.47]}$$

where $S_H = 2.R(P_T.P_{LO})^{0.5}$, $\phi_E(t) = \phi_T - \phi_{LO}$, and it has been assumed that $\omega_T = \omega_{LO}$. $x(t)$ is a white, gaussian noise term arising from the shot noise at the receiver input. It has zero mean and variance given by :-

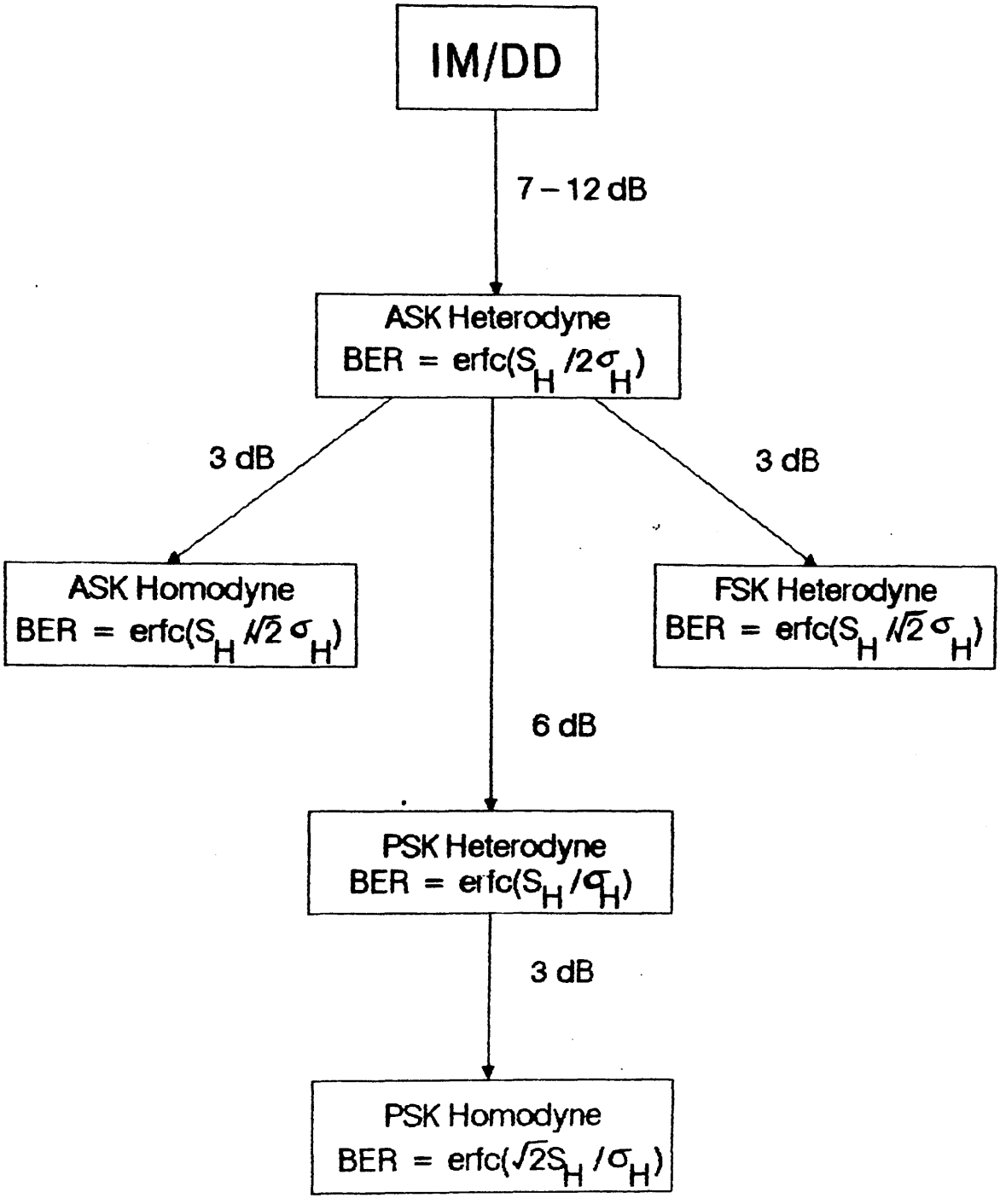


Figure 2.10. Relative Performance of Various Optical Detection Schemes.

$$\sigma^2_x = e.R.P_{LO}.B_B \qquad \text{----[2.48]}$$

thus equating to σ^2_H discussed in the previous section. The phase error variance term is also a zero mean gaussian random variable with variance:-

$$\langle \Delta\emptyset_n{}^2(t) \rangle = \sigma^2_\emptyset \qquad \text{----[2.49]}$$

As with the shot noise term, it is the variance of the phase noise that will determine the final system performance. The variance itself is a function of the bandwidth of the phase fluctuations $\emptyset_n(t)$ (which also determines the laser linewidth) and of the phase tracking circuit.

The inclusion of $\Delta\emptyset_n(t)$ will modify the demodulator output and hence the PDF's $P_0(x)$ and $P_1(x)$:-

$$i_s(t) = \begin{cases} S_H\cos(\Delta\emptyset_n(t)) + x(t) & \text{- mark transmission} \\ -S_H\cos(\Delta\emptyset_n(t)) + x(t) & \text{- space transmission} \end{cases}$$

With the signal still being antipodal the decision level D is unaltered. However, $P_{e0}(x)$ and $P_{e1}(x)$ now include a second random variable, $\Delta\emptyset_n(t)$. For a given $\Delta\emptyset_n$, the probability of error is given by [2.44][8,10]:-

$$BER_\emptyset = \text{erfc}\left[\frac{S_H\cos(\Delta\emptyset_n(t))}{\sigma_H}\right] \qquad \text{----[2.50]}$$

The average bit error rate is obtained by averaging [2.50] over the probability distributions of $\Delta\emptyset_n$, yielding:-

$$BER = \frac{1}{2}\int_{-\infty}^{\infty} \text{erfc}\left[\frac{S_H \cos(\Delta\phi_n(t))}{\sigma_H}\right] p(\phi)\, d\phi \qquad \text{----[2.51]}$$

where $p(\phi)$ is the PDF of $\Delta\phi_n(t)$. The resultant BER curves were calculated by Prahbu[8] for two distributions of $p(\phi)$, gaussian and truncated gaussian. In the majority of practical cases the most useful distribution is the zero mean truncated gaussian (or Tikhonov) distribution. This describes most accurately the band limited nature of the noise process being considered. Bit error rate calculations for this (solid lines) and the zero mean gaussian case (dashed lines) are plotted in figure 2.11 for various values of σ_ϕ. In the Tikhonov distribution case it is shown that for a 10^{-9} BER, a 0.5dB SNR penalty is incurred if $\sigma_\phi = 10^o$ ($\sigma^2_\phi = 0.03$ rads^2). Moreover, if $\sigma_\phi = 12.5^o$ ($\sigma^2_\phi > 0.047$) then a 10^{-9} BER is not achievable. This BER figure is a common specification for commercial communications links and therefore places important performance limitations on practical implementations of phase control loops.

2.6. Optical Detection Systems: Discussion and Conclusions.

This section has defined and quantified the basic noise sources relevant to optical detection. In IM/DD systems these sources were shot noise from signal, background light, and dark current terms, and thermal noise from the electrical components in the receiver. In coherent systems it was shown that in order to maximise the SNR at the receiver input, the local oscillator power should be made large enough to make the resultant shot noise term the dominant source.

The theoretical minimum probability of error, or BER, was evaluated for an IM/DD system and several coherent modulation

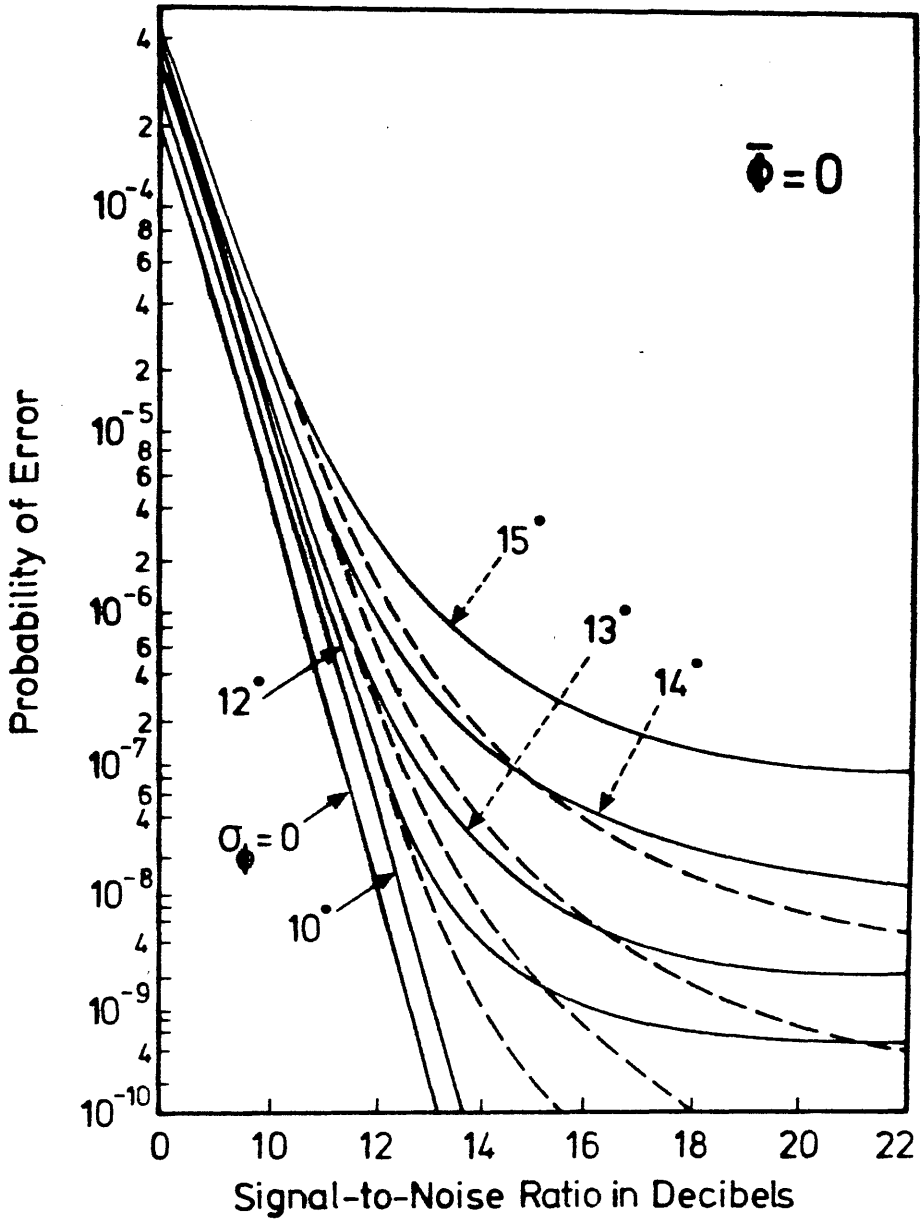


Figure 2.11 Probability of Bit Error for Binary PSK Transmission[8].

Solid Lines denote Tikhonov Distribution. Dashed Lines denote Gaussian Distribution. Legends denote Values of σ_ϕ.

formats, and it was shown that since coherent detection systems can achieve shot noise limited detection, they can offer between 10 and 25dB improvement in receiver sensitivity over direct detection systems depending on operational wavelength, bit rate, and receiver efficiency. It was also shown that the most efficient modulation format and detection scheme was homodyne detection of a phase reversal keyed (PRK) waveform. A comparison of the main results of sections 2.4 and 2.5 is shown in figure 2.10.

Finally, the degradation of BER due to spectral spread in the local oscillator and transmit lasers of coherent systems was discussed. It was shown that an increase in the probability of error will occur proportional to the variance of the phase noise error at the receiver. In the case of a Tikhonov (zero mean, truncated gaussian) phase noise distribution it was shown that if $\sigma^2_\emptyset = 0.03$ rads^2 then a 0.5dB increase in SNR will be required to maintain a BER of 10^{-9}. If, however, $\sigma^2_\emptyset > 0.047$ then a 10^{-9} BER is not achievable. It was stated that $\sigma^2_\emptyset$ is dependent on the linewidth of the sources used and also on the bandwidth of the phase tracking receiver. These results will be used extensively in the analysis of optical phase locked loops in chapters 5 and 6.

References.

1. T.Okoshi et al 'Computation of Bit Error Rate of Various Heterodyne and Coherent Type Optical Communication Schemes.' J. Opt. Commun. vol 2(3) 1983. pp89-96.

2. R.C.Steele, PhD Thesis, University of Glasgow, 1983.

3. Y.Yamamoto 'Receiver Performance Evaluation of Various Digital Optical Modulation-Demodulation Systems in the 0.5-10μm Wavelength Region.' IEEE J. Quant. Elect. 16(11) 1980. pp1251-1259.

4. S.Saito, Y.Yamamoto, T.Kimura 'S/N and Error Rate Evaluation for an Optical FSK Heterodyne Detection System Using Semiconductor Lasers.' IEEE J. Quant. Elect. 19(2) 1983. pp180-193.

5. L.G.Kazovsky. 'Optical Heterodyning vs Optical Homodyning: A Comparison.' Journal of Optical Communications, vol 6(1985) no.1, pp18-24.

6. K.Kikuchi, et al 'Degradation of Bit Error Rate in Coherent Optical Communications Due to Spectral Spread of the Transmitter and Local Oscillator.' J. Lightwave Technol. 2(6), 1984. pp1024-1033.

7. J.Franz 'Evaluation of the Probability Density Function and Bit Error Rate in Coherent Optical Transmission Systems Including Laser Phase Noise and Gaussian Additive Noise.' J.

Opt. Commun. 6(2), 1985. pp51-57.

8. V.K.Prabhu 'Psk Performance with Imperfect Carrier Phase Recovery.' IEEE Trans. on Aero. and Elect. Syst. Vol-AES 12(2) 1976. pp275-286.

9. W.C.Lindsey 'Phase Shift Keyed Signal Detection with Noisy Reference Signals.' IEEE Trans. on Aero. and Electron. Systems. Vol AES.2(4), 1966. pp393-401.

10. B.Glance ' Performance of Homodyne Detection of Binary PSK Optical Signals.' J. Lightwave Technol. 4(2) 1986, pp228-235.

11. S.D.Personick 'Receiver Design for Digital Fibre Optic Communication System' Bell Syst. Tech. Journal vol 52, 1973. pp843-886.

12. J.E.Midwinter 'A Study of Intersymbol Interference and Transmission Medium Instability for an Optical Fibre System' Opt. Quant. Electron. vol 9, 1977. pp299-304.

13. F.G.Stremler 'Introduction to Communication Systems.' 2nd Ed. Addison-Wesely, Reading, Massachusettes, 1982.

CHAPTER 3.

SEMICONDUCTOR LASER PHASE NOISE AND LINEWIDTH REDUCTION.

3.1. Introduction.

It was shown in chapter 2 that the performance of a coherent optical phase locked loop (OPLL) was dependent on the magnitude of the phase instabilities of the lasers used in the system. It would therefore be desirable to utilise lasers with minimal phase noise components or, in other words, narrow linewidths. If consideration is confined to optical fibre communication systems, which will support propagation in the near infra-red, then only two types of laser can be utilised, Helium-Neon (HeNe) gas lasers, or semiconductor lasers (either GaAlAs or quaternary).

On initial inspection, HeNe lasers satisfy a number of the criteria for successful coherent demodulation. Cavity lengths are around 30cms and the lasing medium is not dense. As a result optical confinement is high, cavity losses are low, and hence the cavity Q is high. Linewidths of the order of a few kHz are easily achievable. If the device is carefully handled, then the output emission is also relatively stable. On the other hand, its size makes it cumbersome to manipulate, the optics required to couple light into external components can be sizeable, and the process of coupling into these components is prone to mechanical vibration. Therefore, although the emission of the solitary laser might be quite stable, the overall system output may well be unacceptably transient. In addition, pumping of the gaseous lasing material requires a high level voltage source and, consequently, a power supply that is both complicated and large.

Similarly, the PZT positioners which when attached to the laser mirror are used to implement frequency control over the

output line, require a high voltage drive. Moreover, these devices have a maximum modulation rate of only ~50kHz. Hence, this necessitates the use of an external modulator not only for data encoding, but also at the local oscillator if a receiver loop bandwidth of greater than a few kHz is to be achieved. This introduces yet more mechanical instability into the system.

Semiconductor lasers, on the other hand, are small in size (typically 300μm (length) x 200μm (width) x 100μm (height)), draw a drive current of only one to two hundred milliamps at low voltage levels (<5v), and are also capable of being directly modulated through variation of the drive current. In addition, coupling between laser and fibre can be effected using butt coupled[1,2,3,4] or lens fibres. These are soldered into position using low temperature Indium solder only a few microns away from the laser chip. The laser/fibre combination can then be sold as a package with a fibre pigtail, the pigtail being fused later to the transmission system. Thus, physically, a semiconductor module can be made very small, and mechanically, it can also be made relatively stable[4]. In the future, it should be possible to integrate several active devices such as lasers, pin diodes, electrooptic modulators, etc, onto the one semiconductor wafer. This would substantially reduce coupling losses between components, as well as the overall size of a working system. Indeed, it is possible to envisage a complete optical receiver with local oscillator, directional coupler, pin diodes, and FET amplifier all in the one package, perhaps two to three centimeters square, with a fibre pigtail providing the interface between the receiver and the rest of the system. Thus it can be seen that semiconductor devices hold out a number of benefits for fibre optic communication systems.

If the use of semiconductor lasers is assumed, it is then necessary to characterise and quantify their spectral imperfections. In this chapter, an overview of the theoretical analysis of semiconductor lasers, including expressions for spectral purity and lineshape[5-14], will be presented. The assumption will then be made that the linewidths of solitary semiconductor lasers will not be good enough to satisfy the requirements of coherent communication systems. Consideration will then be given to various techniques that will effect a reduction in the linewidth[1-4,15-22].

3.2. Theoretical Calculations of Semiconductor Laser Linewidth.

There are a number of factors that must be considered when discussing laser noise properties. The cavity Q, or quality factor, of the optical resonator largely determines its fundamental linewidth. This, in turn, is a function of the cavity length, the losses within the cavity, and its overall volume. The losses in the resonator are particularly high in semiconductor lasers due to its high material density, and poor optical confinement. In addition, the effective facet reflectivity is also a vital element in determining the overall optical loss.

Reflectivity is normally induced through simply cleaving the semiconductor facets at appropriate points. This gives rise to the most common type of laser, the Fabry Perot resonator[23-25]. Typical facet reflectivity in these devices is around 30%. Reflectivity can, in addition, be provided by gratings, as is used in distributed feedback laser (DFB) and distributed Bragg reflector (DBR) lasers[16]. The addition of a reflective element placed outside the basic Fabry Perot cavity will also modify its performance. This can produce, under suitable conditions,

desirable line narrowing effects[1,3,4,17,26-31]. These will be discussed in section 3.3. Initially, however, the properties of the solitary laser will be examined.

3.2.1. Basic Laser Equations.

In this section, the basic rate equations for the electromagnetic field in the cavity will be considered. Their dependence on perturbations in the material refractive index induced through carrier density fluctuations will be evaluated, and the effect of these perturbations on the laser lineshape will be discussed.

It is well known that an electromagnetic wave propagating in some medium will obey a wave equation of the form:-

$$\frac{\partial^2 E}{\partial z^2} = \frac{1}{c^2} \cdot \frac{\partial^2 (\varepsilon E)}{\partial t^2} \qquad \text{----[3.1]}$$

where E is the electric field vector, ε is the dielectric constant of the medium, and c is the velocity of light. The E vector for the field in the cavity is denoted by:-

$$E = \beta \exp[j(\omega t - kz)] + \beta \exp[-j(\omega t - kz)] \qquad \text{----[3.2]}$$

where β is the complex amplitude , ω is the angular frequency, k is the propagation constant, and z is the direction of propagation of the wave. In order to accommodate material dispersion (the variance of ε with ω) caused by the time dependence of β, [3.2] is rewritten as[6]:-

$$\varepsilon E = (\varepsilon(\omega)\beta + j\partial\varepsilon/\partial\omega.\dot{\beta}) \exp[j(\omega t - kz)] + c.c. \qquad \text{----[3.3]}$$

where $\dot{\beta} = \partial\beta/\partial t$ and c.c is the complex conjugate of the expression. Substituting [3.2] and [3.3] into [3.1], two equations are obtained, one for β and one for β^*, the complex conjugate of β. Ignoring derivatives of the second order or higher yields an equation for β of:-

$$\frac{2j\omega}{c}\left[\varepsilon + \frac{\omega}{2}\frac{\partial\varepsilon}{\partial\omega}\right]\dot{\beta} = \left[\frac{\omega^2}{c^2}\varepsilon - k^2\right]\beta \qquad \text{----[3.4]}$$

It is important to note here that the dielectric constant, ε, is complex and can be written as the square of a complex refractive index, $n - jn'$. This is brought about by the interaction of the EM field with the constituent molecules of the medium, giving rise to both field dispersion (n), and field attenuation (n').This attenuation is modelled as a modification of the propagation constant k, and is a standard process. The net gain of the resonator can be linked to the refractive index as:-

$$g - \alpha_j = -\frac{2\omega}{c}n' \qquad \text{----[3.5]}$$

where g is the cavity gain per unit length and α_j is the cavity loss per unit length, including facet losses, which are approximated as being distributed uniformly over the whole cavity. At threshold, the point at which lasing begins, $g = \alpha_j$, ε is real (=> $n' = 0$), and therefore ω is chosen so that:-

$$\frac{\omega n}{c} = k \qquad \text{----[3.6]}$$

Changes in the carrier (hole/electron) density of the material will cause n and n' to deviate from their steady state threshold values. Thus, if n'=0 at threshold, and neglecting Δ^2 terms, ε is written as:-

$$\begin{aligned}\varepsilon &= [n + \Delta n - i(n' + \Delta n')]^2 \\ &= n^2 + 2n(\Delta n - i\Delta n') \\ &= n^2 - 2in.\Delta n'(1 + i\alpha) \end{aligned} \qquad \text{----[3.7]}$$

where

$$\alpha \equiv \frac{\Delta n}{\Delta n'} \qquad \text{----[3.8]}$$

The bracketed term on the left hand side of [3.4] is related to the group velocity V_g by:-

$$\begin{aligned}\varepsilon + \frac{\omega}{2}\frac{\partial \varepsilon}{\partial \omega} &= n\left(n + \omega\frac{\partial n'}{\partial \omega}\right) \\ &= \frac{c}{V_g} - n' \end{aligned} \qquad \text{----[3.9]}$$

Substituting [3.6], [3.7], and [3.9] into [3.4] yields an expression for β in terms of cavity gain, loss, and complex refractive index:-

$$\dot{\beta} = \frac{g - \alpha_j}{2} V_g (1 + i\alpha)\, \beta \qquad \text{----[3.10]}$$

In a dispersive medium, the group velocity, V_g, is defined as the ratio of energy flux to energy density in the medium.

Consequently, the rate of stimulated emissions, G, minus the rate of loss, γ', is given by:-

$$G - \gamma' = V_g \, (g - \alpha_j) \qquad \text{----[3.11]}$$

Thus, the final equation for β is:-

$$\dot{\beta} = \frac{[\, G - \gamma' \,]}{2} (1 + i\alpha) \, \beta \qquad \text{----[3.12]}$$

An identical argument realises a similar equation for β^*:-

$$\dot{\beta}^* = \frac{[\, G - \gamma' \,]}{2} (1 - i\alpha) \, \beta^* \qquad \text{----[3.13]}$$

It is now desirable to relate these equations for β to the intensity, I, and the phase, $\emptyset$, of the field in the diode cavity:-

$$I = \beta . \beta^* \qquad \text{----[3.14]}$$

and

$$\emptyset = \frac{1}{2} \ln \left(\frac{\beta}{\beta^*} \right) \qquad \text{----[3.15]}$$

Using [3.12] and [3.13] in [3.14] and [3.15] yields equations for the derivatives with respect to time of I and $\emptyset$:-

$$\dot{I} = \dot{\beta} . \dot{\beta}^* = [\, G - \gamma' \,] \, I \qquad \text{----[3.16]}$$

and

$$\dot{\emptyset} = \frac{\alpha}{2} [\, G - \gamma' \,] \qquad \text{----[3.17]}$$

[3.16] and [3.17] are generally recognised as the simple rate equations for the EM field in the laser cavity. They relate the intensity and phase of the field to the rate of stimulated emission and loss in the cavity. They also show a dependence of the phase of the field on the complex refractive index, which, as has been stated, is a function of the interaction of the field with the constituent molecules of the cavity. This interaction is much greater in semiconductor lasers than in gas lasers as the lasing medium is considerably denser. This is a major reason for the poorer noise properties of these devices.

3.2.2. Spontaneous Emission Events.

In any laser cavity, this material interaction takes the form of spontaneous emission events. In a semiconductor laser, these events are defined as the spontaneous transition of an electron from a high energy level in the conduction band of the semiconductor material, to a lower level in the valance band. This event differs from a stimulated emission in that it is not induced by the field in the cavity. As such it alters discontinuously the phase and intensity of the field as illustrated in figure 3.1[5,6,32].

These events do not only cause an instantaneous phase change. They also produce a delayed phase change resulting from the change in carrier intensity brought about by the loss of a carrier from the upper energy level in the laser. This change then induces a relaxation process to restore the steady state value of the field (Figure 3.2). These oscillations have a period of the order of 1ns, and a decay rate γ. During this time there will be a net gain change of:-

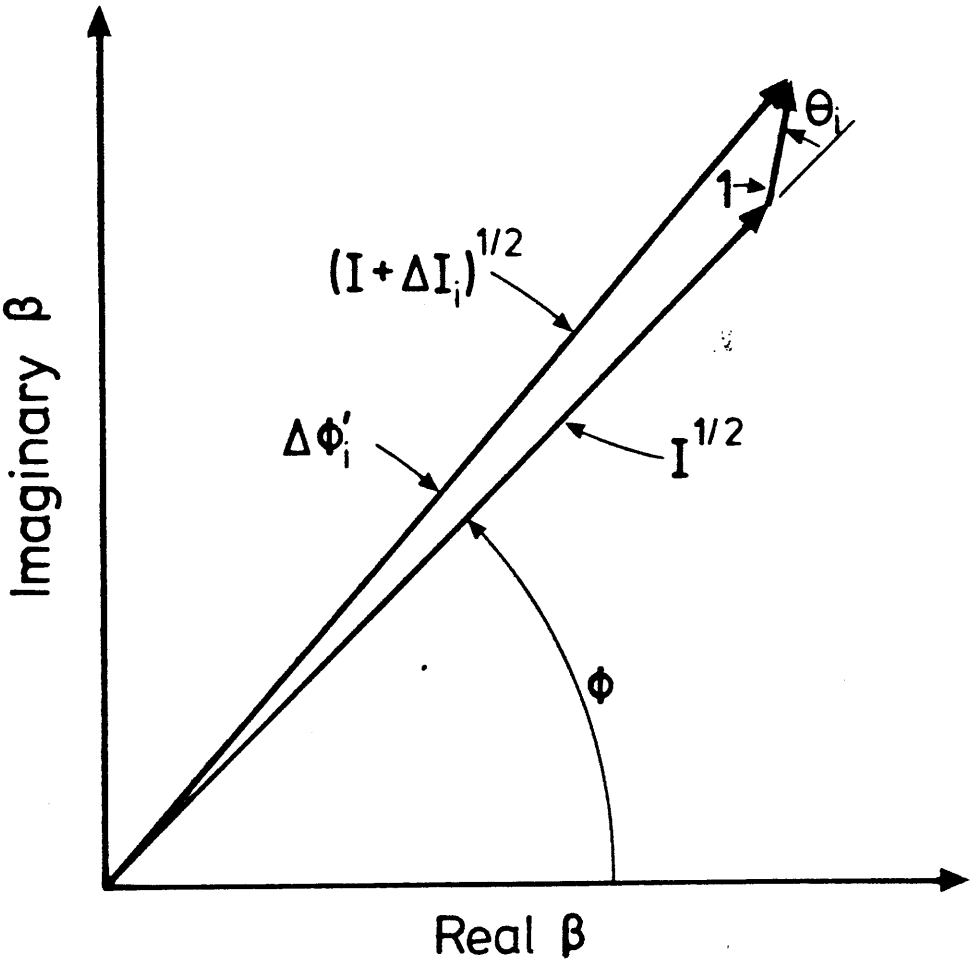


Figure 3.1 Instantaneous Phase Vector Diagram [5].

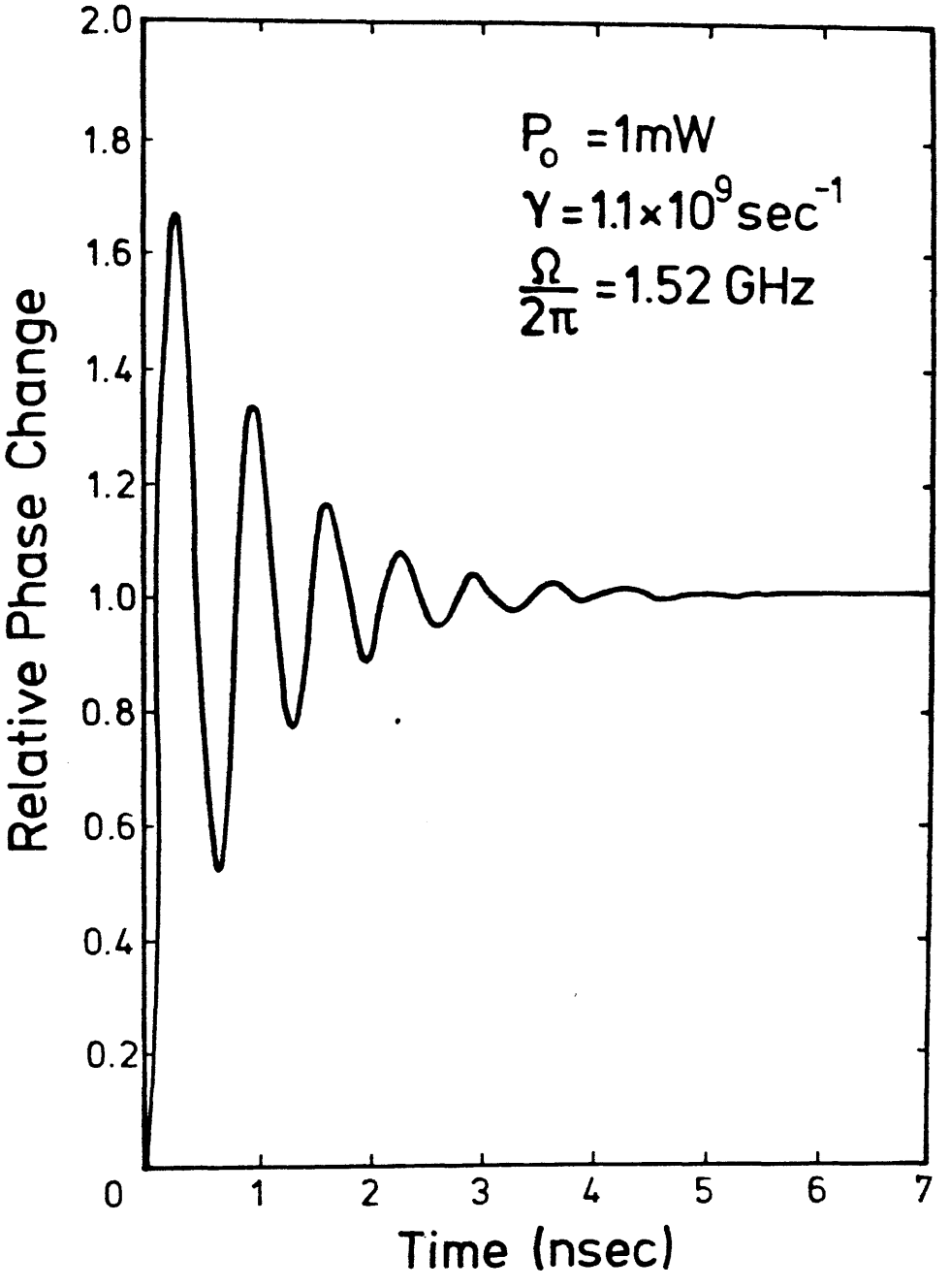


Figure 3.2 Intensity Relaxation Oscillations as a Result of a Spontaneous Emission Event [6].

$$\Delta g(t) = -\left[\frac{2\omega}{c}\right]\Delta n'(t) \qquad \text{----[3.18]}$$

where $\Delta n'(t)$ is the magnitude of the deviation of the imaginary refractive index from its steady state value. This change combined with the change in the real part of refractive index, Δn, gives rise to the ratio α [3.8], which now becomes a measure of the rate of stimulated emissions in the cavity.

3.2.3. Laser Linewidth Formula.

It can be shown that the above noise mechanism produces substantially more line broadening than is predicted by the standard Schawlow-Townes formula[33]. A number of workers[6,7,8] have analysed this phenomena and produced a modified linewidth formula, which corresponds well to reported experimental results[5,7,34]. Although the analysis is beyond the scope of this report, the general procedure applied to obtain this formula from the rate equations will be outlined here. The formula is derived by taking the Fourier transform of $\langle\beta^*(t).\beta(0)\rangle$, the autocorrelation of the complex amplitude of the emitted field after time t. It can be shown[35] that:-

$$\langle\beta^*(t).\beta(0)\rangle = |\beta(0)|^2 \exp(-|t|/t_{coh}) \qquad \text{----[3.19]}$$

where the coherence time, t_{coh}, is the duration of the interval of temporal coherence of the laser radiation, and is given by:-

$$t_{coh}^{-1} = \frac{\langle \Delta\emptyset^2 \rangle}{2t} \qquad \text{----[3.20]}$$

$\langle \Delta\emptyset^2 \rangle$, the variance of the phase perturbations initiated by spontaneous emission, is[6]:-

$$\langle \Delta\emptyset^2 \rangle = \frac{R(1 + \alpha^2)t}{2I} \qquad \text{----[3.21]}$$

R is the rate of spontaneous emission. This method is justified since the absolute phase of the field executes a random walk process under the influence of the phase fluctuation variable, $\Delta\emptyset$. Hence, although the noise process is Poisson, in the limit of large R the distribution of $\Delta\emptyset$ is Gaussian. Having established this, it can be shown that the Fourier transform of [3.19], including [3.20] and [3.21] is indeed Lorenzian with a full width half maximum (FWHM) of:-

$$\delta f = \frac{1}{\pi t_{coh}} = \frac{R(1 + \alpha^2)}{4\pi I} \qquad \text{----[3.22]}$$

Defining I in terms of output power per facet, P_o, group velocity, V_g, and facet loss, α_m yields :-

$$I = \frac{2P_o}{hv.V_g.\alpha_m} \qquad \text{----[3.23]}$$

In addition, defining R in terms of gain, g, and a spontaneous emission factor, n_{sp} :-

$$\frac{R}{V_g} = g.n_{sp} \qquad \text{----[3.24]}$$

will then permit an equation for the linewidth of the solitary laser cavity to be derived:-

$$\delta f = \frac{V_g^2 h\nu.g.n_{sp}\alpha_m(1 + \alpha^2)}{8\pi P_o} \qquad \text{----[3.25]}$$

This equation is $(1 + \alpha^2)/2$ larger than the classic Schawlow-Townes equation[33] and returns values in accord with current experimental results[34]. For this reason α is generally known as the linewidth enhancement factor. Typical values for the terms in [3.25] for a GaAlAs laser are[6,13] $\alpha_m = 0.42m^{-1}$, $h\nu = 1.5eV$, $\alpha_L = 0.45m^{-1}$, $n_{sp} = 2.6$, $\alpha = 4-7$, $g = 0.87m^{-1}$, and $V_g = c/4.33$. At an output power of $P_o = 1mW$, the resultant linewidth is ~28MHz.

3.2.4. Lineshape Anomalies.

The analysis outlined on spontaneous emissions and laser linewidth is a major step forward to quantifying the noise contribution of semiconductor lasers to coherent detection. However, further research into the effects of spontaneous emission[6,7] has shown that the lineshape of the emitted spectra is significantly degraded by this phenomena. Again, the analysis of this effect is lengthy and complicated. However, the results have important implications for coherent systems and therefore the major conclusion of this work will be outlined here.

It has been shown both experimentally by Daino et al[7], and theoretically by Spano et al[8] that the power spectral density of

the instantaneous frequency deviations, $\delta\emptyset/\delta t$, instead of being flat, which would result in the predicted Lorenzian line, does in fact show major peaks at the resonant frequency of the relaxation oscillations (3.2.3) (Figure 3.3). This then translates into a peak in the emission line spectra at that frequency (Figure 3.4) which, although some 10-30dB down on the fundamental line power depending on the facet output power, P_o, and relaxation oscillation damping coefficient, γ[6], results in a significant raising of the noise floor. This peak occurs at between 1.0 and 2.5 GHz depending on laser type, structure, and the above operating parameters.

This result is significant to coherent communication systems since if the peak falls either within the loop bandwidth of a phase tracking receiver, or close to the bit rate of the system, a severe degradation in performance is likely to occur.

3.3. Techniques for Linewidth Reduction.

The above discussion has shown that the spectral linewidth of the solitary laser chip is large and, intuitively, may well be unsuitable for coherent communication systems. A considerable amount of research has been devoted to techniques for improving the spectral purity and stability of these devices. These techniques involve various forms of self-locking of the laser, where a portion of the optical field is coupled back into the gain medium after some delay time, τ. This increases the temporal coherence of the device. The reflecting element can be additional to the original laser facets[1,3,21], or, by means of an antireflection (AR) coating on one facet, replace that facet completely, thereby effectively extending the length of the cavity[4,19,22,36,37]. These two techniques form coupled and

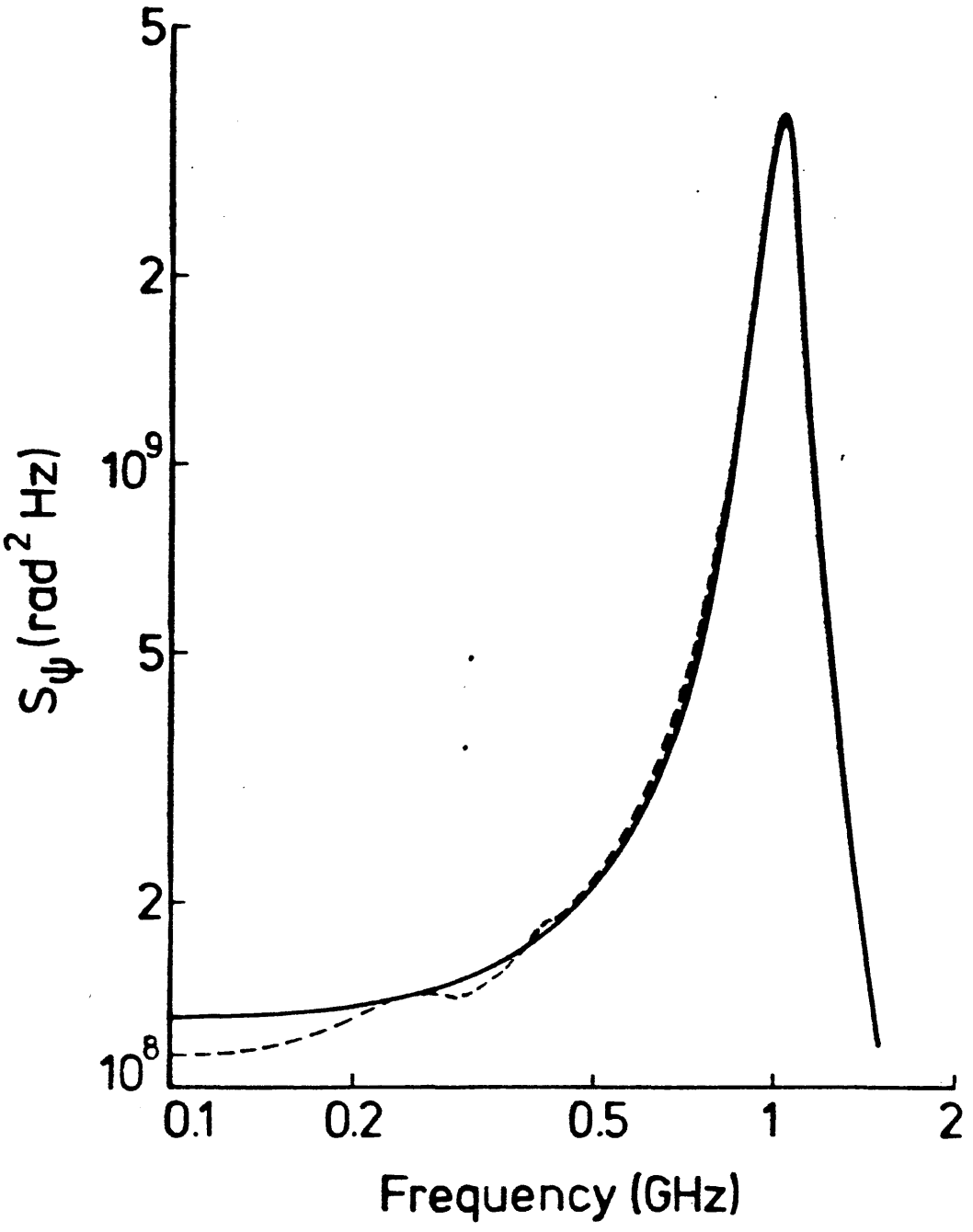


Figure 3.3 Two Sided Power Spectral Density of the Instantaneous Frequency Deviation $S_{\dot{\phi}}(\omega)$ for an ITT LS7709 Laser Diode Operating at 865.5nm [8].

Dashed curve denotes Experimental results.
Solid curve denotes Theoretical results.

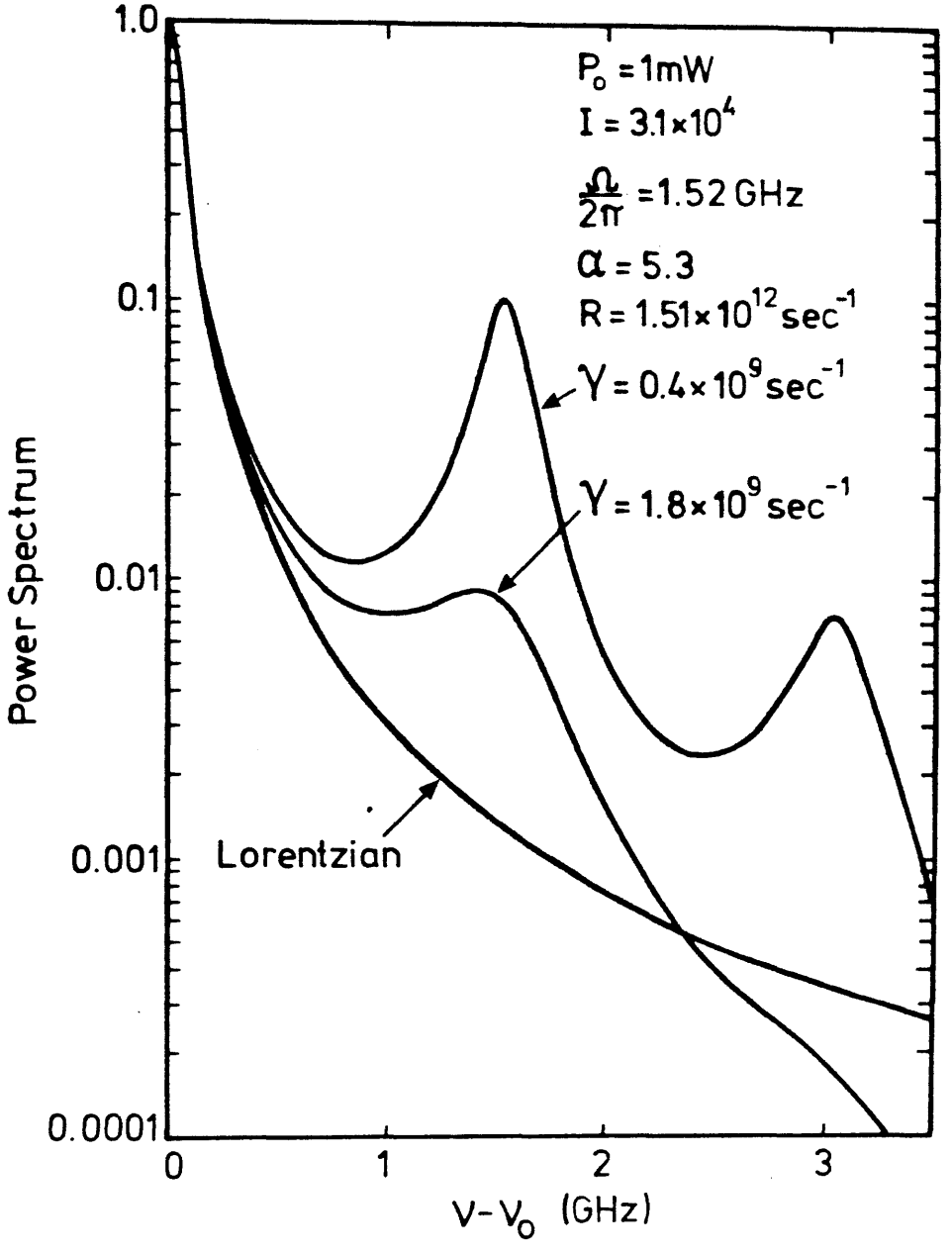


Figure 3.4 Power Spectrum of Laser Line Calculated for Various Conditions of Relaxation Oscillation Damping [6].

external cavities respectively.

An overview of the analysis of optical feedback in a coupled cavity module will be given in this section, focusing primarily on alterations to the cavity field induced by the self-locking process, and on the resultant changes in the rate equations for a solitary laser, given in section 3.2.2.. An expression for the modified linewidth will then be given[17,26,27], followed by consideration of the modification of the phase noise spectrum which also occurs[15,38,39,40].

3.3.1. The Three Facet Optical Model.

3.3.1.1. Rate Equations.

To examine the effect of optical feedback on the overall linewidth of the laser, attention will first be confined to a coupled cavity three facet model as shown in figure 3.5. A partially reflecting surface with reflection coefficient r_m is placed a distance d from the laser output. Provided $r_m \ll 1$, multiple reflections in the passive cavity can be ignored, and the reflected portion of the output adds coherently to the laser field. This field can be described , as in 3.2.2, as:-

$$E(z,t) = \mathrm{Re}\ \{E(t)\ \exp(j\omega_o t)\}\sin kz \qquad \text{----[3.26]}$$

Substituting this into the wave equation [3.1] realises:-

$$\dot{E}(t) = j(\omega_o - \Omega)E(t) + 1/2\ (G - \gamma')(1 - j\alpha)E(t) + kE(t - \tau)e^{j\phi_o} \qquad \text{----[3.27]}$$

where ω_o is the free running laser frequency, Ω is the diode-cavity longitudinal mode frequency, and, in the absence of

feedback, the laser is assumed to have a single mode at $\omega_o = \Omega$. The parameters G, γ', and α are as defined in 3.2.2. The last term in [3.27] arises from the feedback facet: k represents the feedback coefficient and is defined as:-

$$k = \frac{(1 - r_c^2)r_m}{r_c\tau_c} \qquad \text{----[3.28]}$$

and ϕ_o is the passive cavity phase shift:-

$$\phi_o = \omega_o + \phi_m \qquad \text{----[3.29]}$$

Here, r_c is the laser facet reflectivity, and τ_c is the diode-cavity round trip time, while ϕ_m is the phase shift introduced at the reflecting element, and $\tau = 2d/c$ is the passive cavity round trip time. Equating $E = I^{1/2} e^{j\phi}$, and using this in [3.27], the following set of coupled rate equations can be obtained:-

$$\dot{I}(t) = (G - \gamma')I(t) - 2k[I(t)I(t-\tau)]^{1/2}\cos[\phi(t) - \phi(t-\tau) + \phi_o] \qquad \text{----[3.30]}$$

and

$$\dot{\phi}(t) = \frac{\alpha}{2}(G-\gamma') - (\omega_o-\Omega) - k\left[\frac{I(t-\tau)}{I(t)}\right]^{1/2}\sin[\phi(t)-\phi(t-\tau)+\phi_o] \qquad \text{----[3.31]}$$

Comparison of [3.30] and [3.31] with [3.14] and [3.15] show the dependence of the emitted field on the properties of the coupled cavity. Effectively, the output frequency has been pulled by the passive cavity onto a new frequency, Ω, determined by the length

of the cavity and the reflectivity of the external facet. A number of workers[13,15,16,17,36,38] have analysed this process, deriving a modified linewidth formula from the above equations. Agrawal[17] in particular, presents a rigorous analysis using the Langevin noise force approach, and accounts for the effect of spontaneous emissions postulated by Henry[5,6], and outlined in sections 3.2.3 and 3.2.5. This leads to a result that is significantly different from that derived by Kikuchi et al[26] predicting significantly more line narrowing for a given value of $k\tau$.

3.3.1.2. Linewidth Formula.

The analysis presented by Agrawal is beyond the scope of this report. However, his results are particularly relevant to this work. He concludes that the linewidth of a laser in a coupled cavity configuration is related to the solitary laser linewidth as follows:-

$$\delta f = \frac{\delta f_{sl}}{[1 + X\cos(\phi_o - \phi_R)]^2} \qquad \text{----[3.32]}$$

where the feedback parameter X (common to a number of workers) is given by :-

$$X = k\tau(1 + \alpha^2)^{1/2} \qquad \text{----[3.33]}$$

and ϕ_R is given by:-

$$\phi_R = \tan^{-1}\alpha \qquad \text{----[3.34]}$$

It is evident from [3.32] that the extent of line narrowing is

dependent on the phase of the return of the external facet. Maximum reduction occurs when the phase of the return is resonant ie $\emptyset_o + \emptyset_R = 2m\pi$. Under this condition, expression [3.32] reduces to the well known equation:-

$$\frac{\delta f}{\delta f_{sl}} = \frac{1}{[1 + X]^2} \qquad \text{----[3.35]}$$

It would appear obvious then that the larger the value of X, the greater the possible reduction in linewidth. However, analysis of the feedback conditions by Goldberg et al[39] and others[26,27] has shown that above a certain level (~X = 10), and under certain phase conditions, multiple solutions to the wave equation occurs, resulting in multimode operation. In addition to possible multimode operation, line broadening has been observed[39,41] resulting eventually in the phenomena of 'coherence collapse' where the line can widen to some 25GHz. Recent work by Henry[42] has analysed this effect and shown it to be a result of second order fluctuations in phase, and carrier number density, in the cavity resulting from fluctuations in the rate of spontaneous emission. He concludes that increasing the level of feedback still further will reduce this effect to a minimum, although multimode operation of the laser system is still likely to occur.

3.3.1.3. Phase Noise Spectra.

The introduction of optical feedback in general will alter certain characteristics of the solitary laser other than the linewidth. Direct current modulation deviation will be reduced impairing the modules ability to act as a fast FSK transmitter, and increasing the need for external modulators. The phase noise

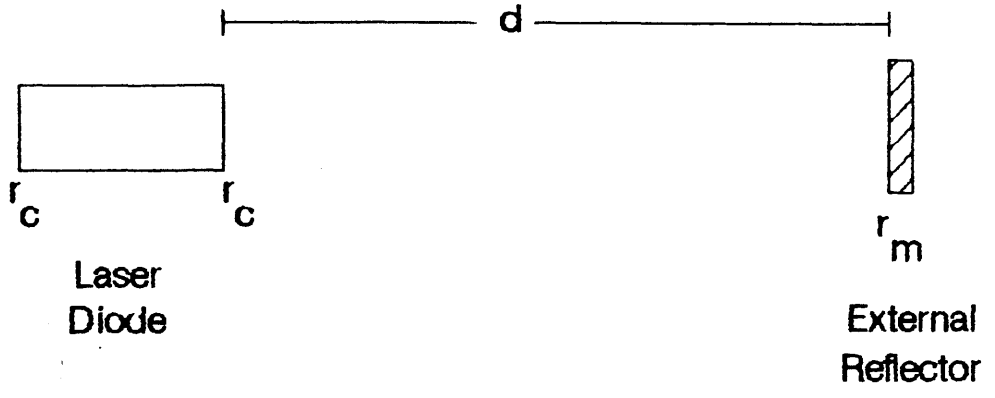


Figure 3.5

Schematic of Three Facet Model of a Coupled Cavity Laser.

power spectral density will also be altered. Analysis of this effect by Spano et al[15] has shown that the shape of this spectral density is dependent on the feedback conditions, that is the magnitude of the feedback signal, k, and the round trip time of the cavity, τ. If τ is larger than the inverse of the relaxation oscillation frequency, f_r, then enhancement of the noise peaks will occur. If, however, τ is less than f_r, then a flattening of the noise peaks will occur. Both these conditions will be in addition to the general reduction of the low frequency portion of the power spectrum which is responsible for the fundamental linewidth reduction. For noise peaks at ~1GHz, this leads to a cavity length of less than 15cm. Figure 3.6 shows theoretical spectra calculated by Spano et al[15] for various conditions of feedback. The suppression of these peaks is significant and desirable for those reasons outlines in 3.2.5. However, this will not be considered a major target in this project, and, therefore, experimental work will not be bound by the above results.

3.3.1.4. Practical Implementation of the Three Facet Model.

There are a number of methods which can be used to realise the practical implementation of this model. The simplest of these is the partially reflecting mirror mounted in the beam path. Alternatively, a diffraction grating or some other optical component such as an external modulator or lens may be aligned to provide suitable feedback. These latter two cases may, however, be less practical owing to the need for alignment to perform their primary function which in turn may not be optimum for producing a stable line narrowed feedback state. A novel idea, with a number of advantages, is the use of the end facet of a length of optical fibre. This was first proposed by Favre et al[1,3,43].

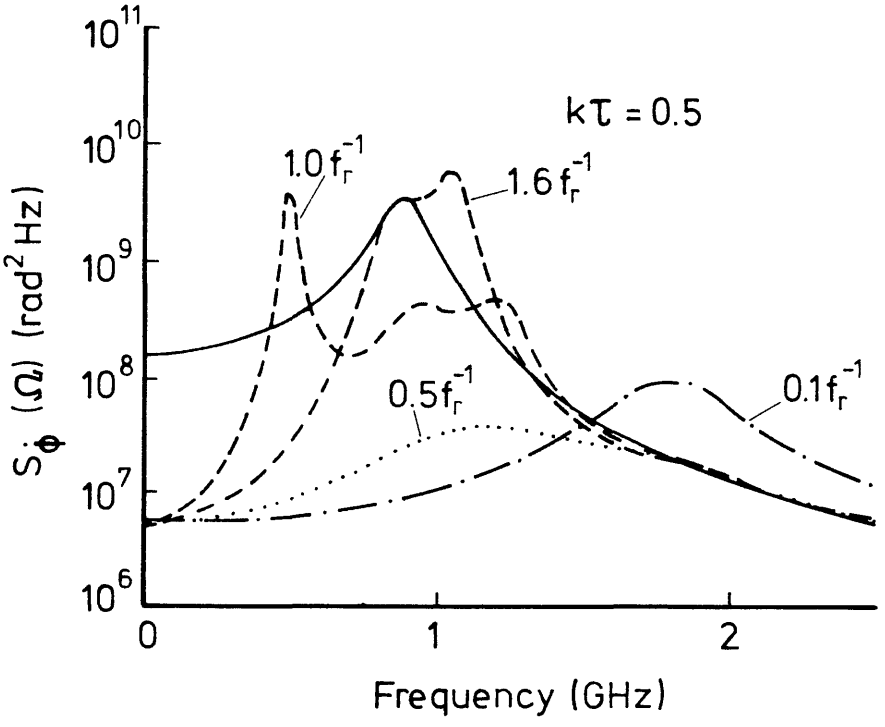


Figure 3.6 Instantaneous Frequency Spectra of an ITT LS709 Laser Under Various Conditions of Optical Feedback [15].

$\tau = 1.6\ f_r^{-1}$ (– – –), $\tau = f_r^{-1}$ (– –), $\tau = 0.5\ f_r^{-1}$ (·····), $\tau = 0.1\ f_r^{-1}$ (— · —). The continuous curve refers to the absence of feedback.

Although this initially appears to be a four facet problem (Figure 3.7), an approximation to a three facet model can be made if the input fibre facet is very close (<1mm) to the laser. The round trip time of this cavity is then very much less than that of r_{f2} and hence, if the k value for the two facets are the same ([3.28]), the dominant cavity will be the longer one. This situation arises naturally since butt coupling the fibre to the laser is the normal coupling process. An advantage of using fibre as the cavity is that its optical length can be very much greater than the physical length of the cavity mount, since the fibre can be coiled up into a relatively small volume. This technique and its merits will be discussed further in chapter 4.

All of the above implementations suffer to a greater or lesser degree from instabilities induced by mechanical, acoustical, or thermal disturbances. Generally speaking the more compact the cavity then the less prone a laser system is to these perturbations. As a result, much research has been done on short cavity systems. Lee et al[44] coupled a 200μm long graded index (GRIN) rod to a conventional 250μm InGaAsP laser chip and obtained a stable single mode output with a linewidth power product of 15MHz.mW, substantially narrower than that of the solitary laser. A similar technique was also attempted by the same group[22] with a distributed feedback (DFB) laser (section 3.3.3). A 4mm long GRIN rod with an AR coating on the near facet and gold on the far facet to increase its reflection coefficient was coupled to the laser (Figure 3.8). The lens was mounted on a PZT device and the amount of feedback was controlled by focusing and defocusing the rod. This resulted in a minimum linewidth power product of 2MHz.mW at 10% feedback power. Modulation characteristics were measured as 200MHz/mA at a modulation frequency of 565MHz.

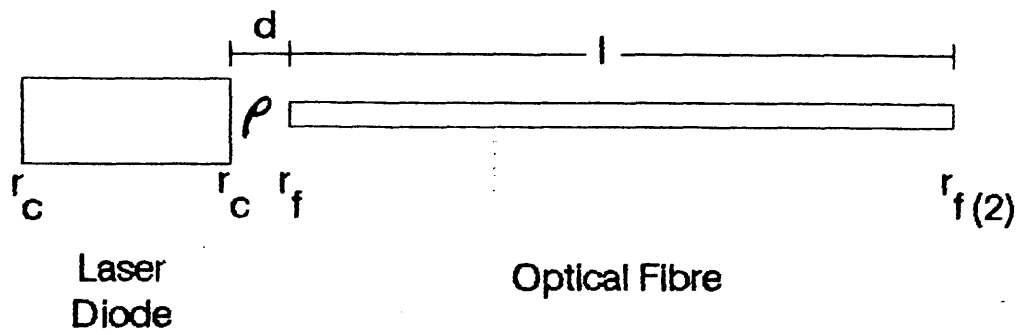


Figure 3.7

Schematic of Four Facet

Fibre Cavity Laser Module.

ρ = Coupling Coefficient

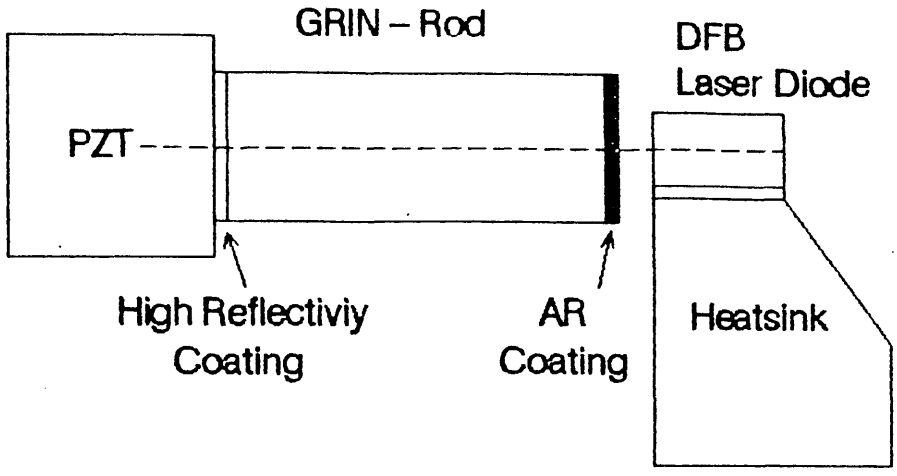


Figure 3.8
A DFB – GRIN Coupled Cavity Laser

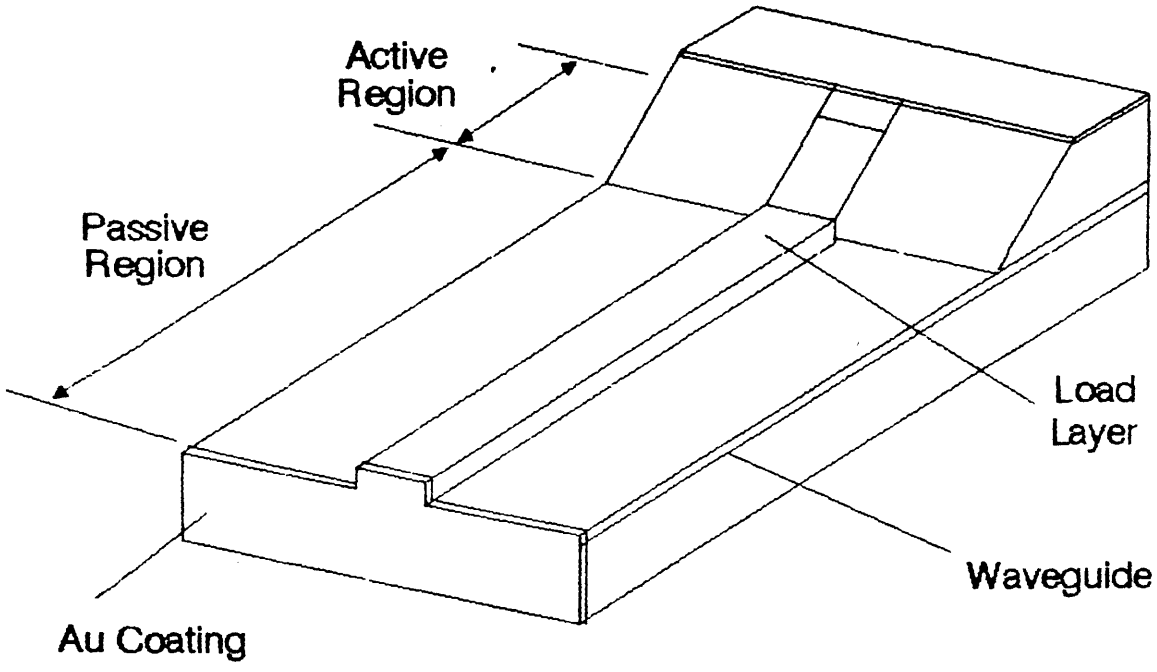


Figure 3.9. An Integrated Passive Cavity Laser.

Taken to its logical conclusion, this work leads to an integrated active and passive cavity on the same semiconductor chip. Such a structure returns the optimum performance in terms of susceptibility to thermal, mechanical, and acoustic transients. The first reported construction of such a structure was by Fujita et al[21]. A 1.5mm long waveguide was formed adjacent to the active region of the device (Figure 3.9). The cleaved facet of the waveguide served as the external mirror and was coated with gold for high reflectivity. A minimum linewidth of 900kHz at 6mW output power was achieved giving a linewidth power product of 5.6MHz.mW.

The final structure which qualifies as a coupled cavity laser is the cleaved coupled cavity or C^3 laser[30]. This consists of two cleaved active regions on the same chip separated by some gap, l. Results from these structures have produced linewidth power products of 3MHz.mW. However, the structure is very sensitive to gap width, and, as a result, it is difficult to optimise and stabilise the output.

3.3.2. External Cavity Laser Structures.

An important group of semiconductor laser modules which have perhaps produced the most promising experimental results to date is the external cavity laser resonators. These differ from the coupled cavity structures in that the middle facet of the coupled cavity is anti reflection coated, resulting in only one resonant cavity, part active, part passive, determining the frequency of the laser mode. The first known report of such a structure was by Glasser[18]. Flemming and Mooradian[37] showed that the linewidth of the external cavity is reduced in inverse proportion to the square of the ratio of the photon time spent in the active section to the time spent in both sections:-

$$\delta f = \delta f_{sl} \left[\frac{L_a/V_{ga}}{L_a/V_{ga} + L_p/V_{gp}} \right]^2 \sim \delta f_{sl} \left[\frac{L_a^2 \, n_{ga}^2}{L_p^2} \right] \qquad \text{——[3.36]}$$

assuming $L_p \gg L_a$. The subscripts a and p refer to active and passive sections of the cavity. L_a and L_p are the respective lengths of the sections, and V_{gp} and V_{ga} are group velocities (Figure 3.10). Equation 3.36 has been regarded as valid for most values of L_p. However, recent work by Sato et al[29] has suggested that for large cavity lengths (> a few centimeters) the phase diffusion coefficient due to spontaneous emission becomes dominant in the linewidth calculation, and the linewidth becomes proportional to L_p^{-1}.

The most significant results from external cavity lasers have been produced by BTRL[4,19] and latterly by Bell Laboratories[45]. Wyatt et al [19] used a diffraction grating as the reflecting element in a 30cm long cavity which gave single mode operation and continuous tuning over 60nm. Fine wavelength adjustment was obtained using an intra-cavity etalon driven by a galvanometer. A minimum linewidth of 10kHz was achieved although acoustic disturbances of the cavity produced a jitter of up to 1.5MHz deviation at audio rates. Matthews et al[4] have since packaged this module on a silica plate and have recorded a long term stability for the unit of better than 150MHz peak drift in 8 hours of operation. Spectral purity and stability of these orders make this unit very useful for coherent systems experiments. Its major drawback apart from its physical size and low output power (~few hundred microwatts) is the need for an integrated optical intra-cavity modulator which would be required to obtain fast

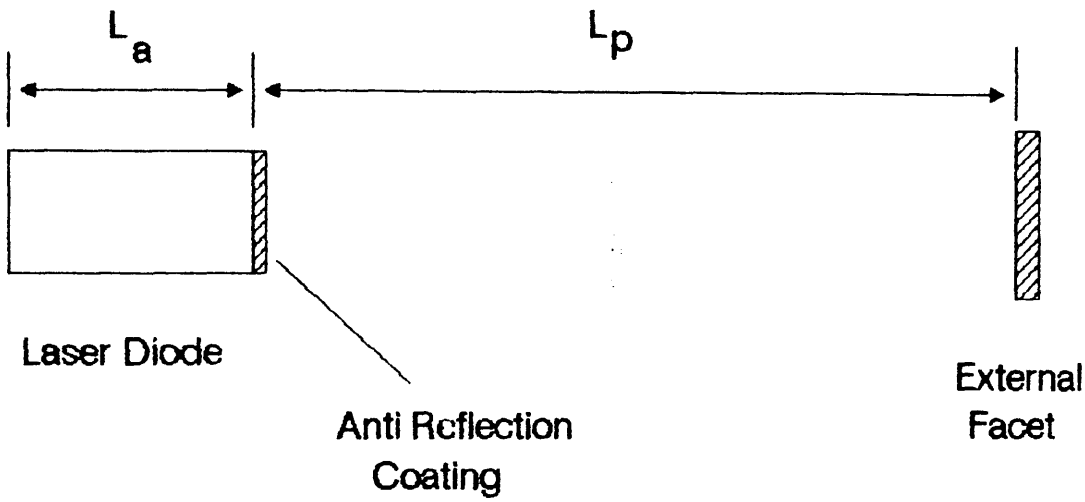


Figure 3.10. External Cavity Laser Schematic.

modulation rates. The introduction of such a device would increase cavity losses and add to mechanical instability by adding an extra two facets to the module. For these reasons it is unlikely that such a device in its current form would prove useful for commercial systems.

3.3.3. DFB and DBR Lasers.

Distributed feedback (Figure 3.11) and distributed Bragg reflector lasers (Figure 3.12) are considered in the long term to be the most promising sources for coherent communication systems. As a solitary entity they operate single mode with moderate tunability. Both configurations rely on coupling between epitaxial gratings and the active lasing medium to provide frequency selectivity and stability. However, present device technology is still only producing results of 44MHz.mW and 63MHz.mW at 1.3 and 1.5μm respectively, with minimum linewidths of 4 and 10MHz being achieved. This is still outside what will be shown to be useful for coherent communication systems. The limiting factor in the production of such devices is the grating definition and fabrication technology. Neither the definition quality nor the length of the grating is sufficient to produce the sub MHz linewidths which are required. The magnitude of the problem can be understood when the grating parameters are considered. This states that the phase of the grating must be consistent to within half a wavelength over its length. At present this can be achieved over a length of 200-300μm. However, to obtain sub-MHz linewidths it will be necessary to write gratings 1mm in length. This is not yet possible with current electron-beam lithographic technology. Therefore, DFB and DBR lasers are considered only suitable for a limited number of coherent experiments.

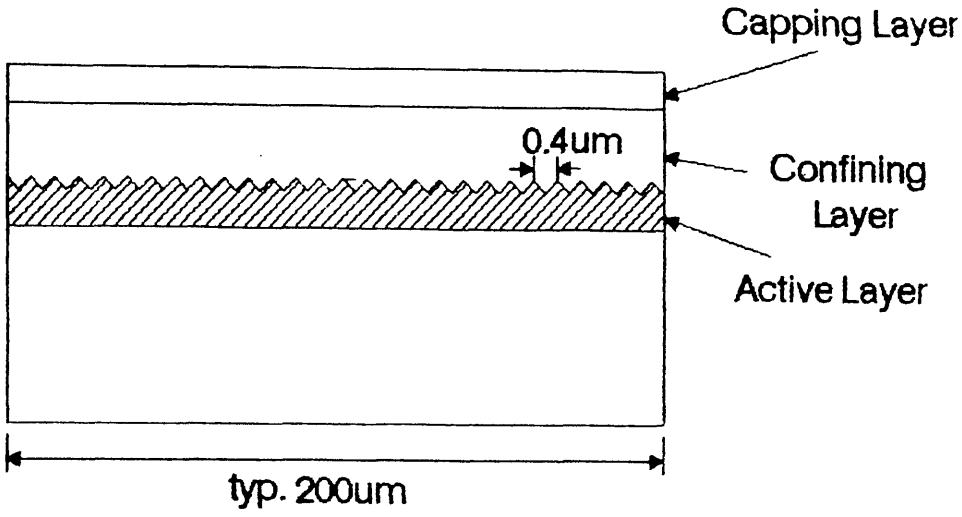


Figure 3.11. Schematic of a Basic DFB Laser.

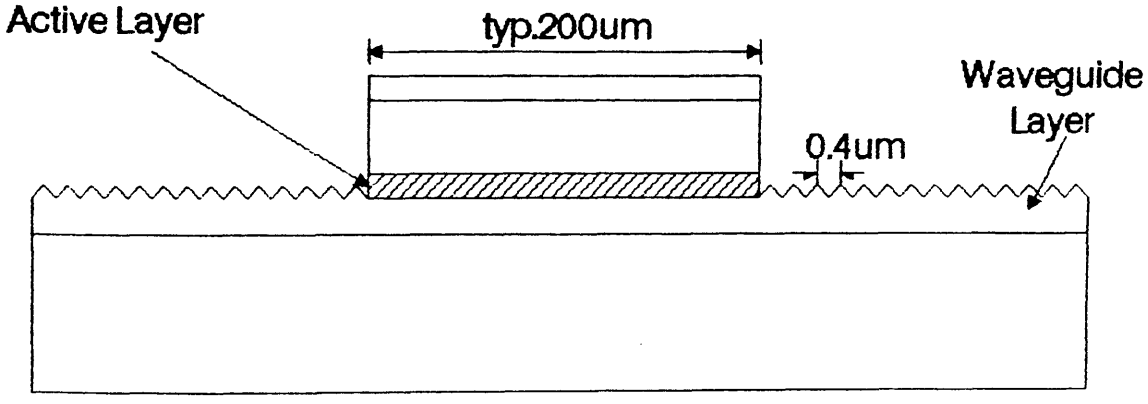


Figure 3.12. Schematic of a Large Optical Cavity DBR Laser.

3.4. Semiconductor Laser Systems: Discussion and Conclusions.

In this chapter consideration has been given to the suitability of a number of different types of lasers for coherent optical communication systems. It has been shown that, although gas lasers have a number of desirable characteristics, they also have certain disadvantages which make then unsuitable for use in these systems. Primarily, these disadvantages are in relation to their size and power supply requirements. Semiconductor lasers, on the other hand do not have these problems. Instead, the major drawback of these sources are the noise properties which they exhibit. These have been considered in some detail, with theoretical expressions being derived for the laser FWHM. In addition, peaks in the emission spectra, which occur as a result of spontaneous emissions in the laser cavity, have been discussed. It was concluded that, although far from ideal at present, these devices can be considered as most promising in the long term.

A number of measures to reduce the noise output of these devices were examined. These techniques for linewidth reduction included partially reflecting mirror cavities, and external grating cavity modules. It was shown that the reduction in linewidth was related in general to the coupled/external cavity length, and the external facet reflectivity, subject to certain conditions being satisfied. DFB and DBR lasers were also considered. However, current technology is capable of producing devices which would be useful for only a minority of coherent communications experiments. The expected linewidth reduction from several relevant techniques were calculated and some experimental results from other workers reported. The most successful technique reported to date is the external cavity laser using a bragg

diffraction grating[4,19]. This exhibited a linewidth of ~10kHz and was continuously tunable over ~60nm. However, implementation of this technique is complex and was not considered viable for this project. A simpler technique was the fibre cavity module. This was considered more promising and attempts to realise such a module will be discussed in chapter 4.

In conclusion, a number of free running laser modules and linewidth reduction techniques have been examined. At present, non have shown themselves to be ideal for use in commercial coherent optical communication systems. Only the distributed feedback lasers, and distributed Bragg reflector lasers show promise of eventually fulfilling this function.

References.

1. F.Favre, D.LeGuen 'Emission Frequency Stability in Single Mode Fibre Optical Feedback Controlled Semiconductor Lasers' Elect. Lett. 19(17) 1983. pp663-665.

2. F.Favre, D.LeGuen 'Autostabilisation Technique for Achieving Highly Stable Resonant Optical Feedback in a Fibre Resonator Loaded Injection Laser' Elect. Lett. 19(24) 1983. pp1046-1048.

3. F.Favre, D.LeGuen, J.C.Simon 'Optical Feedback Effects Upon Laser Diode Oscillation Field Spectrum' IEEE J. Quant. Elect. vol QE-18(10) 1982. pp1712-1717.

4. M.R.Matthews, K.H.Cameron, R.Wyatt, W.J.Devlin 'Packaged Frequency-Stable Tunable 20kHz Linewidth 1.5μm InGaAsP External Cavity Laser' Elect. Lett. 21(3) 1985. pp113-115.

5. C.H.Henry 'Theory of Linewidth of Semiconductor Lasers' IEE J. Quant. Elect. vol QE-18(2) 1982. pp259-264.

6. C.H.Henry 'Theory of the Phase Noise and Power Spectrum of a Single Mode Injection Laser' IEEE J. Quant. Elect. vol QE-19(9) 1983. pp 1391- 1397.

7. B.Daino, P.Spano, M.Tamburrini, S.Piazzolla 'Phase Noise and Spectral Line Shape in Semiconductor lasers' IEEE J. Quant. Elect. vol QE-19(3) 1983. pp266-270.

8. P.Spano, S.Piazzolla, M.Tamburrini 'Phase Noise in Semiconductor Lasers: A Theoretical Approach' IEEE J. Quant. ELect. vol QE-19(7) 1983. pp 1195-1199.

9. R.Schimpe, W.Harth 'Theory of FM Noise of Single Mode Injection Lasers' Elect. Lett. 19(4) 1983. pp136-137.

10. K.Vahala, C.H.Harder, A.Yariv 'Observation of Relaxation Resonance Effects in the Field Spectrum of Semiconductor Lasers' Appl. Phys. Lett. 42(3) 1983. pp211-213.

11. K.Vahala, A.Yariv 'Semiclassical Theory of Noise in Semiconductor Lasers Part II' IEEE J. Quant. Elect. vol QE-19 1983. pp1102-1109.

12. K.Kikuchi, T.Okoshi 'Estimation of Linewidth Enhancement Factor of AlGaAs Lasers by Correlation Measurements Between FM and AM Noises' IEEE J. Quant. Elect. vol QE-21(6) 1985. pp669-673.

13. C.H.Henry 'Phase Noise in Semiconductor Lasers' J. Lightwave. Technol. vol LT-4(3) 1986. pp298-311.

14. H.Sato, T.Fujita, J.Ohya 'Theoretical Analysis of Longitudinal Mode Coupling in External Cavity Semiconductor Lasers' IEEE J. Quant. Elect. vol QE-21(4) 1985. pp284-291.

15. P.Spano, S.Piazzolla, M.Tamburrini 'Theory of Noise in Semiconductor Lasers in the Presence of Optical Feedback' IEEE J. Quant. Elect. vol QE-20(4) 1984. pp350-357.

16. T.P.Lee 'Linewidth of Single-Frequency Semiconductor Lasers for Coherent Lightwave Communications' Invited Paper, 11th European Conference on Optical Communications, Venice, Italy, 1985.

17. G.P.Agrawal 'Line Narrowing in a Single-Mode Injection Laser Due to External Optical Feedback' IEEE J. Quant. Elect. vol QE-20(5) 1984. pp468-471.

18. L.A.Glasser 'A Linearized Theory for the Diode Laser in an External Cavity' IEEE J. Quant. Elect. vol QE-16(5) 1980. pp525-530.

19. R.Wyatt, W.J.Devlin '10kHz Linewidth 1.5μm InGaAsP External Cavity Laser with 55nm Tuning Range' Elect. Lett. 19() 1983. pp110-112.

20. T.P.Lee, C.A.Burrus, D.P.Wilt 'Measured Spectral Linewidth of Variable Gap Cleaved Coupled Cavity Lasers' Elect. Lett. 21(2) 1985. pp53-54.

21. T.Fujita et al 'Narrow Spectral Linewidth Characteristics of Monolithic Integrated Passive Cavity InGaAsP/InP Semiconductor Lasers' Elect. Lett. 21(9) 1985. pp374-376.

22. T.P.Lee et al 'Characteristics of Linewidth Narrowing of a 1.5μm DFB Laser with a Short Grin-Rod External Coupled Cavity' Elect. Lett. 21(15) 1985. pp655-656.

23. G.Burns, I.Nathan 'P-N Junction Lasers' Proc. IEEE July 1964. pp770-794.

24. A.Yariv 'An Introduction to Optical Communications' 2nd Ed, Holt, Reinhart, and Winston 1976.

25. 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. Quant. Elect. vol QE-20(5) 1984. pp478-491.

26. K.Kikuchi, T.Okoshi 'Simple Formula Giving Spectrum-Narrowing Ratio of Semiconductor-Laser Output Obtained By Optical Feedback' Elect. Lett. 18(1) 1982. pp10-11.

27. E.Patzak et al 'Spectral Linewidth Reduction in Semiconductor Lasers by an External Cavity with Weak Optical Feedback' Elect. Lett. 19(22) 1983. pp938-940.

28. M.Tamburrini, P.Spano, S.Piazzolla 'Influence of an External Cavity on Semiconductor Laser Phase Noise' Appl. Phys. Lett. vol 43(5) 1983. pp410-412.

29. H.Sato, J.Ohya 'Theory of Spectral Linewidth of External Cavity Semiconductor Lasers' IEEE J. Quant. Elect. vol QE-22(7) 1986. pp1060-1063.

30. S.Saito, O.Nilsson, Y.Yamamoto ' Oscillation Centre Frequency Tuning, Quantum FM Noise, and Direct Frequency Modulation Characteristics in External Grating Loaded Semiconductor

Lasers' IEEE J. Quant. Elect. vol QE-18(6) 1982. pp961-970.

31. C.J.Nielsen, J.H.Osmundsen 'Linewidth Stabilisation of Semiconductor Lasers in an External Cavity' J. Opt. Commun. vol 5(2) 1984. pp42-45.

32. C.H.Henry 'Theory of Spontaneous Emission Noise in Open Resonators and its Application to Lasers and Optical Amplifiers' J. Lightwave. Technol. vol LT-4(3). 1986. pp288-297.

33. A.L.Schawlow, C.H.Townes 'Infrared and Optical Masers' Phys. Rev. vol 112, 1958.pp1940-1949

34. M.W.Flemming, A.Mooradian 'Fundamental Line Broadening of Single Mode (GaAl)As Diode Lasers' Appl. Phys. lett. 38(7) 1981. pp511-513.

35. M.Lax 'Classical Noise V. Noise in Self Sustaining Oscillators' Phys. Rev. vol 160 1967. pp290-307.

36. R.Lang, K.Kobayashi 'External Optical Feedback Effects on Semiconductor Injection Laser Properties' IEEE J. Quant. Elect. vol QE-16(3) 1980. pp347-355.

37. M.W.Flemming, A.Mooradian 'Spectral Characteristics of External Cavity Controlled Semiconductor Lasers' IEEE J. Quant. Elect. vol QE-17(1) 1981. pp44-59.

38. K.E.Stubkjaer, M.B.Small 'Noise Properties of Semiconductor

Lasers Due to Optical Feedback' IEEE J. Quant. Elect. vol QE-20(5) 1984. pp472-478.

39. L.Goldberg et al 'Spectral Characteristics of Semiconductor Lasers with Optical Feedback' IEEE J. Quant. Elect. vol QE-18(4) 1982. pp555-564.

40. L.Goldberg et al 'Noise Characteristics in Line-Narrowed Semiconductor Lasers with Optical Feedback' Elect. Lett. 17(19) 1981. pp677-678.

41. D.Lenstra, B.H.Verbeek, A.J.den Boef 'Coherence collapse in Single Mode Semiconductor Lasers Due to Optical Feedback' IEEE J. Quant. Elect. vol QE-21(6) 1985. pp674-679.

42. C.H.Henry, R.F.Kazarinov 'Instability of Semiconductor Lasers Due to Optical Feedback from Distant Reflectors' IEEE J. Quant. Elect. vol QE-22(2) 1986. pp294-301.

43. F.Favre, D.LeGuen 'Effect of Semiconductor Laser Phase Noise on BER Performance in an Optical DPSK Heterodyne-Type Experiment' Elect. Lett. 18(22) 1982. pp964-965.

44. T.P.Lee et al 'Measured Spectral Linewidth of Single Frequency 1.3 and 1.5μm Injection Lasers' Elect. Lett. 20 1984. pp1011-1012.

45. R.A.Linke et al 'Coherent Lightwave Transmission Over 150km Fibre Lengths at 400MBit/s and 1GBit/s Data Rates Using DPSK Modulation' Post Deadline Paper, 11th European Conference on

Optical Communications, Venice Italy, 1985.

CHAPTER 4.

EXPERIMENTATION WITH SEMICONDUCTOR LASER MODULES.

4.1. Introduction.

Early in the project the decision was taken that experimental work would be based on the use of 825nm GaAlAs semiconductor lasers, as opposed to 1.5μm GaInAsP devices[1,2,3]. There were three major reasons for this decision. Firstly, 'stable' single mode devices at that wavelength were available commercially in pairs guaranteed matched to within 6nm (In practice the match was much closer). This reduced the amount of temperature-induced frequency pulling required to obtain a beat between the two sources (Section 4.2.2). It was also hoped that the use of devices which were inherently single mode (unlike currently available devices at the other wavelengths) would reduce the development time for a suitable source useful for coherent systems experiments. Finally, experience of working with these devices was already available in the department, and indeed limited success at this wavelength had already been achieved[4].

This chapter will deal with the characterisation of the Hitachi HLP 1400 semiconductor laser. Measurements of longitudinal mode and modulation characteristics will be reported. The measurement of the solitary laser linewidth, and the technique used to make the measurement will also be discussed. Attempts to construct a fibre-coupled cavity laser module will be described and results discussed. A description of a mirror-coupled cavity module which was constructed will also be given. The chapter will conclude with a discussion of the results.

4.2. Characteristics of the Hitachi HLP 1400 Laser Diode.

4.2.1. Device Description.

The Hitachi HLP 1400 is a channel substrate planar (CSP) GaAlAs device. A schematic drawing of its structure is shown in Figure 4.1. The laser chip is approximately 300μm (length) x 200μm (width)x 100μm (height) in size. It is mounted p-side down on an interface block (the heatsink) which, in turn, is mounted on a brass stud (Figure 4.2). One facet of the laser is aligned flush with the heatsink/stud in order to facilitate optical coupling. The other facet is less accessible since the heatsink oversteps the laser chip (Figure 4.3). Bonding wires from the n-type contact are taken to a stress relief block mounted next to the laser chip before carrying on to the main electrical contact point on the brass stud. As these devices are particularly sensitive to static electricity, the electrical contacts are shorted together before they leave the factory. Once removed from their packaging particular care is required in handling, not only to prevent damage due to static, but also to avoid mechanical disruption of the exposed chip. Assuming that the heatsink is earthed, the mounting of the chip p-side down requires the use of a current sink supply. Attention must then be paid to the polarity of any modulating drive source to ensure that it is compatible with the device. In this respect it must be able to operate with a DC offset of -1.4V which is present across the diode.

4.2.2. Modulation and Frequency Tuning Characteristics.

In addition to the two lasers which were used in Steele's project[5], two unused lasers were inherited from him, devices no. 2M0861 and 2M0862. These four devices formed the initial pool of test sources. Four more were later purchased from Hitachi (devices

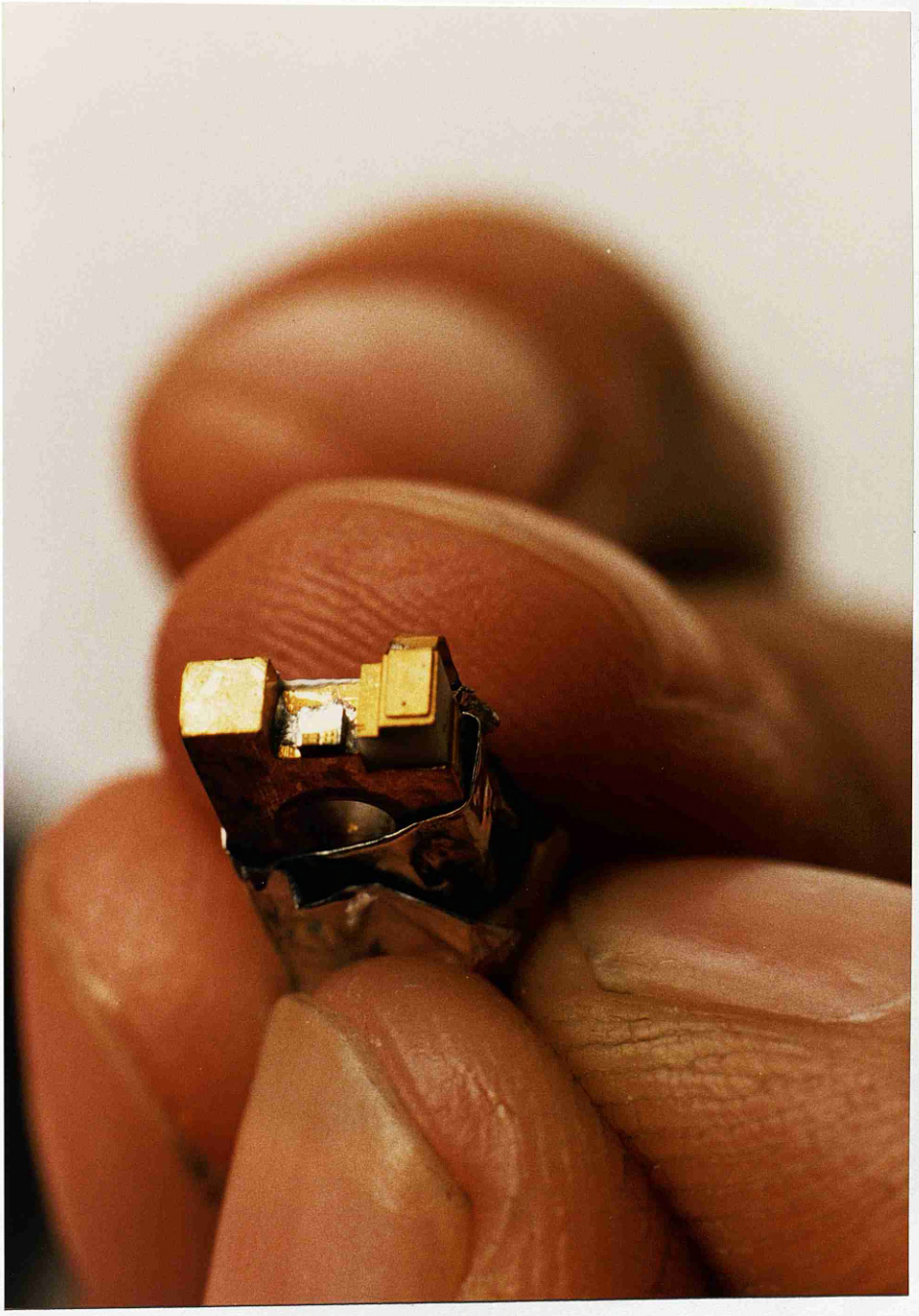


Figure 4.2. The HLP 1400 as Received From the Manufacturer.

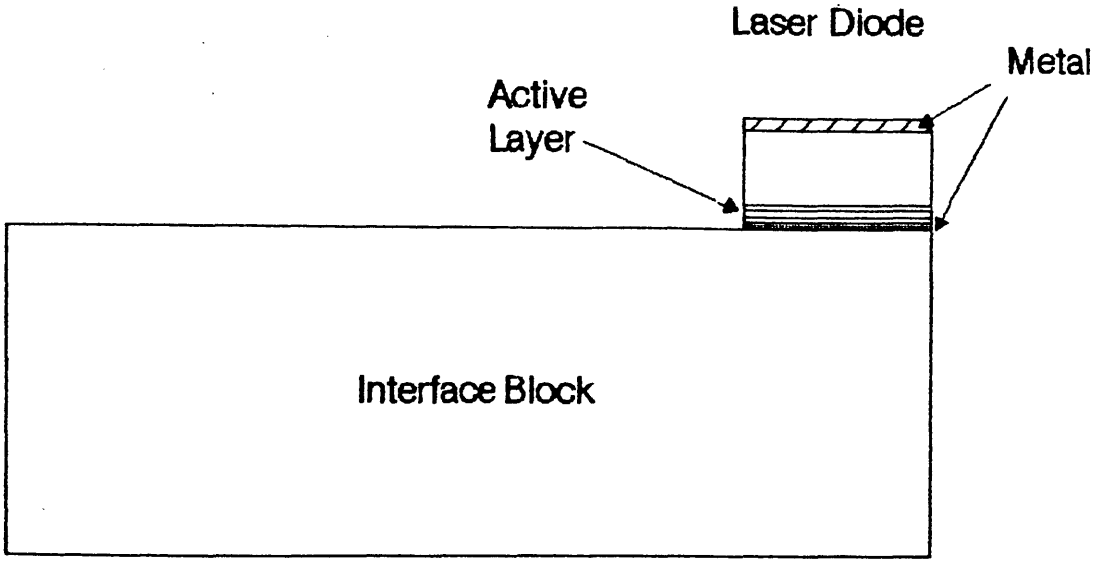


Figure 4.3. Relative Positioning of Laser Chip and Interface Block on an HLP 1400 Laser Diode.

No. 5E3060, 5E3070, 5E3072, and 5E3073). All were found to be single mode over most of their tuning range with nominal emission wavelength of ~825nm at 20°C. Threshold currents (I_{th}) were all around 65mA, while the maximum drive current applied was 110mA ($1.7I_{th}$).

The small signal modulation response (injection current-to-FM transfer function) of the diodes varies little with dc bias current. It comprises of two components, the thermal modulation, and the carrier density modulation of the diode[6,7]. At modulation frequencies below 10MHz there is an enhanced FM response due to thermal modulation of the active layer region of the laser. At higher frequencies, carrier density effects dominate. The theoretical analysis of the modulation response of this and other types of lasers (buried heterostructure (BH), and crank transverse junction stripe (TJS)) are discussed by Kobayashi et al[6]. This section will present a short description of these effects in order to extract from theory a number of fundamental values important to this project.

The thermal response of the diode chip can be analysed using thermodynamic diffusion equations for the laser diode chip and submount structure. Hence, the physical structure of the device, including the proximity of the active region to the mounting stud, has an important effect on the diode modulation response. The transient thermal response time of a typical laser diode structure is around 0.6μs[6]. Therefore, as has been stated, the contribution of thermal effects to the modulation response above a few megahertz is not significant. Both ohmic heating of the diode chip and heat generation in the active layer are responsible in roughly equal parts for modulating the temperature in the active region. This temperature modulation can be expressed in terms of changes

in frequency by[7]:-

$$\frac{1}{f}\frac{df}{dT} = -\frac{1}{L}\frac{dL}{dT} - \frac{1}{n}\frac{dn}{dT} \qquad \text{----[4.1]}$$

where f is the cavity mode frequency, L is the cavity length, n the refractive index of the region, and T the temperature. The first term in the right hand side is a linear thermal expansion coefficient, which for GaAs has a value[6] of $0.6\text{x}10^{-5}$ K^{-1}. At 825nm, this results in a frequency temperature coefficient of -2.18 GHz/K. The second term is a thermal refractive index coefficient which takes a value of $0.99\text{x}10^{-4}$ K^{-1} giving, at 825nm, a frequency temperature coefficient of -36GHz/K. Combining these two terms gives a total DC thermal response of -38.18GHz/K. This is slightly higher than the average measured value for an HLP 1400 of -22 GHz/K.

Carrier density effects are considerably more complicated to analyse. Determination of the effect of small perturbations on the rate equations for the photon density and carrier (electron) density must be undertaken. A detailed analysis of this process is given by Kobayashi et al[6]. In brief, the carrier density dependence of the refractive index is related to the FM response by[7]:-

$$\frac{df}{dN} = -\frac{f}{N}\frac{dn}{dN} \qquad \text{----[4.2]}$$

where, in addition to the terms previously defined, N is the carrier number density. There are two contributions to dn/dN[7]: a free-carrier plasma dispersion term of approximately $-10^{-21}cm^{-3}$,

and a contribution from anomalous dispersion of the band edge absorption which shifts with carrier density at approximately $-10^{20}cm^{-3}$. Together, these produce a flat response up to about a gigahertz when relaxation oscillation effects result in a peak in the response and an abrupt phase change of 180^{o}.

Measurements of the combined effect of these two mechanisms by both Kobayashi et al[6] and Welford and Alexander[7] have been verified by experimental results within this project[8](Figure 4.4). These measurements were performed on device 5E3070 using a scanning Fabry Perot Interferometer. The operation of this instrument will be described in section 4.2.3. Unfortunately, the available equipment allowed for measurements only up to 900MHz where the thermal response dominates. The relaxation resonance effects, which the carrier density response exhibits, were not observed.

A value for the dc injection current coefficient can also be obtained from figure 4.4. This was found to be 4GHz/mA which is in close agreement with figures published elsewhere[7,9,10].

The marked roll-off and phase shift which the diode exhibits must be carefully considered when OPLL systems are being designed. It is quite possible that the loop bandwidth may be small enough to allow approximation of the response to the dc value without undue penalty. Alternatively, in the higher bandwidth cases,some form of equalization may be employed[8].

4.2.3. Modal Characteristics.

Before use in systems experiments, the modal characteristics of each device was required to be measured. This experiment (Figure 4.5) involved splitting the output of the test device using a beam splitter, coupling one portion of the beam to a

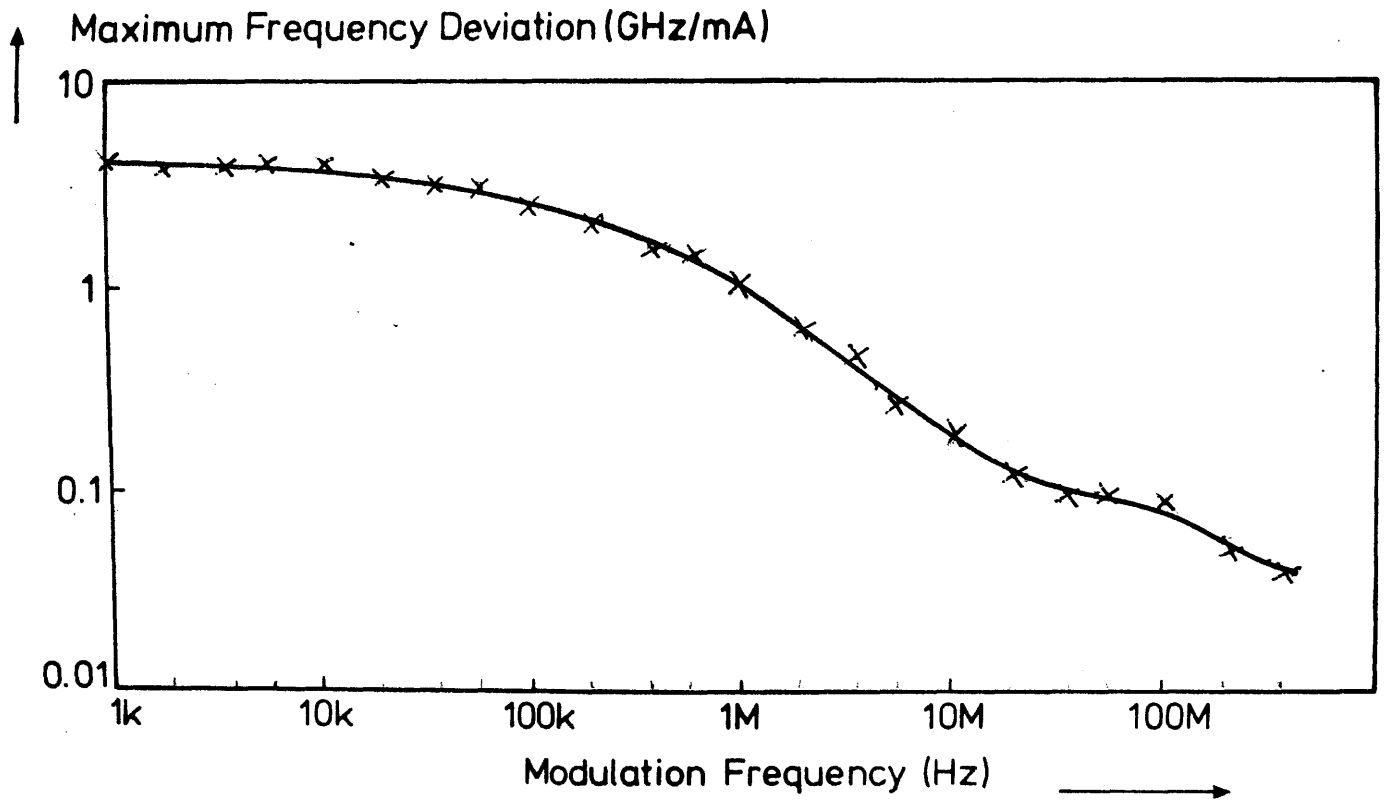


Figure 4.4. Maximum Frequency Deviation Versus Modulation Frequency for an HLP 1400 Laser Diode (No. 5E3070)[8]

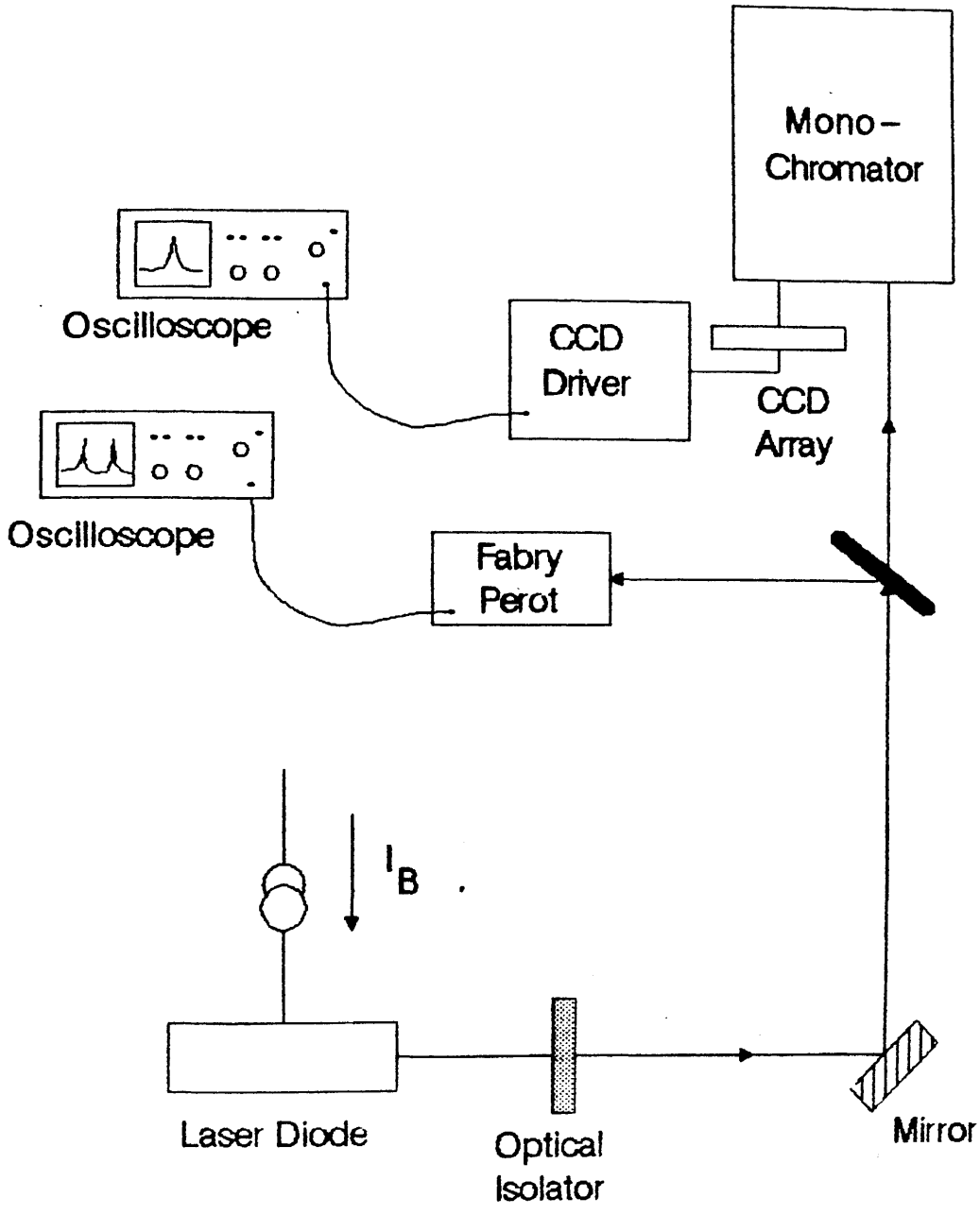


Figure 4.5. Experiment to Measure Absolute Wavelength and Modal Stability of a Test Source.

scanning Fabry Perot spectrum analyser and the other into the monochromator to obtain a measurement of absolute wavelength. These measurements were made with the assistance of a final year project student. Mode plots for three devices are shown in Figures 4.6-4.8. Only one of the devices (5E3072) did not exhibit the phenomenon of mode hopping when tuned across its drive range. Mode hopping involves the discontinuous shift of the output wavelength from one longitudinal mode to another as the device is tuned in frequency. This is normally accompanied by hysteresis in the tuning characteristics. The Hitachi devices that exhibited this did so to varying degrees. As can be seen, device 5E3070 is particularly unstable. Discussions with other workers would indicate that this rate of mode hopping is unusual. It is possible that the way in which the laser chip is mounted results in reflections off the heatsink inducing weak feedback effects which in turn cause this instability.

The scanning Fabry Perot is a spherical mirror Fabry Perot interferometer[11]. Two models were acquired during the course of the project. These were the Tech Optics models SA-2 and SA-300, the former being the most commonly used. They had cavity lengths of 37.5mm and 250mm and free spectral ranges (FSR's) of 2GHz and 300MHz respectively. The mirrors were confocal with a reflection peak at 820nm. One of the mirrors was mounted on a stack of four PZT ceramics giving a scan range of 0.8μm or approximately one wavelength. The maximum finesse of the cavity was 175, giving an element resolution for SA-2 of ~11MHz. This was not sufficient for direct linewidth measurement, but was ideal for examining the mode structure of a laser emission. Moreover, the instrument was very simple to use. It was therefore utilised in this experiment to allow the experimenter to confirm that the laser was operating

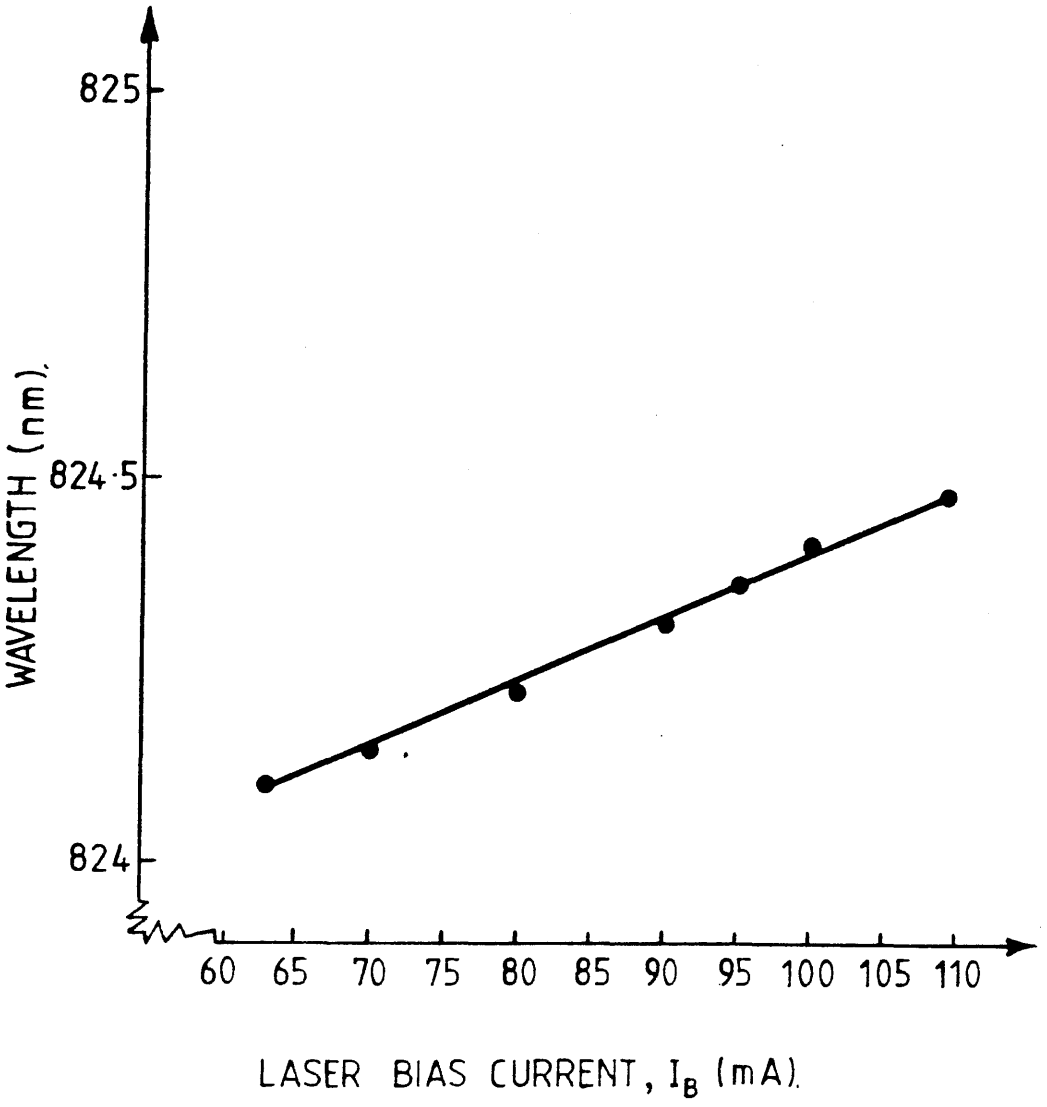


Figure 4.6. Injection Current Versus Wavelength for HLP 1400 No. 5E3072. Operating Temperature = 293.0K.

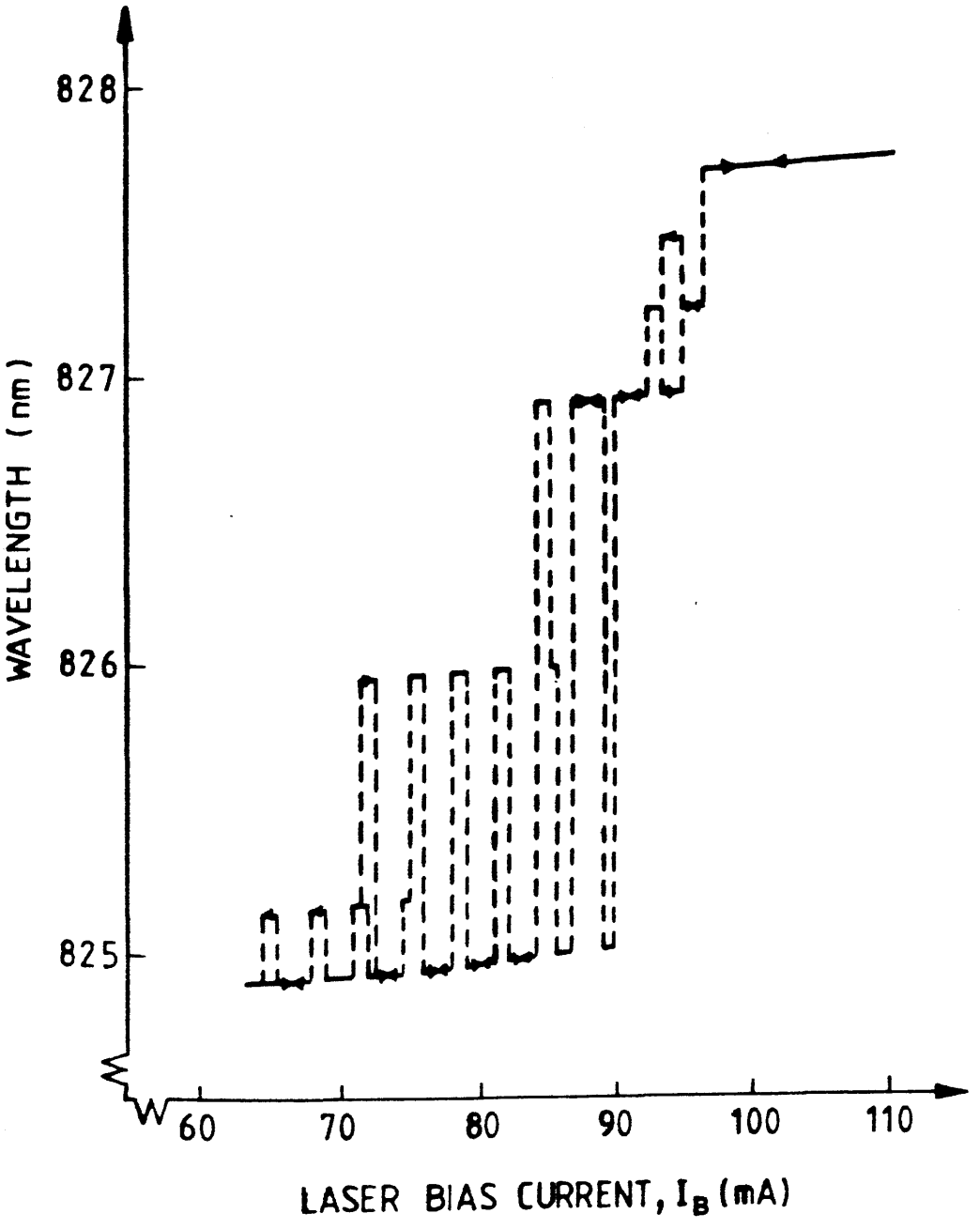


Figure 4.7. Injection Current Versus Wavelength for HLP 1400 No. 5E3070. Operating Temperature = 290.8K.

single mode. The SA-300 had a theoretical resolution of ~2MHz. This figure would have been marginal to measure the free running (unaffected by feedback) laser linewidth. However, because of the length of the cavity, it was very difficult to align accurately and, therefore, a finesse of 175 was very difficult to achieve. Its difficulty in use also made it a secondary choice for monitoring modal stability in this and other experiments.

Once single mode operation was confirmed the absolute wavelength of the line could be measured by the monochromator. The instrument has a total path length of 2m, in the middle of which is placed a Bragg grating with a line density of 1200 lines/mm. The angle of reflection of light from this grating is proportional to wavelength. Thus by rotating the grating, the wavelength selected at the output can be obtained to a resolution of 0.1 Ångstroms. The wavelength is then displayed on a graduated scale on the side of the instrument. When measuring modal characteristics, the output slit was opened wide to allow a 5 Ångstrom band of light to be focused on a CCD array. The output from this array was then displayed on an oscilloscope and a real time spectrum of the test source was observed. By varying the drive current to the laser, and calibrating the display against the monochromator scale, the variation of wavelength with current could be plotted.

Figures 4.6 to 4.8 show typical outputs of devices 5E3072, 5E3070, and 2M0861. From the slope of Figure 4.6, a frequency current coefficient of 5.56GHz/mA was measured. This is in close agreement with the calculated value in section 4.2.2.3. Both the wavelength of individual mode hops and the number of mode hops exhibited by each device varied considerably with temperature and humidity. As a result, day to day measurements were not found to

be consistent even when attempts were made to maintain constant temperature. At no time, however, was 5E3072 found to mode hop except at the very extremities of the drive range.

The temperature control system was implemented around a BBC microcomputer. The temperature sensing element was a platinum resistor in a bridge configuration, mounted below the laser stud, and thermal adjustment was made through Peltier thermoelectric elements. The controller algorithm was a simple three term controller written in Basic. Through observations on the Fabry Perot, control to better than $0.05^{o}K$ was obtained.

From this experiment it was concluded that device 5E3072 would be the most suitable choice for a local oscillator (LO) in any forthcoming systems experiments. The transmit laser would be chosen as that one which lased closest to the LO on the day of the experiment. This situation was not satisfactory but was precipitated by the lack of a reliable spectrally simple source.

4.3. The Self Heterodyne Experiment.

Although modal stability of the sources was of concern at the start of the project, the absolute linewidths of the lasers were considered to be the major problem in achieving a working phase locked loop. It was postulated in reference 4 and subsequently confirmed theoretically[12,13,14] that phase locking of sources without considering the bit error rate of the system would require beat linewidths (the sum of individual linewidths) of less than 10MHz. The attainment of a 10^{-9} BER in an optical link, as discussed in chapter 2, would probably require sub-MHz linewidths. To that end it was necessary to establish an experimental technique to measure the linewidths of the free running sources and to evaluate the effectiveness of any subsequent line narrowing

measures that might be taken. A suitable scheme was proposed by Okoshi et al[15]. This has subsequently been called the self heterodyne technique.

4.3.1. Principal of Operation.

It is possible to make high resolution spectral measurements of two optical sources by heterodyning the pair and observing the beat spectra. However, it is not a trivial task to obtain a stable local oscillator on an arbitrary wavelength. The Okoshi method relies on splitting the output of the test source and using one portion as the local oscillator. This can then be made incoherent with the second branch by delaying it for a time, τ_D, which is longer than the coherence time, t_{coh}, of the laser. The second branch is then shifted by some frequency, f_s, which is greater than the FWHM of the line. The two portions of the beam are then recombined and focused onto an avalanche photodiode (APD). The resultant beat is then centred on f_s and is proportional to the linewidth. Since both halves of the signal are equally noisy, the measured beat spectra is larger than the original. This proportionality factor is $2^{1/2}$ if the spectrum is Lorenzian, 2 if it is gaussian. In practice, a factor somewhere in between is usual since the spectrum comprises of 1/f noise up to phase fluctuations of 1MHz producing a gaussian component, and is Lorenzian beyond that (subject to the variations described in chapter 3)[16]. This results from acoustic and mechanical perturbation of experimental components. However, for the purposes of the foregoing measurements, a Lorenzian spectrum will be assumed.

4.3.2. Experimental Implementation.

The experiment was implemented as shown in Figure 4.9. An optical isolator, consisting of a 30dB polarizer and a $\lambda/4$ plate, was placed between the test source and a 50/50 beam splitter. The beam splitter was optically flat to $\lambda/10$. One branch of the beam was then coupled into 9.5km of single mode optical fibre using a standard X40 microscope objective. The fibre was step index fibre pulled for this project by British Telecom Research Laboratories, Ipswich. It had a loss of 2dB/km at 830nm and a cutoff wavelength of 700nm. It therefore supported only one mode at 830nm. A second 0.5km length of triangular profile fibre was also obtained from BT and was used initially to set up the experiment. This, however, was marginally dual moded and was eventually replaced. The delay time induced by the longer length was nominally 44μs. Favre et al[17] calculated that τ_D is required to be five times smaller than the coherence time of the laser in order that no coherence penalty is paid in the beat spectrum. Using this and [3.22], the resolution of the experiment can then be evaluated. The minimum linewidth that can be resolved in this instance is 36kHz. Another X40 objective was then used to couple light out of the fibre, through the recombining beam splitter and onto an APD.

The second branch from the source was passed through a second $\lambda/4$ plate to orientate the polarization vector into a plane which would maximise propagation through the frequency translation device. This was an Lithium Niobate acoustooptic Bragg cell. It was constructed in the department by A. Dawar and S. Al-Shukri and kindly donated for use in this experiment. It consisted of a planar waveguide, manufactured by proton exchange, overlaid with gold electrodes. These fingers, constructed to a standard design, were used to launch the acoustic wave across the device.

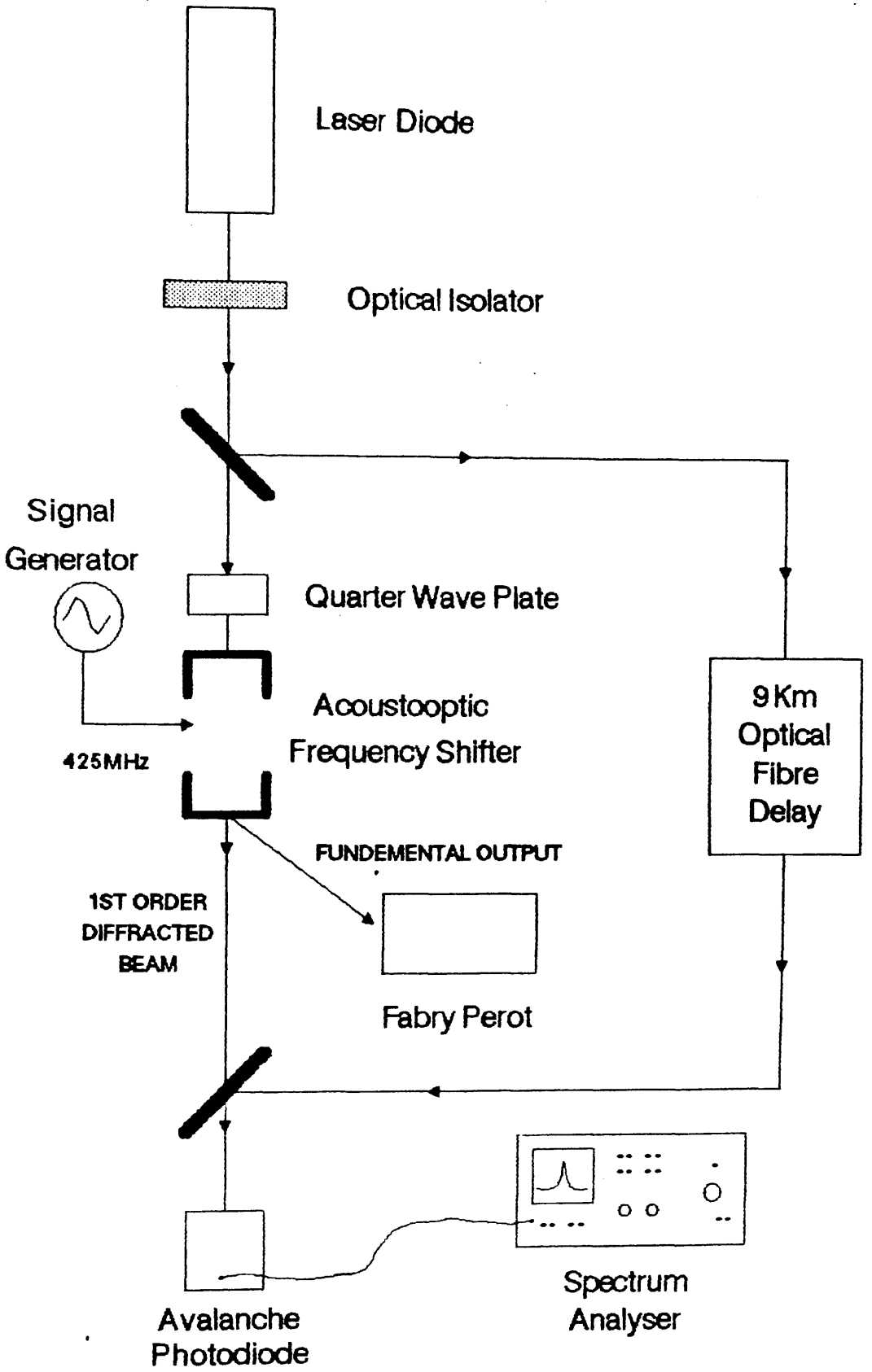


Figure 4.9. The Self Heterodyne Experiment.

Unfortunately, as is shown by network analysis (Figure 4.10a), the electrodes were not matched to 50Ω at its designed resonant frequency of ~425MHz. An inline inductance of 38nH was added (Figure 4.10b). However, an ideal (50 + j0)Ω load impedance was not achieved. The resultant mismatch later manifested itself as RF interference on the beat spectrum recorded by the spectrum analyser.

An RF drive level of +26dBm was considered to be the maximum safe level to avoid the destruction of the device. This resulted in a peak diffraction efficiency of 30% at 830nm being achieved. As was stated above, propagation in the device was very sensitive to the input polarization vector. In order to obtain the greatest degree of manipulative freedom the device was mounted on its side (Figure 4.11). As a result, the polarization vector had to be rotated into the horizontal for maximum efficiency and then back to circular before being combined with the fibre output. This was a lossy but necessary procedure. Rutile prisms were used to couple into and out of the device.

The output of the device was then collimated and directed through a 50/50 beam splitter where recombination took place. It should be noted that at this point power levels were frequently of the order of a microwatt. A further 3dB was then lost in the recombination process. The spatial alignment of the beams impinging on the APD was critical to obtain good mixing, and many hours were spent simply adjusting beam alignment at the beam splitter. This problem was exaserpated by drift in both the diffracted beam and in the fibre coupling optics which were not of the highest quality. However, with practice, results on a test source could eventually be obtained within a day.

Note should be made of the monitoring of the Bragg cell

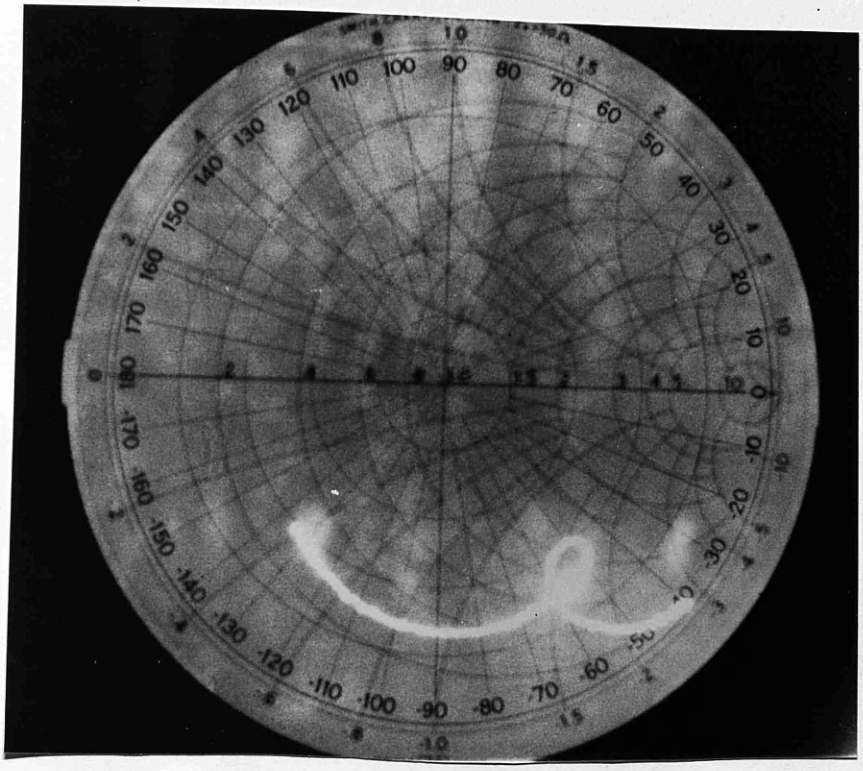


Figure 4.10a. Smith Chart Impedence Measurement of Unmatched Acoustooptic Bragg Cell.

Bright Spot Denotes Impedence at 425MHz.

Measurements Made on an HP 8410 Network Analyser, and an HP 8745 S – Parameter Test Set.

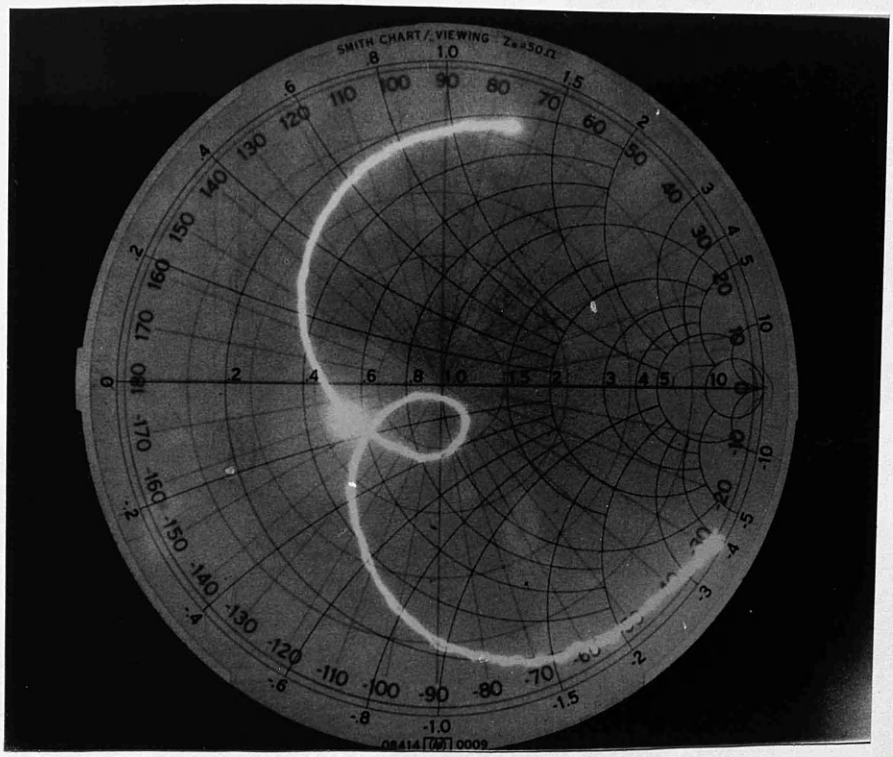


Figure 4.10b. Smith Chart Impedence Measurement of Matched Acoustooptic Bragg Cell.

Bright Spot Denotes Impedence at 425MHz.

Measurements Made on an HP 8410 Network Analyser, and an HP 8745 S – Parameter Test Set.

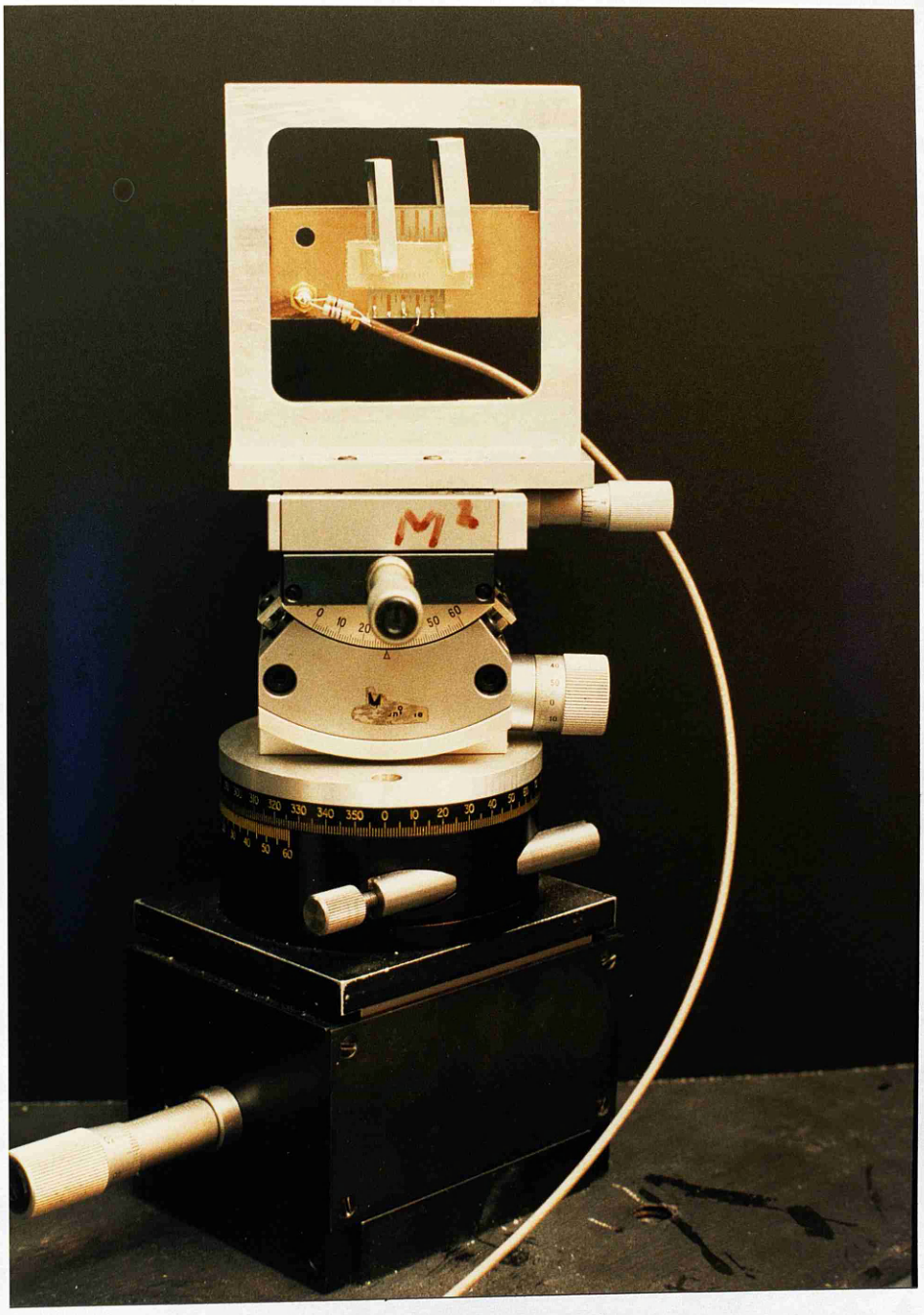


Figure 4.11. The Acoustooptic Bragg Cell And Mount.

fundamental output by the Fabry Perot. This was done in order that the limited input power available at the fibre and Bragg cell inputs was not wasted. Single mode operation could then be confirmed and extraneous external reflections from the fibre resulting from poor operation of the isolator could be eliminated through the use of this device.

4.3.3. Measurements of the Linewidth of Free Running Lasers.

Initial results with this experiment were poor and it was some months before reliable measurements were made. During this period, the length of 0.5km fibre was used to increase output power and therefore ease the mixing problems. This reduced experimental resolution to ~700kHz. However, at that stage it was not envisaged that linewidths of that order would be encountered. Measurements were made on the laser modules used by Steele[4] in his original heterodyne experiments. These modules were now four years old and observations on the Fabry Perot suggested that one of them in particular could be suffering from spectral ageing. This was confirmed by the results of the self heterodyne (Figure 4.12). The measured 3dB linewidth was estimated at ~50MHz, much larger than was predicted in chapter 3. Instabilities in the measured spectra was a major feature of these results. The main sources of instability were found to be:

a) Frequency Jitter of the test source. Observations on the Fabry Perot showed that test sources could exhibit frequency jitter in their emissions. This was caused by mechanical and thermal disturbances of the laser chip as a result of normal activities in an around the laboratory. Experiments late at night could show a reduction in jitter from a mean deviation of several hundred MHz to as low as 50MHz. A poorly designed power supply with a high

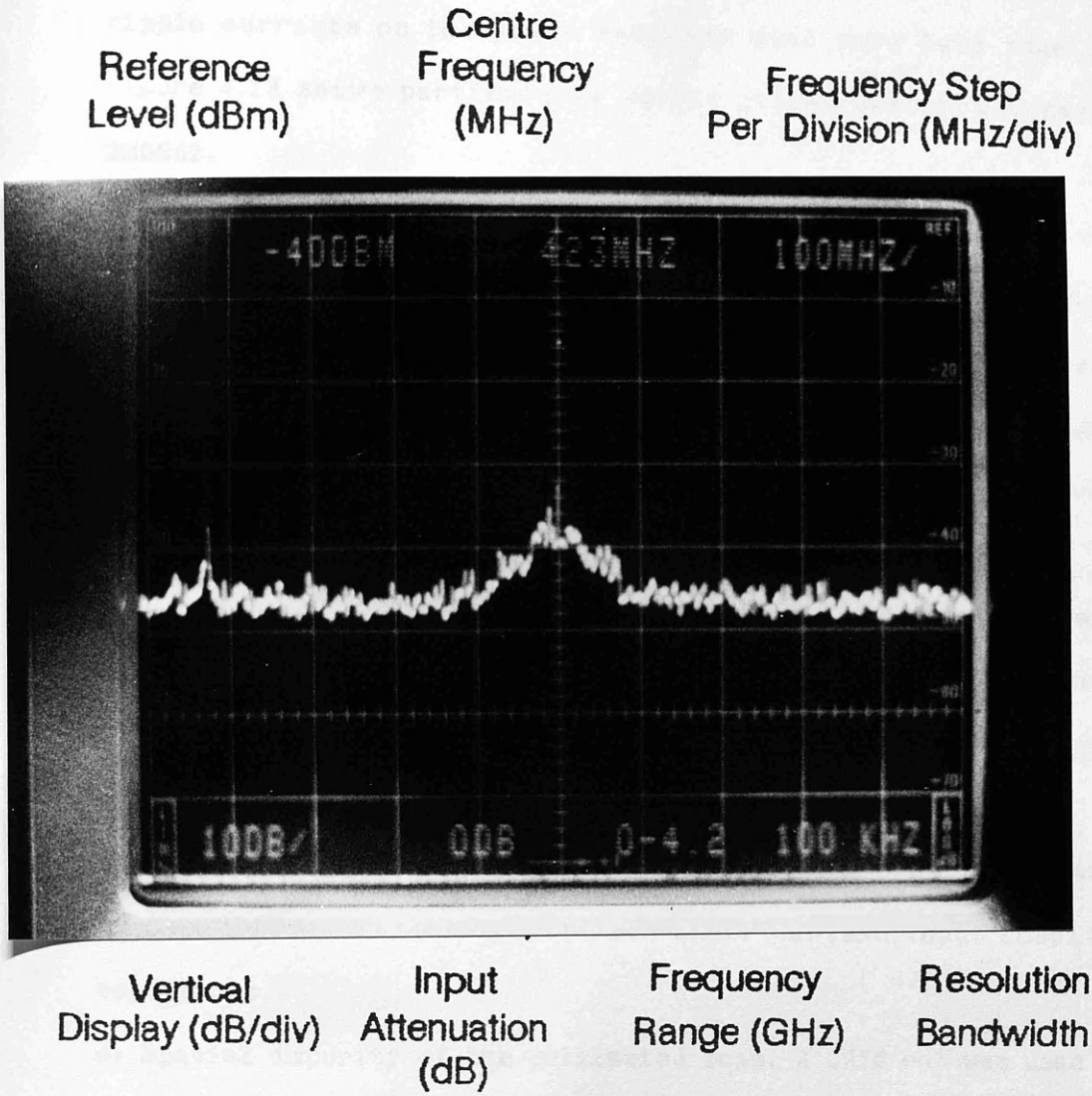


Figure 4.12. Linewidth of Free Running 'Steele' Laser Diode.

$$I_B = 1.6\, I_{th}$$

ripple current could also contribute to this phenomenon. However, ripple currents on the power supplies used were less than 1μA. Figure 4.13 shows particularly severe jitter measured on device 2M0862.

b) Extraneous reflections from experimental equipment. Even with an optical isolator in place, care had to be taken to avoid reflections, usually off the fibre input facet. Many early results showed misleadingly narrow linewidths as a result. Figure 4.14a and b show the effectiveness of the optical isolator when measuring the linewidth of one of the Steele lasers.

c) Thermal drift of the diffracted beam. The diffraction angle in the Bragg cell appeared to vary as the device warmed up. Moreover, the diffraction efficiency also decreased. This necessitated the periodic realignment of the beam.

d) Dual mode operation of the 0.5km fibre length. Two fibre modes could be observed at certain times under certain input coupling conditions.

e) Spatial impurity of the collimated beam. A GRIN rod was used to collimate the lasers output. This produced concentric rings about the central beam spot which proved problematic when coupling into the fibre. Spatial filtering alleviated this problem.

f) Mode hopping of the source. If the source was not operating on a stable mode then obviously fluctuations in the output spectrum would occur.

g) Mechanically and acoustically induced fluctuations in path length. These perturbations of optical components were initially considered to be negligible since the frequency of fluctuation was low (~kHz) and would not add significantly to the noise already present. However, it was quickly found that the coupling optics (as has been mentioned) were considerably less stable than was

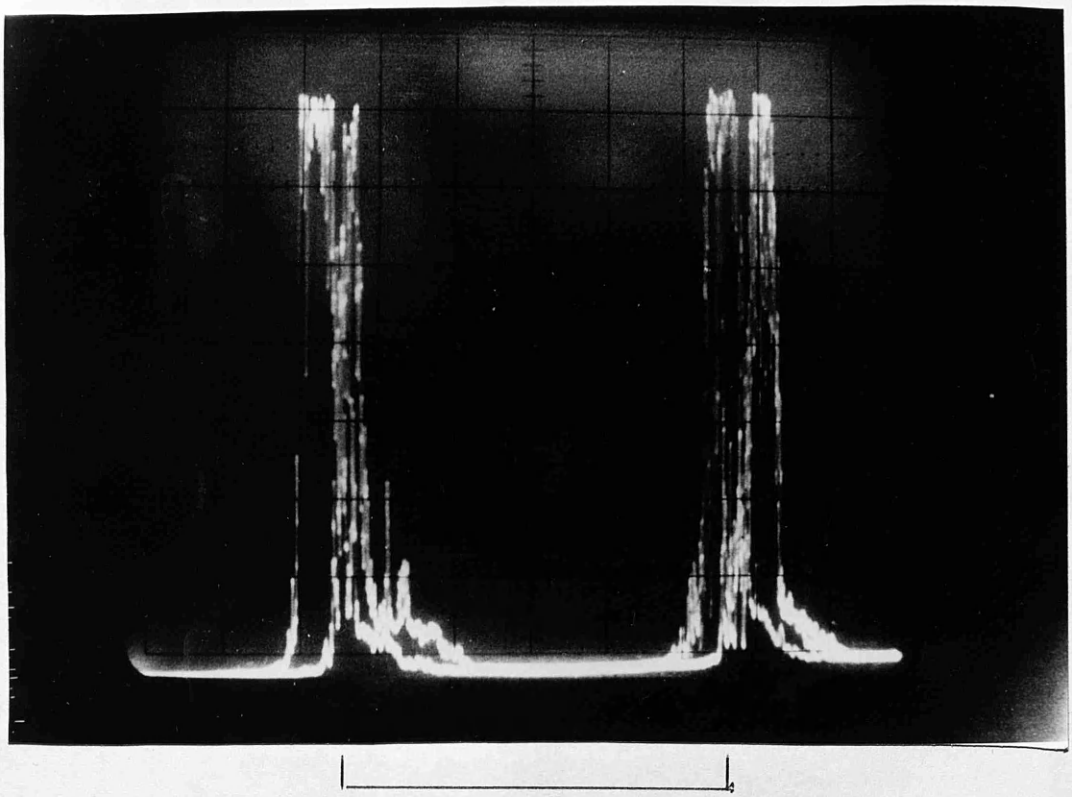


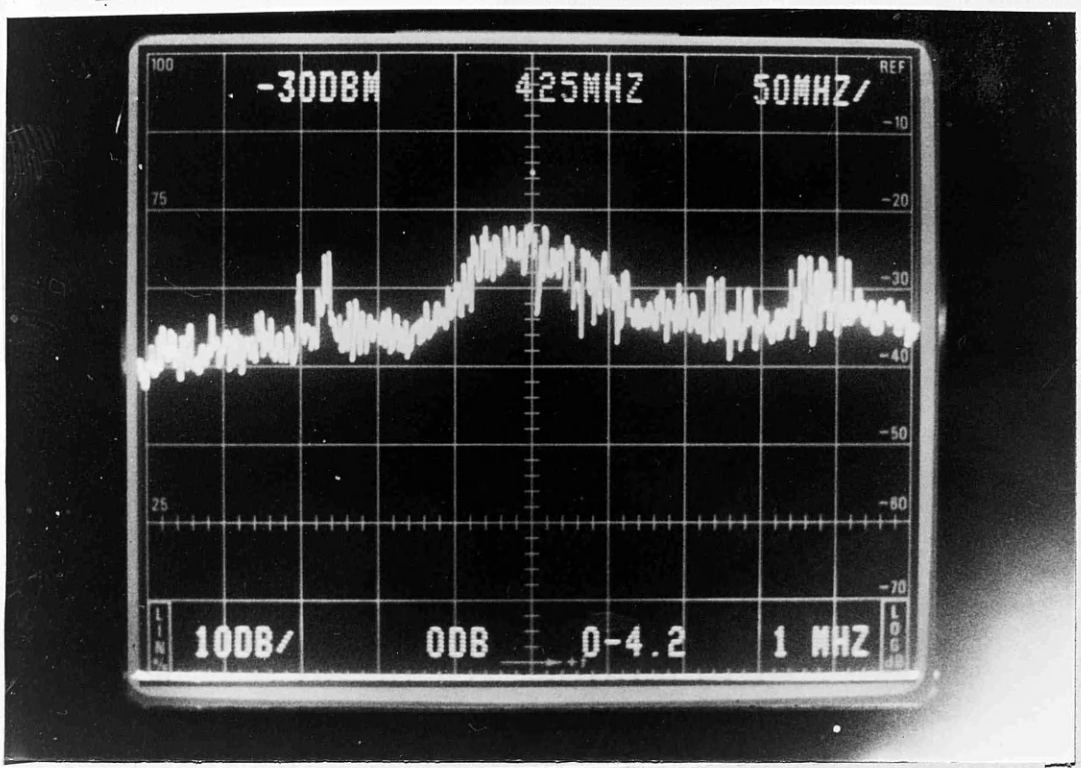
Figure 4.13. Frequency Jitter Exhibited By Device 2M0862.

$I_B = 1.6\ I_{th}$

Reference Level (dBm)

Centre Frequency (MHz)

Frequency Step Per Division (MHz/div)



Vertical Display (dB/div)

Input Attenuation (dB)

Frequency Range (GHz)

Resolution Bandwidth

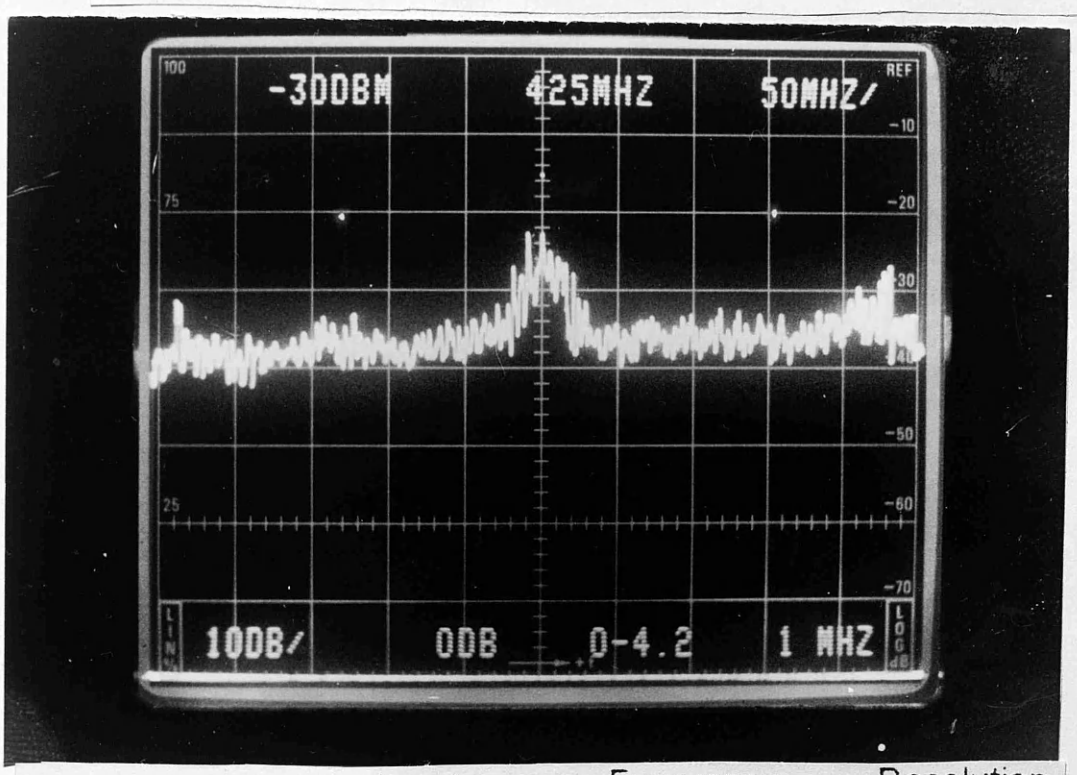
Figure 4.14a. Laser Linewidth Measurement With Optical Isolator.

$I_B = 1.6\ I_{th}$

Reference Level (dBm)

Centre Frequency (MHz)

Frequency Step Per Division (MHz/div)



Vertical Display (dB/div)

Input Attenuation (dB)

Frequency Range (GHz)

Resolution Bandwidth

Figure 4.14b. Laser Linewidth Measurement Without Optical Isolator.

$I_B = 1.6\ I_{th}$

required. It was also found that the fibre suffered from microphonic perturbations, and that the disturbance of free standing optical components such as mirrors, beam splitters, and the Bragg cell were adding in a limited fashion to these effects. These would be a limiting factor in later experiments.

Measured spectra were therefore found to be variable and unstable and many measurements had to be made before a realistic result could be obtained. A limited amount of processing was available on the spectrum analyser (Tektronix model 492) to remove extraneous noise components. The functions available from this instrument were an instantaneous freeze of a sample set, time average over the period, or a maximum sample hold over the period. Normally the first function was chosen as it was considered more representative of the actual linewidth. Its use on photographs of spectra will be assumed unless otherwise stated.

Measurements on the new lasers show a considerable improvement on the 50MHz linewidth of the Steele modules. Figure 4.15 shows a typical spectrum from device 2M0862. From this and other measurements, the individual linewidths of each of the new modules were found to be around 10-15MHz. As a result of the low powers available at the APD, most of the linewidth measurements had to be made at $1.6I_{th}$ in order to be significantly above the noise. It was therefore not possible to experimentally confirm the $1/P_o$ dependence of the linewidth of these devices.

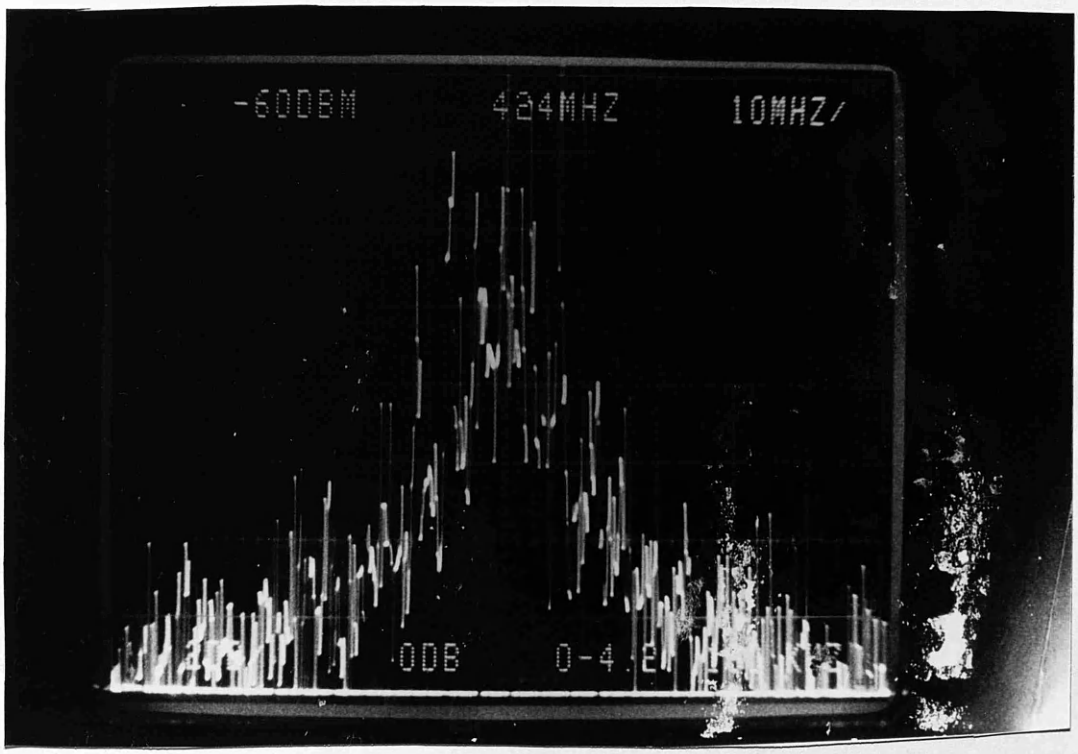
4.4. The Fibre Cavity Experiment.

The use of a short length of optical fibre as a passive cavity in a laser module was first demonstrated by Favre et al[10,17,18]. In an attempt to improve on previous systems in this project[5], this approach was considered to be promising since the

Reference Level (dBm)

Centre Frequency (MHz)

Frequency Step Per Division (MHz/div)



Vertical Display (dB/div) [2 dB/div]

Input Attenuation (dB)

Frequency Range (GHz)

Resolution Bandwidth [100 kHz/div]

Figure 4.15. Self Heterodyne Spectrum of Device 2M0862.

$I_B = 1.6\ I_{th}$

results reported indicated that a stable, controllable and compact cavity had been achieved. Designs were therefore drawn up to enable a length of optical fibre to be coupled to the rear of an HLP 1400 diode while maintaining a strong collimated forward beam for linewidth monitoring and experimental purposes. The results obtained from this experiment will be reported here.

4.4.1. Theoretical Analysis.

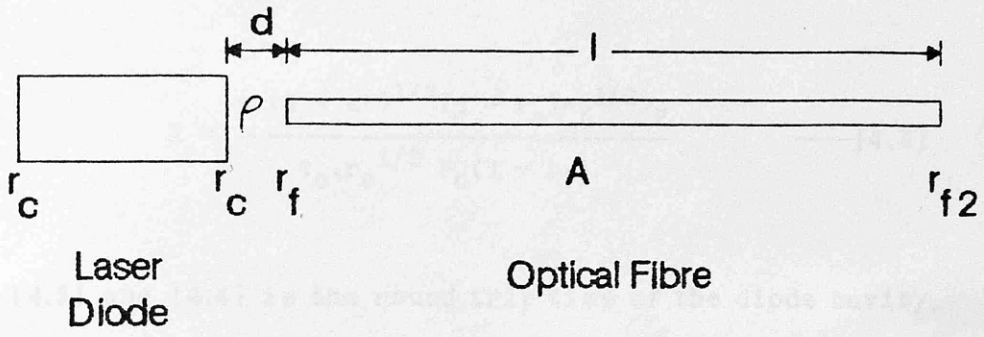
The model used in the analysis of this cavity has already been mentioned in chapter 3 (Figure 3.7). If the diode facet reflectivity is denoted by r_c and the fibre facet by r_f, then an expression for the power coupled back into the laser cavity relative to that reflected internally, can be derived:-

$$(\eta)^{1/2} = \frac{(1 - r_c)(1 - r_f)\rho A.r_f^{1/2}}{r_c^{1/2}} \qquad \text{----[4.3]}$$

where ρ is the fibre input coupling coefficient (Figure 4.16) and A is the fibre attenuation coefficient. In this experiment, only a short length of fibre is used. Hence, A will be assumed to be unity. From rms values of refractive index, r_f and r_c take values of 0.035 and 0.32 respectively[17]. The major variable in equation [4.3] is then ρ. This relates the power output from the diode facet, P_o, to that output at the far fibre facet, P_F:-

$$\rho = \frac{P_F}{P_o(1 - r_f)^2} \qquad \text{----[4.4]}$$

By using expression [3.33]:-



. Figure 4.16.

Schematic of Four Facet

Fibre Cavity Laser Module.

ρ = Laser – Fibre Coupling Ratio.
A = Fibre Attenuation Coefficient.

$$X = k\tau(1 + \alpha^2)^{1/2} \qquad \text{----[3.33]}$$

and expressing k in terms of ρ[17]:-

$$k = (\rho)^{1/2}/\tau_c \qquad \text{----[4.5]}$$

an expression for the linewidth reduction factor in terms of fibre output power can be obtained:-

$$X = \frac{\tau.(1 + \alpha^2)^{1/2}(1 - r_c)r_f^{1/2}P_F}{\tau_c.r_c^{1/2}\ P_o(1 - r_f)} \qquad \text{----[4.6]}$$

τ_c in [4.5] and [4.6] is the round trip time of the diode cavity, as defined in chapter 3. For an HLP 1400, τ_c has been measured[17] as 7.41×10^{-12}. Assuming $\alpha = 4$ (section 3.2.4), a fibre length of 50cm, and n=1.4 for silica, then [4.6] becomes:-

$$X = \frac{441.04\ P_F}{P_o} \qquad \text{----[4.7]}$$

This, substituted in [3.32], is a convenient measure of expected linewidth reduction for given output conditions.

The fibre cavity format also has additional benefits in operation. While the laser diode has a negative frequency-temperature coefficient, the fibre has a positive coefficient. This leads to the expression under optical feedback conditions of:-

$$\frac{\Delta V_{OF}}{\Delta T} = 2.2\ \frac{X - 10}{X + 1}\ \text{GHz/K} \qquad \text{----[4.8]}$$

where $\Delta V_{OF}/\Delta T$ is the change in emission frequency with temperature of the system. It can be seen that for X=10, the output frequency does not drift with temperature, assuming mode hopping between external cavity modes (separated by $c/2.n_F.L_F$) can be ignored. To achieve this second condition, however, requires the fibre diode temperature to be controlled to within[18] $\pm 4 \times 10^{-3}$K for a 50cm cavity. Similarly, the frequency-current coefficient is reduced by a factor of:-

$$\frac{\Delta V_{OF}}{\Delta I} = \frac{3}{1 + X} \text{ GHz/mA} \qquad \text{----[4.9]}$$

Temperature control to that degree can be achieved with careful engineering. In addition, careful module design is also required to provide adequate stabilisation of d, the fibre/ laser separation. A change of $\lambda d/2l$, where l is the diode length, will induce a cavity mode hop. Thus, for a separation of 300μm, this distance must be maintained to $\lambda/2$.

4.4.2. Experimental Results.

Since the self heterodyne experiment is transparent to low frequency mode hopping, it was decided that initial fibre cavity designs need not initially incorporate exceptionally high degrees of mechanical and thermal stability. Thus, an initial proving module was constructed (Figure 4.17). This consisted of a copper fibre mount around which 50cm of fibre was wound and clamped after having been cleaved at both ends. The module was then positioned at the rear facet of the laser. Owing to the close proximity of the fibre and the laser, a microscope was used to monitor the

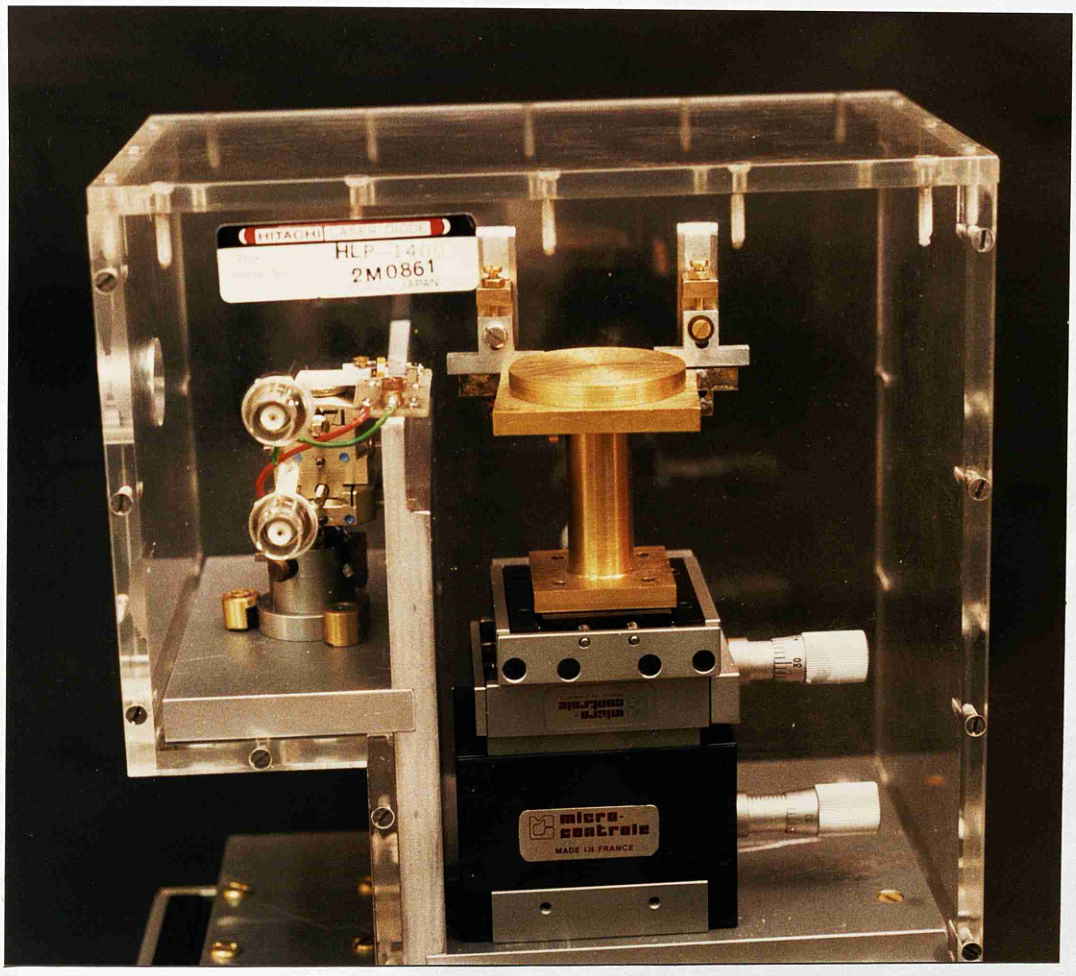


Figure 4.17. The Fibre Cavity Module.

final positioning adjustments. Initial experiments were carried out on the older Steele lasers so that operating experience could be gained without endangering new devices. Measurements of rear facet output power, P_O, versus drive current had previously been made (Figure 4.18). Thus, by measuring the fibre output power, P_F, an expected value of X could be calculated from [4.6].

Results from both the calculations and the self heterodyne experiment are shown in figure 4.19. In all of these measurements, I_D= 110mA and, hence, P_O = 4.4mW. It can be seen that the measured reduction was much less than the calculated values. Considering the noise present in the experiment from sources already listed (e.g. 1/f noise enhancement of the spectra), this anomaly was not unexpected. Moreover, difficulty was experienced when trying to isolate the fibre output from all other extraneous light sources, especially direct pickup from the laser diode. In view of these factors, the two sets of results then show reasonable agreement.

The major conclusion of this experiment, however, was that the maximum P_F obtained was less than 3.5μW. This produced a maximum linewidth reduction factor of X = 0.34. This falls far short of the desired figure of X = 10. The reason for such a low coupling factor stems from the mounting of the laser. The laser chip, as described in section 4.2.1, is mounted p-side down with the active layer only 4μm above the heatsink. The closest the fibre can be placed to the chip, with the core at the level of the active layer is 2mm. As the beam divergence from the non-collimated diode is 10^o x 25^o, ρ is very small. Even if some form of collimation was employed, the front fibre facet would be too far away from the chip to remain an insignificant reflecting facet. If the fibre is moved closer on top of the heatsink, then, because the 8μm core is surrounded by 121μm of cladding, the core

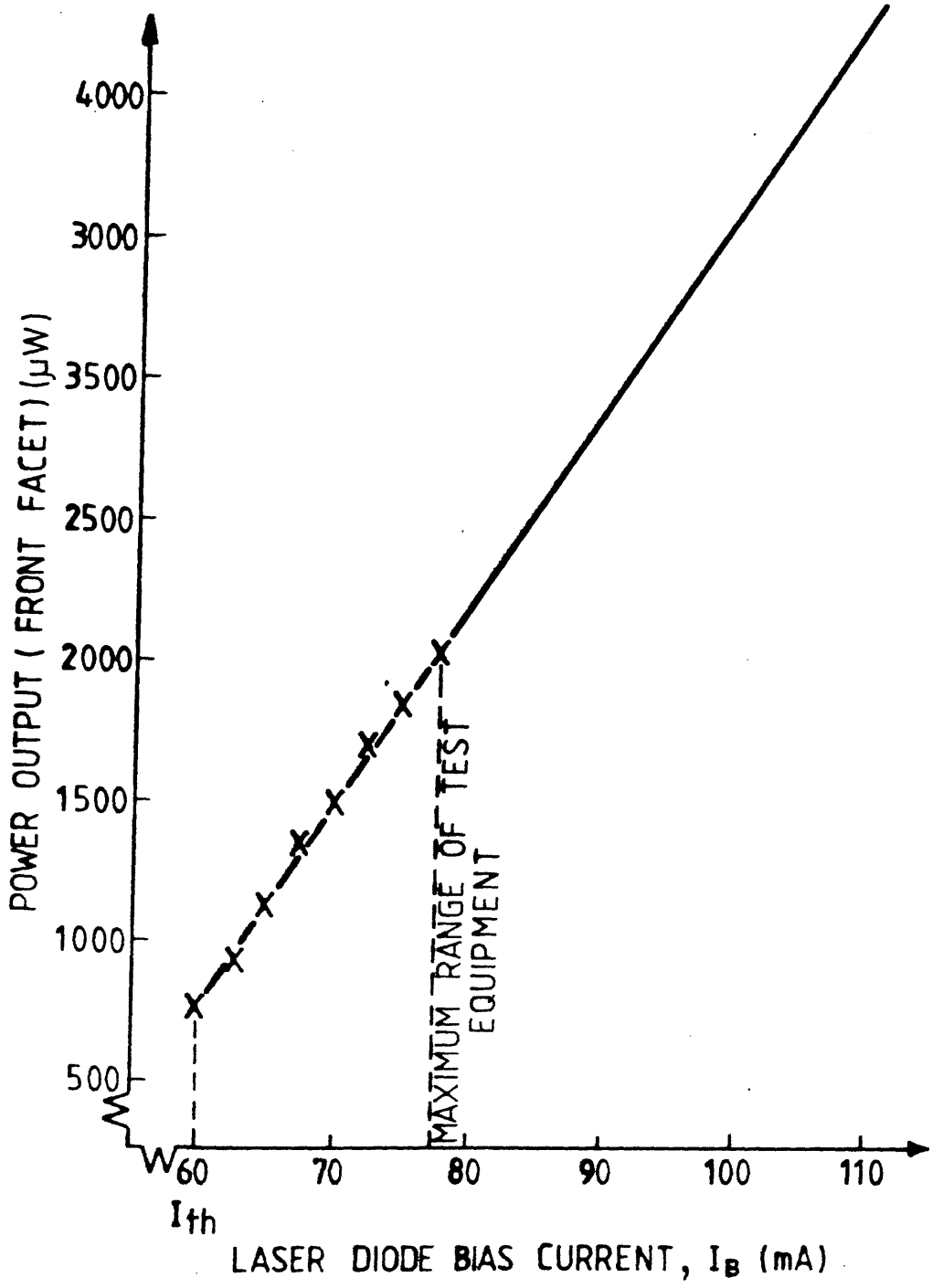


Figure 4.18. Injection Current versus Output Power for an HLP 1400 Laser Diode.

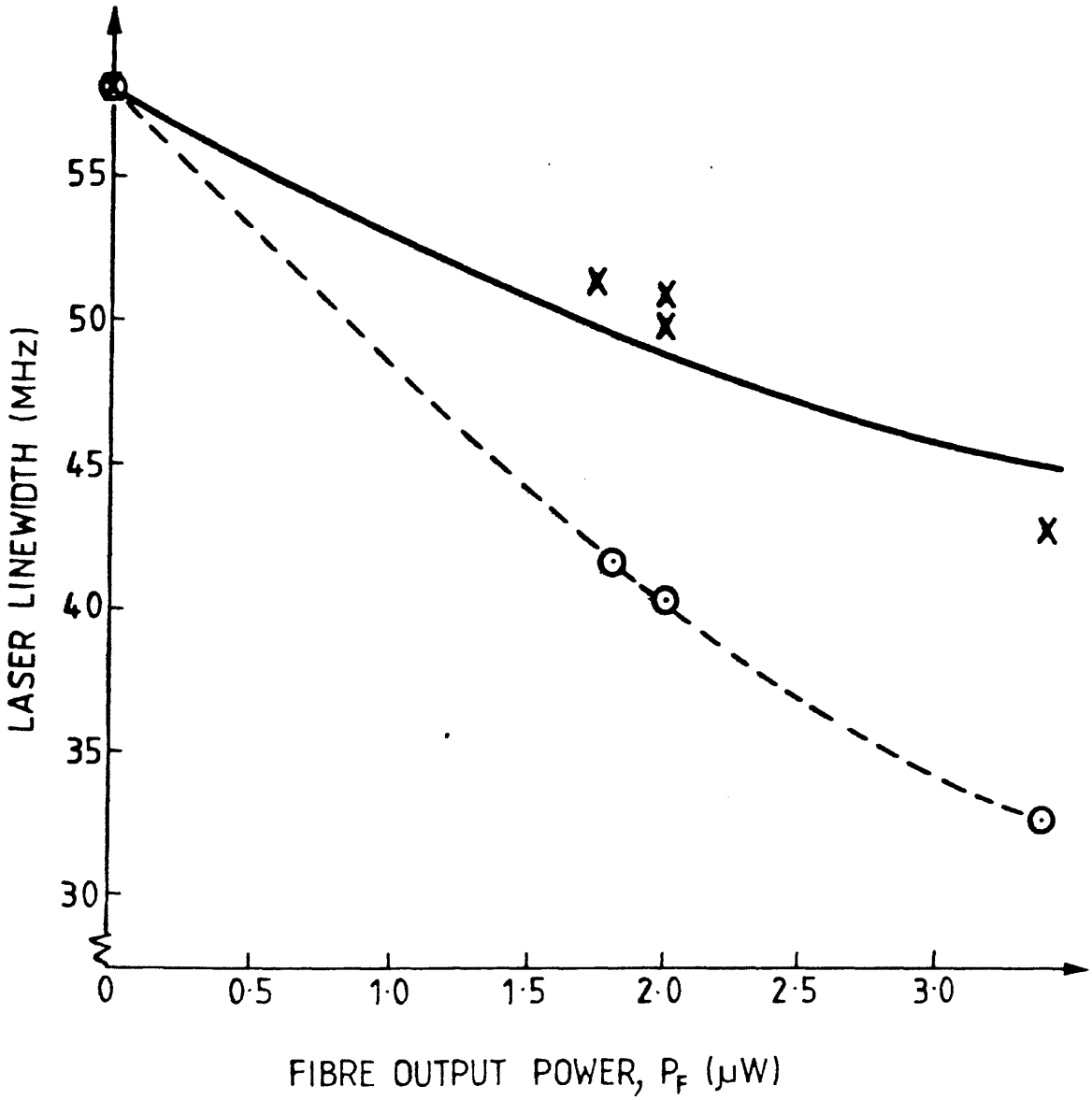


Figure 4.19. Linewidth Reduction as a Result of Optical Feedback from a Passive Fibre External Cavity.

Solid Line = Experimental Results.

Dashed Line = Theoretical Results.

is totally out of alignment with the active region. Therefore, some method of bringing the core into alignment with the active layer is required. Immediately, coupling to the front facet of the laser can be ruled out since that collimated output would be required for monitoring and processing purposes. In addition, any attempt to structurally remove part of the heatsink would inevitably lead to device destruction. Attempts were therefore made to reduce the total diameter of the fibre by chemical etching of the cladding.

4.4.3. Fibre Etching.

The etch used was a 7:1 solution of water and hydrofluoric acid. Depending on the temperature of the solution, this could etch fibres down to ~10μm in a few minutes. It was desirable, however, not to etch too quickly since the fibre could easily disintegrate completely if left in a heated solution fractionally too long. All etching was therefore done at room temperature. It was decided that in order to make a significant impact over the unetched fibre, etching to a diameter of around 10μm should be attempted. Two lengths of fibre, 50cm long, each had one end suspended in solution along with a number of short test fibres used to monitor the etch rate. After 6.1/2 hours the fibres were removed, carefully washed and potted in wax. This allowed end polishing of the fibre to take place in an attempt to achieve a clean orthogonal facet. The fibre was then removed, cleaned and then placed in the holder. It was noticed that when the fibres were removed from the etch that, apart from being very fragile as expected, the process had induced bending of the etched length of the fibre. This was a major problem when it was mounted in the fibre holder as it was impossible to align the bend in the plane of the junction, thus aligning the core with the active region of

the laser. Attempts to do so only resulted in the destruction of the etched length. Attempts to align the fibre with the bend not in the junction plane also failed, usually due to insufficient movement in the micropositioners that held the cavity.

Further lengths of the fibre were then etched down to 45μm in an attempt to reduce fibre bending. This it succeeded in doing and the same procedure was then followed. Success was achieved in coupling light into this length of fibre. However,the maximum measured output was 0.6μW, less than that achieved with the full diameter fibre. This was probably a result of non-orthogonal end polishing of the fibre and, hence, increased coupling losses.

4.4.4. Increasing Rear Facet Reflectivity.

Having failed to achieve any significant increase in X from etching the fibre, consideration was given to the effect of increasing the rear fibre facet reflectivity. This could be achieved by evaporating a metal coating (e.g. aluminium) onto the fibre end. The effectiveness of this action was first calculated by modifying [4.3] in [3.33]. If the reflectivity of the rear facet is designated r_{f2} then [4.3] becomes:-

$$(\eta)^{1/2} = \frac{(1 - r_c)(1 - r_f)\rho A . r_{f2}^{1/2}}{r_c^{1/2}} \qquad \text{----[4.9]}$$

The increase in X can then be calculated by comparing [4.9] and [4.3] in [3.33]. This results in:-

$$\frac{X'}{X} = \left[\frac{r_{f2}}{r_f} \right]^{1/2} \qquad \text{----[4.10]}$$

where X' is the modified linewidth reduction factor containing r_{f2}. If $r_{f2} = 1$ and $r_f = 0.032$ then X is increased by a factor of 5.5. Given that the maximum X obtained from the fibre cavity module was 0.34, this still falls far short of the desired value of 10. In practice, a value of 1 for r_{f2} was never achieved as thick coatings of aluminium would not adhere to the small diameter fibre. Thus no significant improvements in X were observed when this technique was implemented.

It was concluded, therefore, that primarily as a result of the way in which the laser chip was mounted, a useful reduction in the laser linewidth was unlikely to be achieved using a rear facet coupled fibre cavity. Since, as has been stated, it was impractical to mount the cavity in front of the laser, other methods of linewidth reduction were then considered.

4.5. Partially Reflective Mirror Cavities.

The simplest form of optical feedback system is the partially reflective mirror cavity. It was decided that, when the fibre cavity results did not look promising, two co-workers, Mr M.J.Fletcher and Mr W.C.Michie, would undertake the design and construction of two mirror cavity modules. 20cm and 30cm were chosen as the cavity lengths. Figure 4.20 shows the 30cm module incorporating device 5E3070.

The mirror consisted of a microscope cover slip polished at an angle on one side to spatially separate the reflections from front and back surfaces. The mirror mount was initially adjusted using two fine threaded brass screws. This was not a particularly stable arrangement but was suitable for initial experiments. Later models would use conventional micrometer actuators. With careful adjustment of the mirror, substantial linewidth reductions could

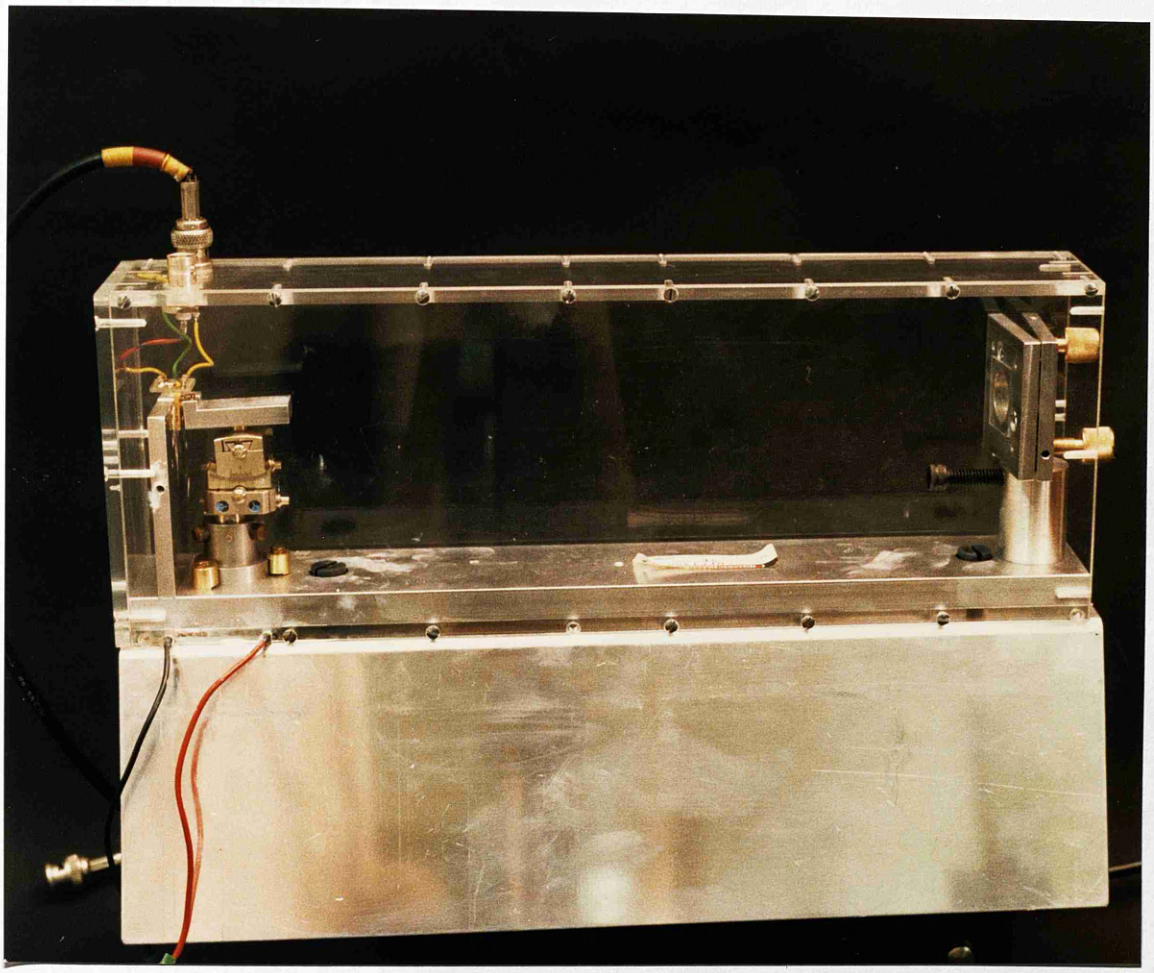


Figure 4.20. The Passive Mirror Cavity Laser Module.

External Cavity Length = 30cm.

Installed Diode is Device No. 5E3070.

be obtained. Figure 4.21 shows the results of measurements on module 5E3060 with a 20cm cavity. The 3dB linewidth is of the order of ~750kHz. However, long term stability of this line was poor and acoustic and thermal perturbations would cause the effect to be lost after a few minutes.

Short term stability was also affected by acoustic or mechanical transients on the mirror inducing path length fluctuations. From equation [3.32] it can be seen that changes in path length of fractions of a wavelength cause substantial variations in linewidth reduction. In addition, competition between several external cavity modes was also apparent from observations on the scanning Fabry Perot (Figure 4.22). However, the observed frequency jitter of the line was substantially reduced (cf. Figure 4.13).

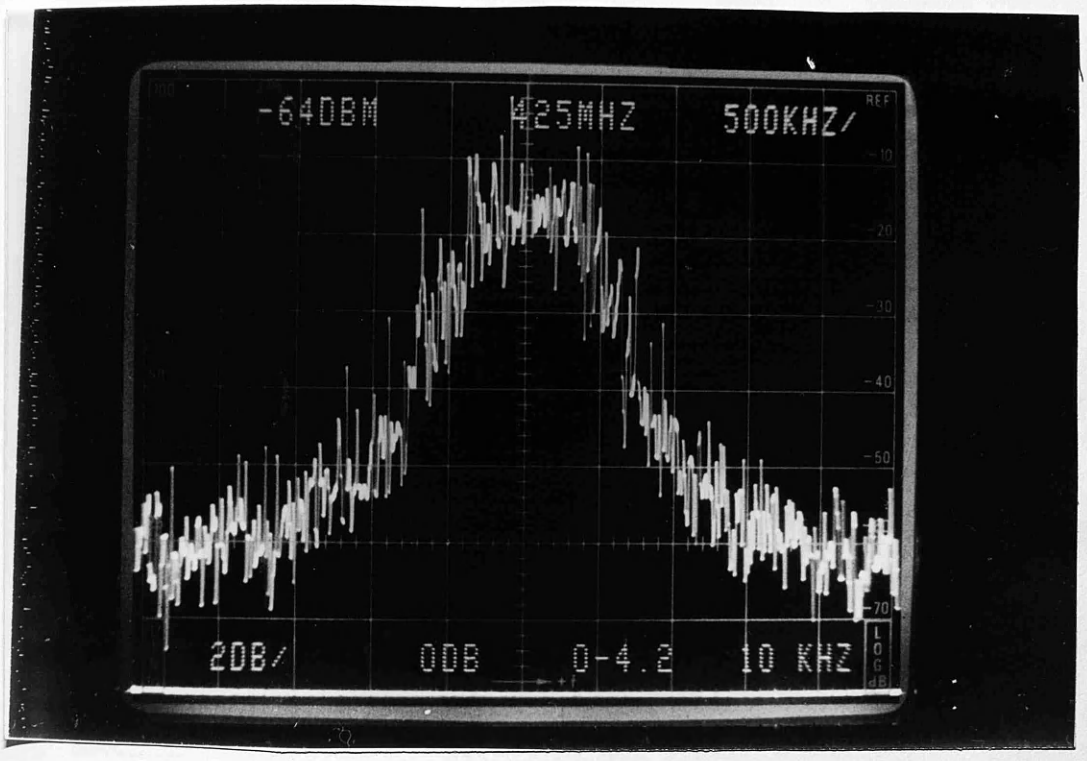
An additional problem was the lack of tunability of these sources. With any resonant cavity, the output emission is only tunable over a small percentage of the cavity mode spacing, that percentage being proportional to the Q of the cavity. Therefore, unless the effective length of the cavity can be easily altered, as in the fibre and external grating cases, continuous tunability, and hence the selection of an arbitrary wavelength, is not possible. Moreover, unless the lengths of the cavities could be determined to within $\lambda/2$, there would be no guarantee that the two modules could be made to lase on the same wavelength.

One possible solution might be to use one line narrowed cavity and one free running tunable laser, namely device 5E3072. This still presupposes that a 10-15MHz beat linewidth would be within practical limits for phase locking. Whether this is the case will be discussed in detail in chapter 7 in the light of the analysis presented in chapters 5 and 6.

Reference Level (dBm)

Centre Frequency (MHz)

Frequency Step Per Division (MHz/div)



Vertical Display (dB/div)

Input Attenuation (dB)

Frequency Range (GHz)

Resolution Bandwidth

Figure 4.21. Linewidth Reduction in a Mirror Cavity Module.

Cavity Length = 20cm

Laser Diode = Device No. 5E3060

Estimated FWHM ~ 750kHz

$I_B = 1.6\, I_{th}$

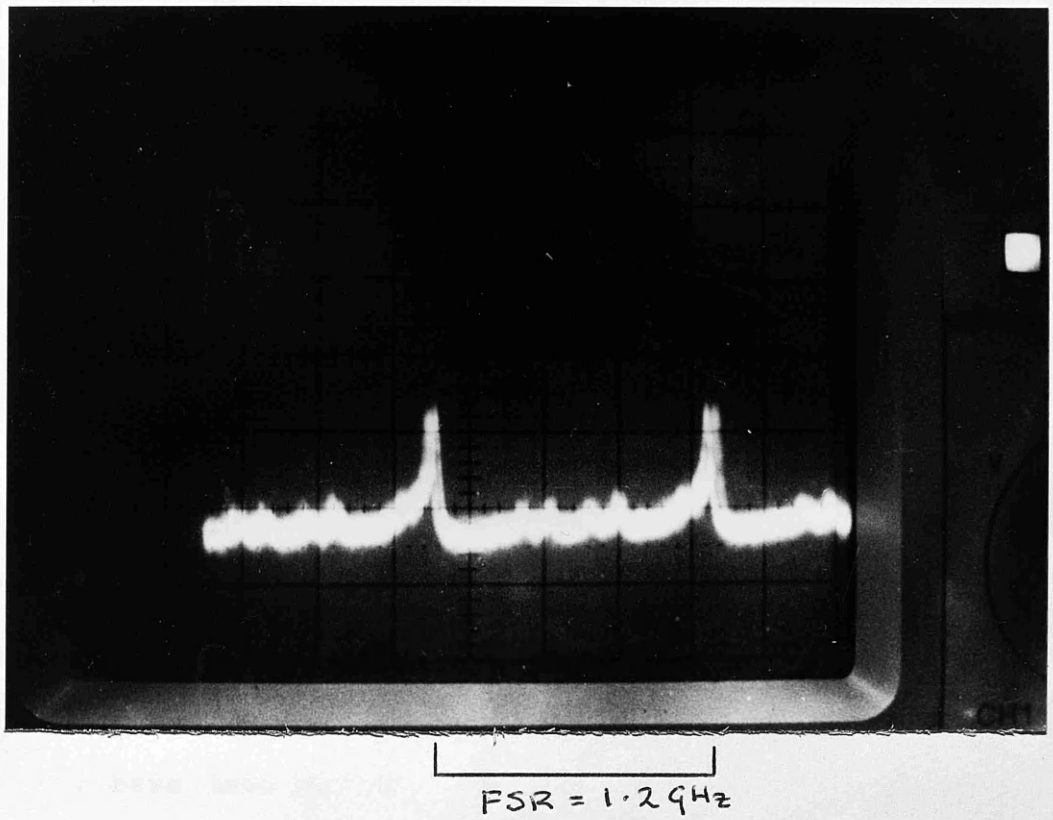


Figure 4.22. Frequency Jitter and Cavity Mode Competition Exhibited By Module 5E3060 under the Influence of Optical Feedback.

Cavity Length = 60cm

$I_B = 1.6\, I_{th}$

4.6. Semiconductor Laser Modules: Discussion and Conclusions.

This chapter has presented a description of the chosen semiconductor laser, a Hitachi HLP 1400 GaAlAs device operating nominally at 825nm. The modulation response of the laser has been theoretically discussed and experimentally measured. The roll-off in the response due to thermal effects in the device has been observed, and the relaxation oscillation resonance due to carrier effects discussed. The DC thermal response has been calculated as -38GHz/K and measured as -22GHz/K. The DC injection current coefficient has been also measured as 4GHz/mA.

An experimental technique to measure the linewidth of a laser module has been constructed. This experiment, which was first demonstrated by Okoshi et al[15], is an optical autocorrelation technique, and has become more commonly known as the self heterodyne experiment. Results on the measurement of free running lasers have been presented showing that each device has a linewidth of 10-15MHz. Two older devices tested exhibited substantially wider linewidths of some 50MHz. This is probably due to the phenomena of spectral ageing. A technique to determine single mode operation of the source and to measure the peak deviation of any frequency instability and jitter has also been established with the acquisition of a scanning Fabry Perot interferometer.

A theoretical overview of a fibre cavity linewidth reduction system has been presented. A relationship between the fibre rear facet output power, P_F, and the linewidth reduction factor, X, has been derived. Experimental implementation of such a cavity was attempted but was hindered primarily as a result of the physical structure of the HLP 1400 laser. The inaccessibility of the rear

facet resulted in insufficient power being coupled into the fibre cavity resulting in no meaningful linewidth reduction being observed. A small reduction, consistent with the measured maximum value of P_F was recorded. Attempts to increase the laser/fibre coupling coefficient, ρ, by fibre etching did not succeed. No other suitable laser was available at the time of the experiment to attempt to rectify this problem.

A mirror coupled cavity module was then constructed and tested. This exhibited a minimum linewidth of ~750kHz which is much more appropriate for coherent optics experiments. However, this narrow mode did not exhibit long term stability. Neither did it exhibit a substantial tuning range, a very desirable feature for a coherent communications source. A number of contributory factors to modal and FWHM instability were discussed. These included mechanical frailty of the mirror mount, and mechanical and acoustic perturbations of the cavity.

Practical measures to improve on the device described here involve mainly the elimination of these perturbing sources. These measures would include

a) establishing a laboratory far from sources of vibration, close to ground level;

b) Conducting experiments on an air-isolated, vibration-free optical table;

c) Obtaining the most stable and accurate optical positioners available. Piezoelectric (PZT) positioners are considered essential in obtaining sub-micron resolution for optical elements.

Action on these points would then provide immediate improvements in the system described above possibly providing a usable source in a relatively short timescale.

However, the production of the ideal source for coherent

communications experiments requires a completely new approach to the problem. The result of that approach would produce a device with a linewidth of a few kHz, an absolute modal stability of a few MHz and a continuous tuning range of several nanometers in order that an arbitrary wavelength could be selected. Wyatt et al[19] have come close to this mark with the external grating cavity laser, and others have followed this lead[3,20]. However, the devices produced here fall far short of the above performance, and to attempt a more complex module would not be possible in the time allocated to the project. It would be recommended that future workers in the field consider constructing an external grating semiconductor laser module. In conclusion, it is therefore necessary to continue with laser modules which fall short of the desired specifications.

References,

1. D.J.Malyon et al 'PSK Homodyne Receiver Sensitivity Measurements at 1.5μm' Elect. Lett. 19(4) 1983. pp144-146.

2. R.Wyatt et al '140MBit/s Optical FSK Fibre Heterodyne Experiment at 1.54μm' Elect. Lett. 20 1984. pp912-913.

3. R.A.Linke et al 'Coherent Lightwave Transmission Over 150km Fibre Lengths at 400MBit/s and 1GBit/s Data Rates Using DPSK Modulation' Post Deadline Paper, 11th European Conference on Optical Communications, Venice, Italy, October 1985.

4. R.C.Steele 'Optical Phase-Locked Loop Using Semiconductor Laser Diodes' Elect. Lett. 19(2) 1983. pp69-70.

5. R.C.Steele, Ph.D. Thesis, University of Glasgow, 1983.

6. S. Kobayashi et al 'Direct Frequency Modulation in AlGaAs Semiconductor Lasers' IEEE J. Quant. Electron. vol QE-18(4) 1982. pp582-595.

7. D.Welford, S.B.Alexander 'Magnitude and Phase Characteristics of Frequency Modulation in Directly Modulated GaAlAs Semiconductor Diode Lasers' IEEE J. Lightwave Technol. vol LT-3(5) 1985. pp1092-1099.

8. J.O.Askautrud, Final Year Project Report, Dept. of Electronics and Electrical Engineering, University of Glasgow, 1986.

9. 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. Quant. Elect. vol QE-20(5) 1984. pp478-491.

10. F.Favre, D.LeGuen 'Emission Frequency Stability in Single Mode Fibre Optical Feedback Controlled Semiconductor Lasers' Elect. Lett. 19(17) 1983. pp663-665.

11. M.Hercher 'The Spherical Fabry Perot Interferometer' Appl. Optics. vol 5(5) 1968. pp951-966.

12. L.G.Kazovsky 'Decision-Driven Phase-Locked Loop for Optical Homodyne Receivers: Performance Analysis and Laser Linewidth Requirements' IEEE J. LIghtwave Technol. vol LT-3(6) 1985. pp1238-1247.

13. L.G.Kazovsky 'Balanced Phase-Locked Loops for Optical Homodyne Receivers: Performance Analysis and Laser Linewidth Requirements' IEEE J. LIghtwave Technol. vol LT-4(2) 1986. pp182-194.

14. M.A.Grant, W.C.Michie, M.J.Fletcher 'The Performance of Optical Phase-Locked Loops in the Presence of Non-Negligible Loop Propagation Delay' submitted to the IEEE J. of Lightwave Technol.

15. T.Okoshi, K.Kikuchi, A.Nakayama 'Novel Method for High Resolution Measurement of Laser Output Spectrum' Elect. Lett.

16(16) 1980. pp630-631.

16. J.Salz 'Coherent Lightwave Communications' AT+T Technology Journal vol 64(10) 1985. pp2153-2209.

17. F.Favre, D.LeGuen, J.C.Simon 'Optical Feedback Effects Upon Laser Diode Oscillation Field Spectrum' IEEE J. Quant. Elect. vol QE-18(10) 1982. pp1712-1717.

18. F.Favre, D.LeGuen 'Effect of Semiconductor Laser Phase Noise on BER Performance in an Optical DPSK Heterodyne-Type Experiment' Elect. Lett. 18(22) 1982. pp964-965.

19. R.Wyatt, W.J.Devlin '10kHz Linewidth 1.5μm InGaAsP External Cavity Laser With 55nm Tuning Range' Elect. Lett. 19(3) 1983. pp110-112.

20. F.Favre et al 'External Cavity Semiconductor Laser with 15nm Continuous Tuning Range' Elect. Lett. vol 22(15) 1986. pp795-796.

CHAPTER 5.

THE EFFECT OF PHASE NOISE ON AN OPTICAL HOMODYNE RECEIVER.

5.1. Introduction.

It was concluded in chapter 4 that the laser sources developed in this project were unlikely to be optimum for phase-locked coherent detection systems . They exhibited line broadening and frequency jitter which was expected to be, at best, detrimental and, at worst, destructive to any phase locking process. This claim, however, is at this stage only intuitive. It is the purpose of this chapter and chapter 6 to quantify the effects of phase noise on an optical phase-locked loop. In line with the project objective, this analysis will be based on a PSK homodyne receiver, although the techniques used are equally applicable to other loop types (Appendix 5B).

Various receiver designs suitable for the homodyne demodulation of phase shift keyed signals have been developed for use in microwave systems. The total suppression of the carrier which occurs with phase reversal keyed (PRK) modulation determines that some form of carrier recovery involving a non-linear squaring element must be used. The resultant 2ω term devoid of its modulation component can then be tracked by a standard PLL. Unfortunately, no such element is available for use at optical frequencies. Recovery must therefore either rely on some residual carrier component being transmitted ($\Delta\theta < \pm\pi/2$), with an associated power penalty in reception, or employ a data synchronising loop which effectively 'wipes out' the incoming data signal at base bend. The former solution has previously been investigated[1] and a number of additional penalties in practical implementation identified[2]. It was therefore decided that a data

synchronising system should be the main object of study. The system chosen was a Costas loop[3].

A description of loop operation will be given in section 5.2. In section 5.3, expressions for the loop phase error, θ_E, and total phase error variance, σ_E^2, will be derived. Optimisation of the loop parameters (in particular the loop bandwidth, or natural frequency, ω_n) in relation to a desired loop performance will then be performed in section 5.4 using the phase error variance expression. Initial conclusions on the performance of the Costas loop in the presence of phase noise will then be drawn before further analysis is presented in chapter 6.

5.2. Description of the Costas Loop.

A block diagram of the classical microwave implementation of the Costas loop is shown in Figure 5.1. It consists of a conventional PLL based around the Q arm of the loop, in parallel with a data detection arm (the I arm). The modulation function, m(t), takes the values ± 1 for mark and space states. The input signal is split and mixed in phase and in quadrature with representations of the local oscillator signal. Low pass filters in each arm remove the 2ω terms. When the loop is in lock, the loop phase error, ϕ_E $(=\theta_T-\theta_{LO})$, is small. Hence m(t) is recovered at data out (Figure 5.1) and the Q arm signal shows only the small deviations of θ_E. This signal still, however, contains the antipodal modulation function which, when time averaged, produces zero output. Hence, if this signal was fed back directly to the local oscillator, phase lock would not be achieved. m(t) is therefore removed by combining I and Q arm signals in the third mixer. The mixer output is then proportional to $2\theta_E$, after linearising for small θ_E. This signal is filtered through the loop

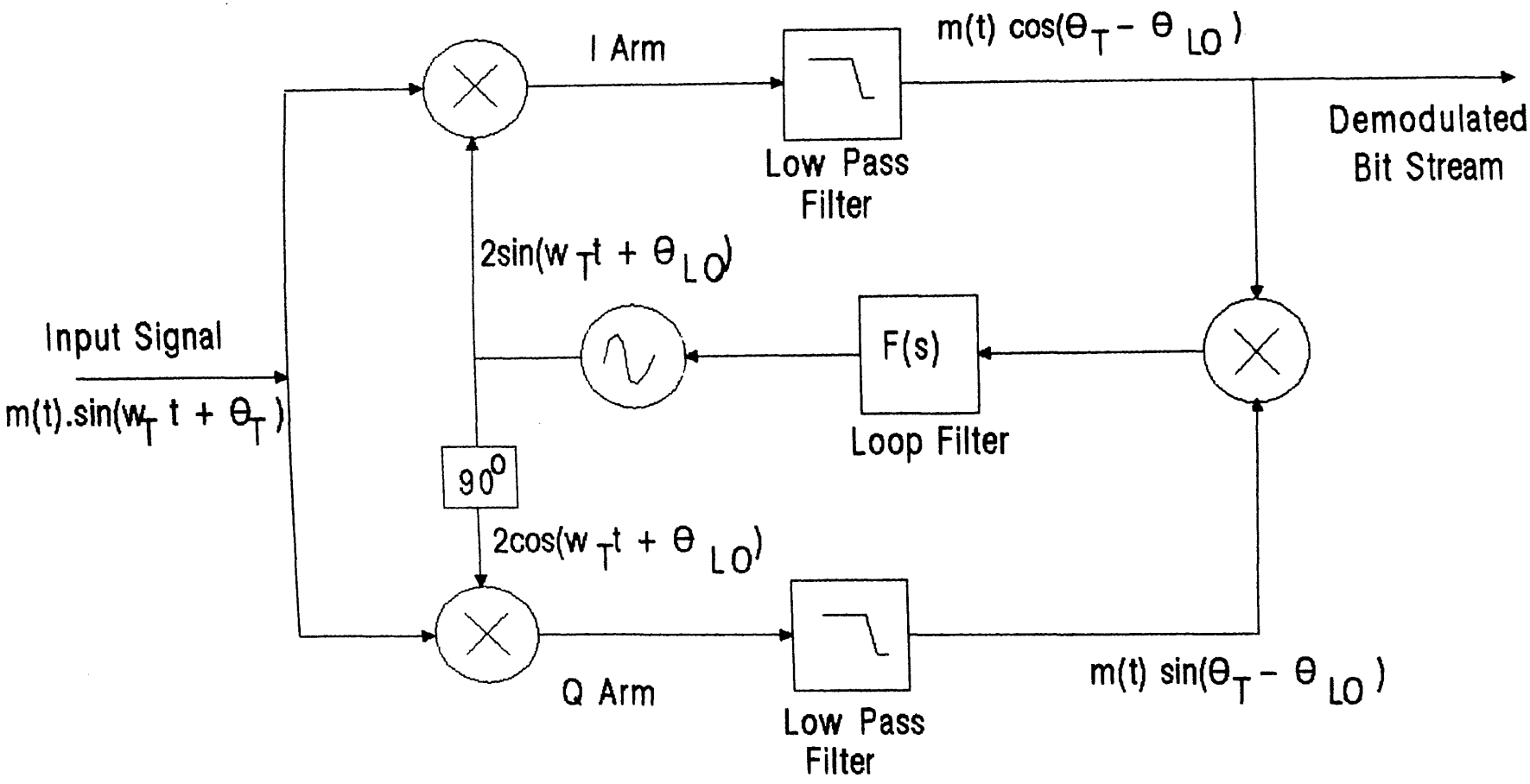


Figure 5.1. The Microwave implementation of a Costas Loop.

*NOTE: While $H_B(s)$ should be designed not to corrupt the received beat signal, that is, the high frequency pole should at least be greater than the bit rate, an excessively large passband in the I and Q arms of the loop is undesirable. The precise criteria for determining the high frequency cutoff of these filters is still the subject of current research. Initial results of this research are briefly discussed in section 6.5.

filter, F(s), and applied to the local oscillator.

5.3. The Optical Costas Loop: An Analytical Model.

5.3.1. General Considerations.

It is not possible to directly implement this loop description in the optical regime. A number of modifications are required. Figure 5.2 shows the modified loop on which the analytical model described here will be based. The fundamental changes are those involving the input signal mixers. Instead of a direct optical mixer, the local oscillator and input signals are spatially combined in beam splitters and focused onto a photodiode. This diode then responds to the beat of the two signals while at the same time performing a low pass filtering function, removing the 2ω terms. If a band pass filter with transfer function $H_B(s)$ is then placed after each diode, the DC terms in the diode photocurrents are removed. It is assumed here that both high and low frequency poles of these filters should not encroach upon the modulation spectrum of the beat signal.* In that way, the modulation and phase error signals are recovered uncorrupted since the phase error signal is effectively modulated onto the bit stream. Moreover, the noise terms in this process can be considered as gaussian or truncated gaussian functions, both of which are well understood.

A second major modification of the loop is the introduction of the coupling ratio, α. This describes the ratio of received signal power delivered to each of the two arms. In the model used here, α is utilised in such a way as to describe the action of combining the two signals in a 90^{o} hybrid of the form described by Leeb[4]. In this analysis, P_{LO} is X% of the total transmitted local oscillator power, where X is the reflectance of the beam combiner

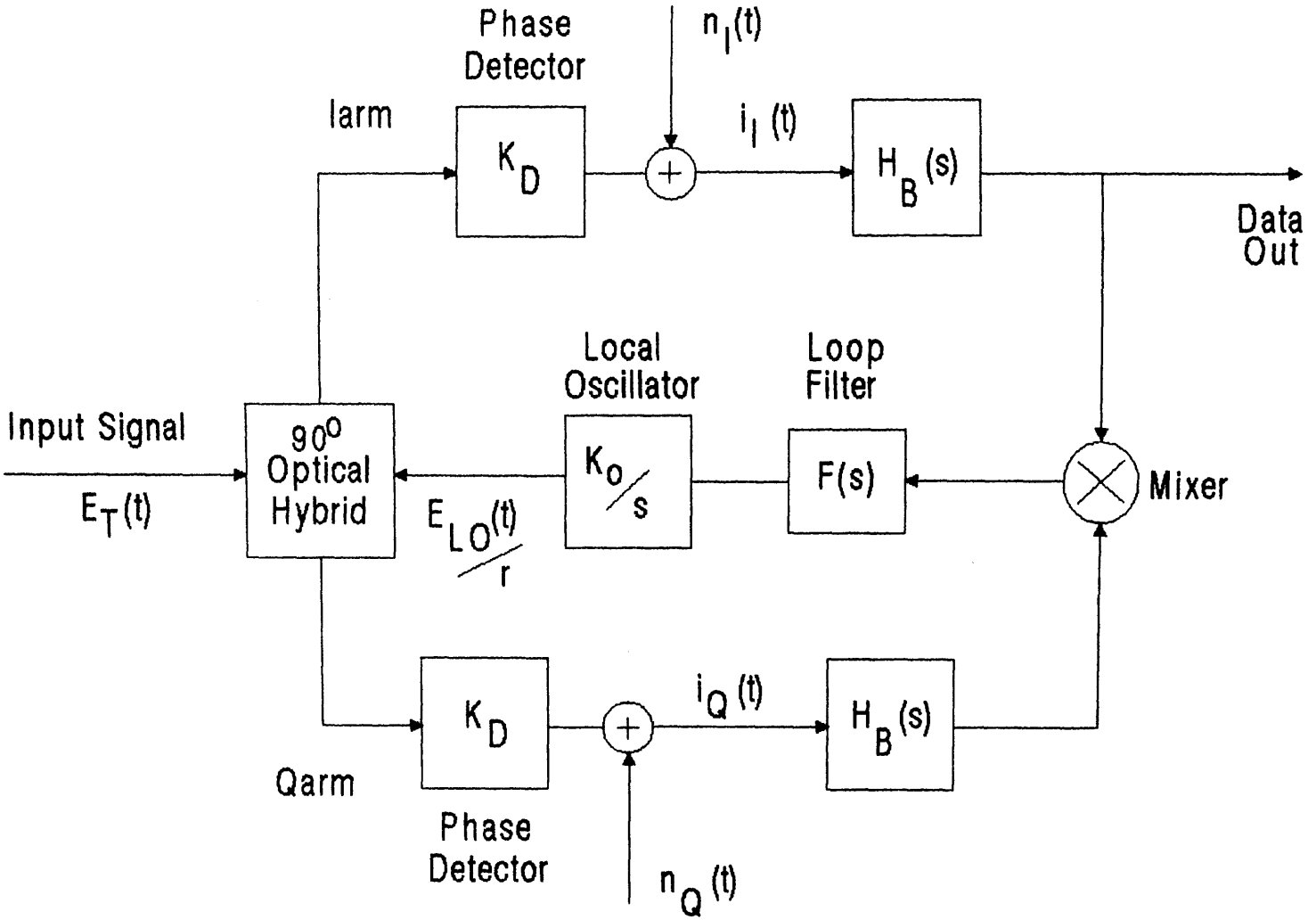


Figure 5.2. Schematic Diagram of an Optical Costas Loop.

r^2 = Power Reflection Coefficient of the Beam Combiner in the 90^o Hybrid

in the optical hybrid (chapter 7). This hybrid is the only practical solution available for this task at the present time.

To model the system, an expression for the total phase error in the loop will first be derived. In common with other standard linear PLL analyses[5,6], the linearising assumption of small θ_E when the loop is in lock will be made, and an expression for θ_E in terms of the loop transfer function, H(s), obtained. Noise performance of the loop will then be quantified by deriving an expression for the phase error variance. The noise bandwidth integrals in this expression will be evaluated for several forms of loop filter in appendix 5A. One filter type, used frequently in practice, will then be chosen and expressions for optimum loop parameters derived. Finally, a number of examples will be given, showing the degradation in performance as a result of large laser line broadening.

5.3.2. System Phase Error.

Figure 5.2 shows the schematic of the optical system to be analysed. Let the input signal $E_T(t)$ be:-

$$E_T(t) = m(t)(2P_T)^{1/2}\sin(\omega_T t + \theta_T) \qquad \text{----[5.1]}$$

where ω_T is the angular frequency of the carrier, m(t) is a modulation function taking the values ± 1, and θ_T is the phase of the signal given by:-

$$\theta_T = \phi_T + \Psi_T(t) \qquad \text{----[5.2]}$$

where ϕ_T is the absolute phase of the signal and $\Psi_T(t)$ is some gaussian distributed random variable which represents the laser

phase noise process described in chapter 3. The double sided power spectral density (PSD) of $\not\theta_T(t)$ is given as:-

$$S_T = \frac{\delta f_T}{2\pi f^2} \qquad \text{rads/Hz} \qquad \text{----[5.3]}$$

A similar set of equations can be attributed to the local oscillator signal, $E_{LO}(t)$:-

$$E_{LO}(t) = (2P_{LO})^{1/2}\sin(\omega_{LO}t + \theta_{LO}) \qquad \text{----[5.4]}$$

$$\theta_{LO} = \not\theta_{LO} + \not\theta_{LO}(t) \qquad \text{----[5.5]}$$

and

$$S_{LO} = \frac{\delta f_{LO}}{2\pi f^2} \qquad \text{rads/Hz} \qquad \text{----[5.6]}$$

Splitting [5.1] and [5.4] in the appropriate ratio, and combining them on the two detectors yields expressions for the photocurrents in each arm at points A:-

$$i_I(t) = R\,[m^2(t).P_T(1-\alpha) + m^2(t).P_{LO}/2] + m(t).K_{DI}.\cos(\theta_T-\theta_{LO}) + n_I(t) \qquad \text{----[5.7]}$$

and similarly:-

$$i_Q(t) = R[m^2(t).P_T\alpha + m^2(t).P_{LO}/2] + m(t).K_{DQ}.\cos(\theta_T-\theta_{LO}) + n_Q(t) \qquad \text{----[5.8]}$$

In these expressions homodyne detection is assumed ($\omega_T = \omega_{LO}$), and K_{DI} and K_{DQ} are the effective diode gains for the respective arms of the receiver, and are defined as:-

$$K_{DI} = R.(2.P_T.P_{LO}.(1 - \alpha))^{1/2} \qquad \text{----[5.9]}$$

and

$$K_{DQ} = R.(2.P_T.P_{LO}\alpha)^{1/2} \qquad \text{----[5.10]}$$

Also $n_I(t)$ and $n_Q(t)$ are additive gaussian noise terms representing the shot noise produced in the photodetection process. Assuming $P_{LO} >> P_T$, these terms have double sided PSD's of:-

$$S_I = R.e.P_{LO}/2 \qquad \text{----[5.11]}$$

and

$$S_Q = R.e.P_{LO}/2 \qquad \text{----[5.12]}$$

The respective photocurrents are then filtered, removing the dc terms, as described, and combined in an ideal double balanced mixer. In practice, some form of saturation of the I arm port will be required to enable the device to function[7]. However, given the assumptions that will be made, this will not alter the noise statistics of the loop. For the analysis here, the mixer output current is given by:-

$$i_m(t) = K_{DI}K_{DQ}m^2(t)/2 \sin 2(\theta_T - \theta_{LO})H_B^2(s) + n_m(t) \qquad \text{----[5.13]}$$

where $n_m(t)$ contains the mixer noise product terms to be described shortly. $i_m(t)$ is then operated upon by the loop filter, F(s), before being fed back to the local oscillator. The output phase of the oscillator, ϕ_{LO}, is given by:-

$$\phi_{LO} = K_o \int i_m(t).F(s)dt$$

$$= K_oF(s)i_m(t)/s \qquad \text{----[5.14]}$$

Substituting for $i_m(t)$ in [5.14] and linearising for small θ_E yields:-

$$\phi_{LO} = K_oF(s)[\ K_{DI}K_{DQ}(\theta_T-\theta_{LO})H_B^2(s) + n_m(t)\ ]/s \qquad \text{----[5.15]}$$

Let $\phi_E = \phi_T-\phi_{LO}$, $\nu = \nu_T-\nu_{LO}$, and $H_B(s)=K_B$, a pure gain term, an assumption which is justified in light of the previous discussion on $H_B(s)$. Then [5.15] becomes:-

$$\phi_{LO} = \frac{K_oF(s)K_{DI}K_{DQ}K_B^2\phi_E}{s} + \frac{K_oF(s)K_{DI}K_{DQ}K_B^2\nu}{s} + \frac{K_oF(s)}{s}\ n_m(t)$$

$$\text{----[5.16]}$$

If the total phase error, θ_E, is now defined as:-

$$\theta_E = \phi_E + \nu = \phi_T - \phi_{LO} + \nu \qquad \text{----[5.17]}$$

and letting $K = K_{DI}K_{DQ}K_B^2$, then substituting [5.16] into [5.17] and rearranging, an expression for θ_E in terms of ϕ_T and noise terms can be obtained:-

$$\theta_E = \frac{\phi_T.s}{s + K_o.K.F(s)} + \frac{\nu.s}{s + K_o.K.F(s)} - \frac{n_m(t).K_o.F(s)}{s + K_o.K.F(s)}$$

$$\text{----[5.18]}$$

It is now possible to define a loop transfer function, H(s),

describing the action of the approximated linear phase-locked loop:-

$$H(s) = \frac{K.K_O.F(s)}{s + K.K_O.F(s)} \qquad \text{----[5.19]}$$

[5.18] can now be expressed in terms of the loop transfer function, H(s):-

$$\emptyset_E = (1 - H(s))\emptyset_T + (1 - H(s))\emptyset - \frac{H(s)}{K} n_m(t) \qquad \text{----[5.20]}$$

Thus the total system phase error has been described in terms of an input signal phase plus two noise terms representing the combined (beat) laser phase noise, and received shot noise. The performance of the system under the influence of each of these perturbations can now be determined.

5.3.3. Derivation of Phase Error Variance.

The parameter used to quantify the noise response of the system is the phase error variance, σ_E^2. For an ergodic process, x(t), the phase error variance is described as:-

$$\sigma_x^2 = E\{x^2(t)\} = R_x(0) \qquad \text{----[5.21]}$$

where E{.} denotes the mathematical expectation process, and $R_x(0)$ is the autocorrelation function evaluated at $\tau = 0$. The object of this section is to determine the magnitude of the signal power required to maintain a set performance in the presence of a non zero phase error variance. This power level is commonly expressed

as a power penalty excess over the theoretical minimum required to obtain that performance (section 2.4) (P_T-P_{TH}).

It is firstly assumed that in [5.20], $\emptyset_T$ is slowly varying, and therefore does contribute significantly to σ_E^2. It can also be seen that the second term is a function only of the system linewidth, while it can be shown that the third term contains both phase and shot noise components. Confining interest to the third term in [5.21], $n_m(t)$ can be expanded from [5.7] and [5.8] as:-

$$n_m(t) = m(t)n_Q(t).K_{DI}.K_B^2.\cos\emptyset_E + m(t)n_I(t).K_{DQ}.K_B^2.\sin\emptyset_E + n_Q(t)n_I(t).K_B^2 \qquad \text{----[5.22]}$$

Again it can be assumed that θ_E is small when the loop is in lock, and hence $\cos\theta_E \sim 1$ and $\sin\theta_E \sim \theta_E$. The variance of [5.22] is obtained by taking the autocorrelation:-

$$R_m(0) = \frac{K_{DI}^2.K_B^4}{K^2} R_Q(t) + \frac{K_{DQ}^2.K_B^4.\theta_E^2}{K^2} R_I(t) + \frac{R_{IQ}(t)}{K^2} \qquad \text{----[5.23]}$$

where $R_I(t)$, $R_Q(t)$, and $R_{IQ}(t)$ are the autocorrelations of $n_I(t)$ and $n_Q(t)$, and the cross-correlation of $n_I(t).n_Q(t)$ respectively. For most practical cases, terms two and three in [5.23] will be negligible in comparison to term one, since $\theta_E^2 << 1$ and $R_{IQ}(t)$ will be small. Thus, dependence of $n_m(t)$ on shot noise alone is established. Since the autocorrelation function and power spectral density form a Fourier transform pair, the variance of $n_m(t)$ can be established in terms of defined PSD's ([5.11] and [5.12]). Substituting the Laplace variable, $s = j2\pi f$, and utilising [5.10] yields:-

$$\sigma_m^2 \ (=\sigma_{SN}^2) = \int_{-\infty}^{\infty} \frac{K_{DI}^2 . K_B^4 . S_Q(0)}{K^2} |H(f)|^2 \, df$$

$$= \frac{e}{4R . P_T \alpha} \int_{-\infty}^{\infty} |H(f)|^2 \, df \qquad \text{——[5.24]}$$

A similar treatment for the phase noise term, Ø, in [5.20] yields an expression for its variance of:-

$$\sigma_{PN}^2 = \frac{\delta f_{TOT}}{2\pi} \int_{-\infty}^{\infty} \left| \frac{1 - H(s)}{f} \right|^2 df \qquad \text{——[5.25]}$$

where δf_{TOT} $(=\delta f_T + \delta f_{LO})$ is the FWHM of the beat signal at the detectors. An expression for the total phase error variance of the Costas loop now follows:-

$$\sigma_E^2 = \sigma_{PN}^2 + \sigma_{SN}^2$$

$$= \frac{\delta f_{TOT}}{2\pi} \int_{-\infty}^{\infty} \left| \frac{1 - H(s)}{f} \right|^2 df + \frac{e}{4R . P_T \alpha} \int_{-\infty}^{\infty} |H(f)|^2 \, df$$

——[5.26]

Evaluation of this expression will lead to a determination of the loop performance when subjected to varying source linewidths, input powers, and loop filters. The form of the loop filter, F(s), is a major factor in evaluating the integrals in [5.26]. The shot noise bandwidth integral, I_{SN}, is well defined in texts for a number of major filter types[5]. However, the phase noise bandwidth integral, I_{PN}, is peculiar to the analysis of optical phase locked loops and has not thus far been investigated except for a single case[6,7]. Appendix 5A evaluates and tabulates both integrals for three types of loop filter. The most common choice of F(s) is the

first-order active lead-lag filter, producing a second order loop as defined by H(s). This facilitates the greatest manipulation of the loop parameters. From table 5A.2:-

$$I_{SN} = \frac{\omega_n}{4\xi}(1 + 4\xi^2) \qquad \text{----[5.27]}$$

and

$$I_{PN} = \frac{\pi^2}{\xi\omega_n} \qquad \text{----[5.28]}$$

Substituting [5.27] and [5.28] into [5.26] yields:-

$$\sigma_E^2 = \frac{\pi\delta f_{TOT}}{2\xi\omega_n} + \frac{\omega_n e(1 + 4\xi^2)}{16RP_T a\xi} \qquad \text{----[5.29]}$$

[5.29] is a complete analytical solution for the phase error variance in a Costas loop incorporating a lead-lag active filter. By inspection, it can be seen that minimising σ_E^2 will require optimisation since I_{SN} is proportional to ω_n and I_{PN} is proportional to $1/\omega_n$. This will be demonstrated in the following design procedure for loop implementation.

5.3.4. Design Procedure for an Optical Costas Loop.

In designing a Costas loop a number of parameters must be specified from the outset. These are primarily the system bit rate and the desired system performance. Using the data in chapter 2, the minimum theoretical power which must be received at the I (data) arm detector, P_{TH}, to achieve that performance can then be calculated. To ease discussion, a system bit rate of 565 Mbit/sec at a bit error rate (BER) of 10^{-9} will be specified. From sections

2.4.3 and 2.4.4, P_{TH} = -60.2dBm. Obviously, ideal detection is unlikely to be achieved in practice. Firstly, as a result of non-zero phase error variance, some penalty will be paid (section 2.5). In addition, mismatch of the received polarisation vectors of E_T and E_{LO}, plus non-ideal decision circuitry will add to system losses. Reasonable figures for these penalties are 0.5dB each. In the case of phase tracking errors this will result in σ_E^2 = 0.03 rad^2[8]. Adding these to P_{TH} results in a required received power at the I arm detector of P_{DATA} = -59.2dBm.

Additional power is also required in the Q arm for phase tracking. Careful selection of the coupling ratio, α, determines that this power, $P_T\alpha$, is sufficient to maintain the desired phase tracking performance. Thus, from figure 5.2, the total power at the input to the receiver is:-

$$P_T = \frac{P_{DATA}}{(1 - \alpha)} \qquad \text{----[5.30]}$$

Finally, two loop parameters can be specified. For optimum tracking performance, the loop damping factor takes the value of ξ = 0.707. Also, the responsivity of the photodiodes used can be found from manufacturers data sheets. A practical value for a silicon PIN diode is R = 0.85.

It is now appropriate to determine an expression for the optimum value of loop bandwidth. Differentiating [5.29] with respect to ω_n and setting the resultant equal to zero yields:-

$$\omega_{n.opt} = \left[\frac{8R.P_T\alpha.\pi\delta f_{TOT}}{e(1 + 4\xi^2)} \right]^{1/2} \qquad \text{----[5.31]}$$

Substituting [5.31] back into [5.29] gives the corresponding expression for the minimum value of σ_E^2, σ_{EMIN}^2:-

$$\sigma_{EMIN}^2 = \left[\frac{\pi\delta f_{TOT}.e(1 + 4\xi^2)}{8\xi^2.R.P_T\alpha} \right]^{1/2} \qquad \text{----[5.32]}$$

Since σ_E^2 has already been specified, [5.30] can be substituted into [5.32] and, by solving for α, an expression for the required coupling ratio for the specified performance is obtained:-

$$\alpha = \frac{\pi\delta f_{TOT}.e(1 + 4\xi^2)}{8\xi^2.R.P_{DATA}.\sigma_{EMIN}^4 + \pi\delta f_{TOT}.e(1 + 4\xi^2)} \qquad \text{----[5.33]}$$

Thus, by using [5.30], [5.33], and [5.31] in turn, the loop parameters for a given performance can be completely determined.

Table 5.1 shows results for three values of linewidth which would be typical of external grating cavity lasers (20kHz), partially reflective mirror coupled cavities (1MHz) as discussed in chapter 4, and a coupled cavity module used with a free running laser (15MHz).

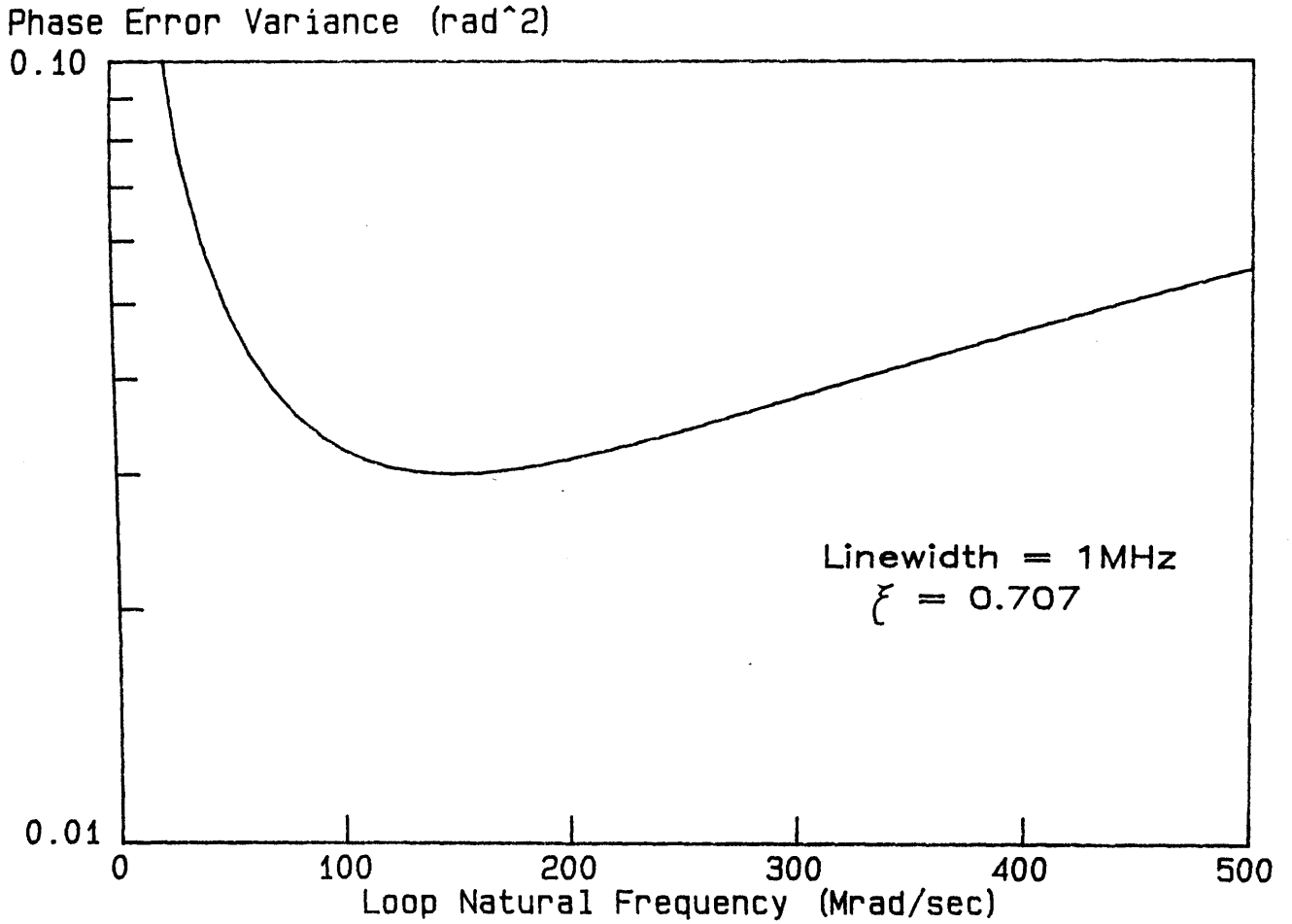
Having now defined all the loop parameters, it is now also possible to plot equation [5.29] to observe what effect a non-optimum loop bandwidth will have on loop performance. Figure 5.3 shows the function plotted for a 1MHz linewidth. From this it can be seen that less of a penalty is incurred if ω_n is too large rather than too small. This indicates that the phase noise term imposes the more stringent requirements on system performance. Hence, the approach of improving system performance by reducing source linewidths is justified.

Linewidth	δf_{TOT} = 20kHz	δf_{TOT} = 1MHz	δf_{TOT} = 15MHz
P_{DATA}	-59.2 dBm	-59.2 dBm	-59.2 dBm
α	8.13×10^{-3}	0.29	0.86
P_T	-59.16dBm	-57.7dBm	-50.7dBm
$W_{n\ opt}$	2.96Mrad/sec	148 Mrad/sec	2.22 Grad/sec
Total System Power Penalty	1.04 dB	2.5 dB	9.5 dB

Table 5.1. Optimum Loop Parameters For Three Given Beat Linewidths

BER = 10^{-9}, f_B = 565 MBits/sec, σ_E^2 = 0.03rad^2, R = 0.85, ζ = 0.707.

Figure 5.3. Phase Error Variance vs Loop Bandwidth for a 1MHz Beat Linewidth.



5.4. Costas Loop Analysis: Discussion and Conclusions.

From table 5.1 it can be seen that large values of beat linewidth give rise to proportionally large values of optimum loop bandwidth, $\omega_{n.opt}$. If the beat linewidth, δf_{TOT}, can be maintained below ~100kHz, then detection sensitivity very close to the quantum limit can be achieved. At 1MHz beat linewidth, the performance obtained is still better than that obtained from heterodyne PSK detection (Appendix 5B). However, the value of ω_n required is large in comparison to that normally attained in PLL's. At a 15MHz beat linewidth, however, the specified value of loop bandwidth would require very high standards of engineering to achieve. It should also be noted that this value is likely to encompass the noise peaks in the laser phase noise spectrum (section 3.2.6), as well as the bit rate, f_B, and the third harmonic of f_B. Performance degradation beyond that specified here is therefore inevitable. Finally, the quoted value of alpha is likely to render the assumptions made in [5.22] invalid. It is probable that term two of [5.22] will no longer be insignificant. As a result, the shot noise term would be dependent on θ_E leading to a significant modification of the analysis. If, on the other hand, a decision driven loop was employed[7], then this problem would be circumvented. The only modification required in the decision driven loop analysis would be the reduction of the dc loop gain, K, by a factor of $1/K_{DI}$.

In spite of these problems, a Costas loop employing a 15MHz beat linewidth should not yet be rejected. If a BER of 10^{-9} is not required, then the loop parameters in table 5.1 can be relaxed somewhat. It is possible to reliably acquire and maintain solid phase-lock with a variance of $0.15rad^2$[5]. Indeed, the linear analysis presented here is valid up to $\sim 0.25rad^2$[9], where phase

locking may still be possible. Without accounting at all for the performance of the link, a loop phase error variance of, for example, 0.15rad^2 can be obtained with $P_T\alpha = -65.3\text{dBm}$, and a loop bandwidth, $\omega_n = 444.3\text{Mrad}$. This value of ω_n is still very high but phase-lock with a smaller value would be very difficult to achieve.

In conclusion, a noise analysis of an optical PRK homodyne receiver has been presented. Equations describing coupling ratio, required received power, and optimum loop bandwidth for a specified performance have been derived. It has been shown that to obtain near quantum limited detection, beat linewidths of less than 100kHz are required. Performance improvements over quantum limited PSK heterodyne detection can be achieved with beat linewidths up to 1MHz. However, beyond 1MHz, BER = 10^{-9} performance ceases to become a practical proposition. Finally, it has been shown that above a 15MHz beat linewidth, obtaining phase-lock without regard to receiver performance is still exceptionally demanding.

APPENDIX 5A: EVALUATION OF NOISE BANDWIDTH INTEGRALS.

This Appendix will undertake the evaluation of the noise bandwidth integrals of the form:-

$$I_{SN} = \int_{-\infty}^{\infty} |H(f)|^2 \, df \qquad \text{----[5.1A]}$$

and

$$I_{PN} = \int_{-\infty}^{\infty} \left| \frac{1 - H(f)}{f} \right|^2 df \qquad \text{----[5.2A]}$$

where H(f) is the frequency domain representation of the loop transfer function, H(s). This transfer function takes the form of:

$$H(s) = \frac{K.F(s)}{s + K.F(s)} \qquad \text{----[5.3A]}$$

where K is some DC gain term and F(s), a filter function. This is then mapped to the frequency plane by the Laplace variable $s = j2\pi f$. The integrals I_{SN} and I_{PN} will be evaluated for three loop types, a first-order passive lag filter, with transfer function:-

$$F_L(s) = \frac{1}{1 + s\tau_1} \qquad \text{----[5.4A]}$$

a first-order passive lead-lag filter, with transfer function:-

$$F_{LL}(s) = \frac{1 + s\tau_2}{1 + s\tau_1} \qquad \text{----[5.5A]}$$

and a first-order active lead-lag filter, with transfer function:-

$$F_A(s) = \frac{1 + s\tau_2}{s\tau_1} \qquad \text{----[5.6A]}$$

Typical implementations of each of these transfer functions are shown in Figures 5A.1a-c. The solution to the passive lag case will be derived first. Substitution of [5.4A] into [5.3A] yields an equation for H(s):-

$$H(s) = \frac{K/\tau_1}{s^2 + s/\tau_1 + K/\tau_1} \qquad \text{----[5.7A]}$$

This equation is of the form:-

$$H(s) = \frac{\omega_n^2}{s^2 + 2\xi\omega_n s + \omega_n^2} \qquad \text{----[5.8A]}$$

where ω_n is the natural frequency of the loop, or loop bandwidth, and ξ is the loop damping factor. For [5.7A], $\omega_n = (K/\tau_1)^{0.5}$ and $\xi = 1/\{2(K.\tau_1)^{0.5}\}$. For most practical cases, ξ will lie in the range $0 =< \xi =< 1$. Optimum tracking performance is obtained when $\xi = 1/2^{0.5}$. From [5.8A], an expression for 1-H(s) in terms of ω_n and ξ is easily obtainable:—

$$1 - H(s) = \frac{s^2 + 2\xi\omega_n s}{s^2 + 2\xi\omega_n s + \omega_n^2} \qquad \text{----[5.9A]}$$

Using the change of variable, $s = j2\pi f$, [5.1A] and [5.2A] become:-

$$I_{SN} = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} |H(s)|^2 \, ds \qquad \text{----[5.10A]}$$

$$I_{PN} = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \left| \frac{1 - H(s)}{s} \right|^2 ds \qquad \text{----[5.11A]}$$

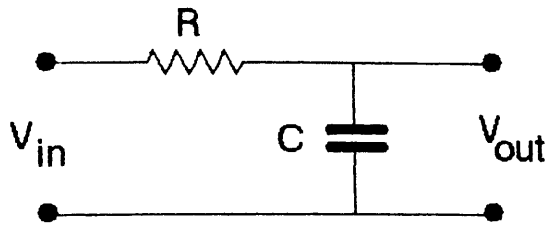
These integrals can be solved with the identities tabulated in table 5A.1[13]. Substituting [5.8A] into [5.10A] yields:-

$$I_{SN} = \frac{\omega_n}{4\xi} = \frac{K}{2} \qquad \text{----[5.12A]}$$

Similarly [5.9A] into [5.11A] yields a solution for I_{PN}:-

$$I_{PN} = \frac{\pi^2}{\xi\omega_n} [4\xi^2 + 1] \qquad \text{----[5.13A]}$$

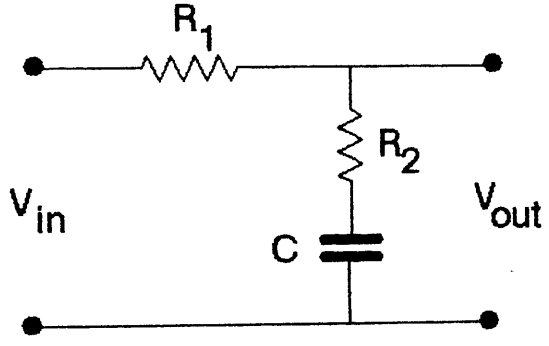
An identical argument is used to derive the solutions to the other filter cases. Results for all three filters are tabulated in table 5A.2. It should be noted from these results that the lead-lag filters will provide better tracking performance by virtue of their ability to alter the damping factor, ξ, independently of ω_n. This gives much greater scope to the loop designer to improve system performance under adverse noise conditions. Note should also be made that the passive lead-lag case will tend to the active case for large values of K.



$$F_L = \frac{1}{1 + s\tau_1},$$

$$\tau_1 = RC.$$

a) A Passive First – Order Lag Filter

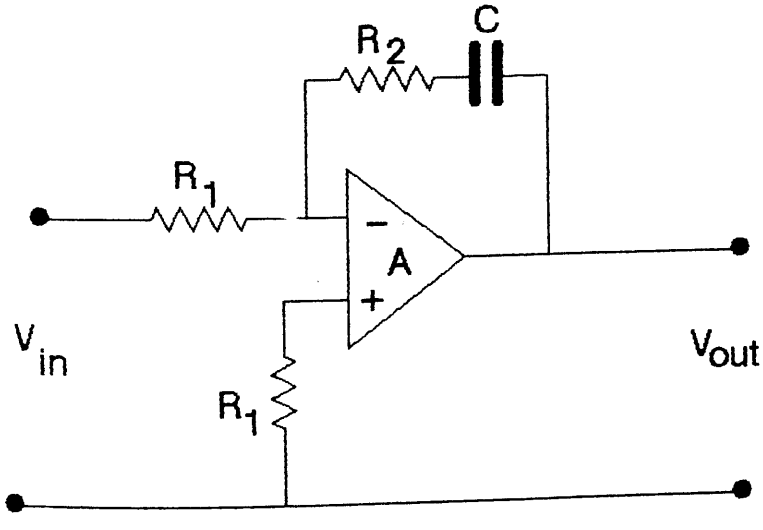


$$F_{LL} = \frac{1 + s\tau_2}{1 + s\tau_1},$$

$$\tau_1 = (R_1 + R_2)\,C,$$

$$\tau_2 = R_2C.$$

b) A Passive First – Order Lead-Lag Filter



$$F_A = \frac{1 + s\tau_2}{s\tau_1},$$

$$\tau_1 = R_1C,$$

$$\tau_2 = R_2C.$$

c) An Active First – Order Filter.

Figure 5A.1.

$$I_n = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} H_n(s).H_n(-s)\, ds$$

$$H_n(s) = \frac{c_0 + c_1 s + c_2 s^2 + ... + c_{n-1} s^{n-1}}{d_0 + d_1 s + d_2 s^2 + ... + d_{n-1} s^{n-1}}$$

$$I_1 = \frac{c_0^2}{2.d_0.d_1} \qquad I_2 = \frac{c_0^2.d_2 + c_1^2.d_0}{2.d_0.d_1.d_2}$$

$$I_3 = \frac{c_2^2.d_0.d_1 + (c_1^2 - 2.c_0.c_2)d_0.d_2 + c_0^2.d_2.d_3}{2.d_0.d_3.(d_1.d_2 - d_0.d_3)}$$

Table 5A.1. Table of Integral Solutions.

$F(s)$	$H(s)$	$1 - H(s)$	ω_n^2	ξ	I_{SN}	I_{PN}
$\frac{1}{1 + s\tau_1}$	$\frac{\omega_n^2}{s^2 + 2\xi\omega_n s + \omega_n^2}$	$\frac{s^2 + 2\xi\omega_n}{s^2 + 2\xi\omega_n s + \omega_n^2}$	$\frac{K}{\tau_1}$	$\frac{1}{2\,K.\tau_1}$	$\frac{K}{2}$	$\frac{\pi_2}{\xi\omega_n}\,[4\xi^2 + 1]$
$\frac{1 + s\tau_2}{1 + s\tau_1}$	$\frac{\omega_n^2 + \omega_n^2.\tau_2.s}{s^2 + 2\xi\omega_n s + \omega_n^2}$	$\frac{s^2 + s(2\xi\omega_n - \omega_n^2.\tau_2)}{s^2 + 2\xi\omega_n s + \omega_n^2}$	$\frac{K}{\tau_1}$	$\frac{1}{2\omega_n\tau_1} + \frac{\omega_n.\tau_2}{2}$	$\frac{\omega_n}{4\xi}\,[1 + (2\xi - \frac{\tau_1}{K})^2]$	$\pi^2\,\left[\frac{4\xi_2 + 1}{\xi\omega_n} + \frac{\tau_2(\omega_n\tau_2 - 4\xi)}{\xi}\right]$
$\frac{1 + s\tau_2}{s\tau_1}$	$\frac{\omega_n^2 + 2\xi\omega_n s}{s^2 + 2\xi\omega_n s + \omega_n^2}$	$\frac{s^2}{s^2 + 2\xi\omega_n s + \omega_n^2}$	$\frac{K}{\tau_1}$	$\frac{\omega_n\tau_2}{2}$	$\frac{\pi^2}{\xi\omega_n}$	$\frac{\omega_n}{4\xi}\,(1 + 4\xi^2)$

Table 5A.2. Table of Loop Parameters and Noise Bandwidth Integrals for Three Types of First Order Loop Filters.

APPENDIX 5B: THE HETERODYNE PLL.

It is relevant to subsequent experiments that the application of the forgoing noise analysis to heterodyne OPLL's be considered. Figure 5B.1 shows a schematic of the system to be discussed. $E_T(t)$ and $E_{LO}(t)$ will take similar forms to [5.1] and [5.4] except that $\omega_{LO} = \omega_T + \omega_{ref}$, where ω_{ref} is the intermediate frequency of the loop which must be greater than f_B. It can be seen that the loop is very similar to the Q arm of the Costas loop. There is, however, no coupling ratio to consider. Instead, in the PSK case, there must be a residual 'pilot' carrier transmitted with the data signal. Power in the carrier component, P_C, contains no information and takes no part in the data decision process. It is therefore 'wasted' power, or a power penalty. P_C does, however, determine the accuracy of the tracking circuit, as does $P_T\alpha$ in the Costas loop analysis.

The analysis is developed in the same way as was done with the Costas loop. $H_B(s)$ is again chosen to pass frequencies of $\omega_{IF} \pm f_B$ without distortion. Again it is approximated to some gain K_B across the passband. The mixer is assumed to be ideal and the reference source noiseless, both reasonable assumptions in practice. If an actual mixer has a conversion gain figure, this will simply modify the DC loop gain, $K.K_o$. A similar expression to [5.20] can be developed for $\theta_E = \theta_T - \theta_{LO}$:-

$$\theta_E = (1 - H(s))\emptyset_T + (1 - H(s))\emptyset - \frac{n(t).H(s)}{K_D.K_B} \qquad \text{----[5.1B]}$$

where $K_D = 2R.(P_{LO}.P_C)^{0.5}$, $K_D.K_B = K$, and n(t) is a zero mean gaussian distributed random variable representing shot noise from the detection process with double sided PSD :-

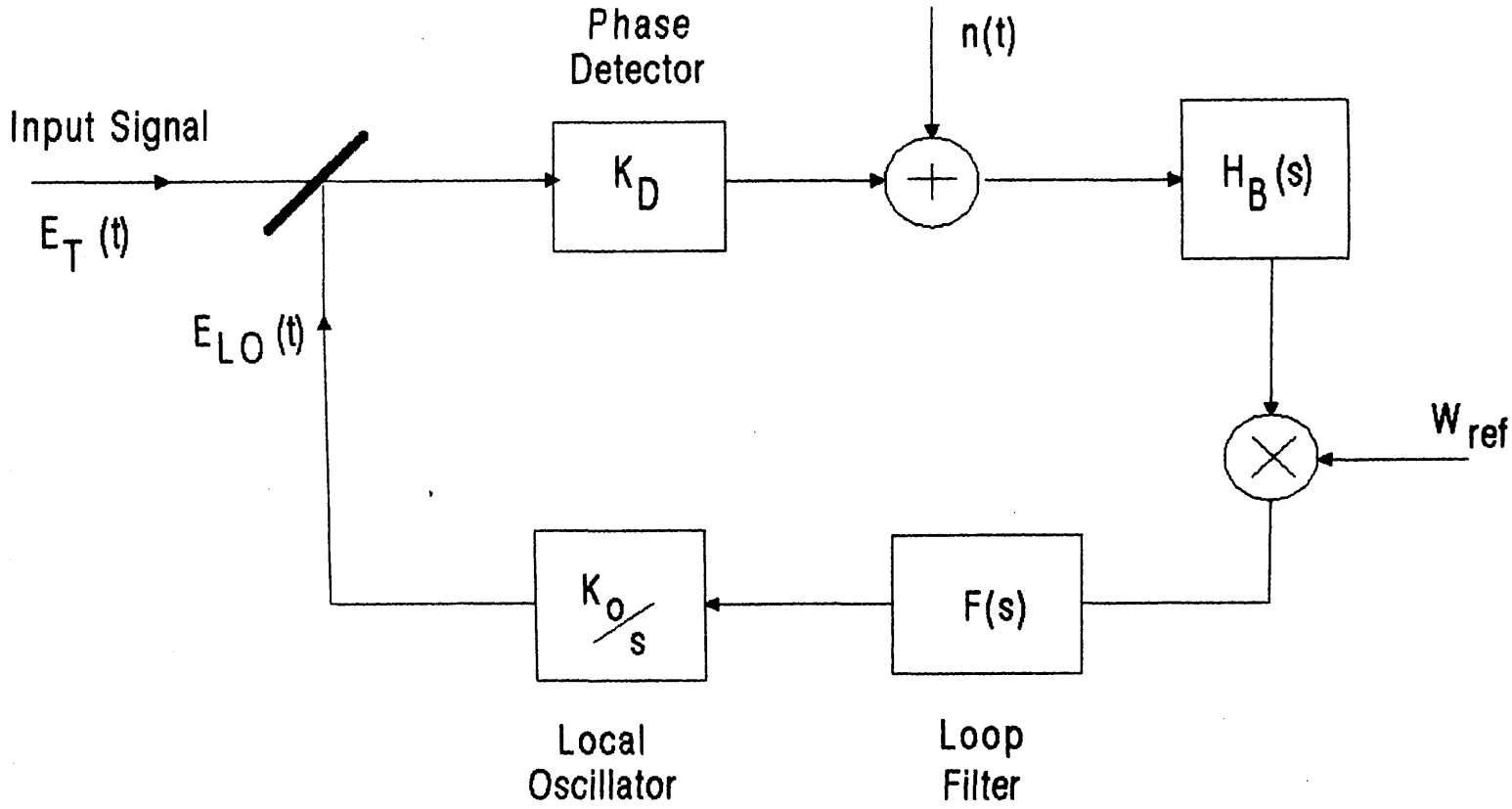


Figure 5B.1. The Optical Heterodyne PLL.

$$S_{SN} = 2.e.R.P_{LO} \qquad \text{----[5.2B]}$$

The factor of two occurs as a result of the shot noise bandwidth being doubled due to detection at an intermediate frequency (section 2.4.5). Again it has been assumed in [5.2B] that $P_{LO} >> P_C$. Taking the Fourier transform of the autocorrelation of the noise process leads to an expression for the phase error variance of the loop:-

$$\sigma_E^2 = \frac{8.\pi.\delta f_{TOT}}{2.\xi.\omega_n} + \frac{e.\omega_n(1 + 4\xi^2)}{8.R.P_C.\xi} \qquad \text{----[5.3B]}$$

By differentiating with respect to ω_n and setting the resultant to zero, an expression for the optimum loop bandwidth is obtained as before:-

$$\omega_{n.opt} = \left[\frac{4.\pi.\delta f_{TOT}.R.P_c}{e.(1 + 4\xi^2)} \right]^{0.5} \qquad \text{----[5.4B]}$$

Substituting back into [5.2B] gives:-

$$\sigma^2_{E.MIN} = \left[\frac{e.(1 + 4\xi^2).\pi.\delta f_{TOT}}{4.R.P_C.\xi^2} \right]^{0.5} \qquad \text{----[5.5B]}$$

By specifying $\sigma^2_{E.MIN}$ to obtain a given loop performance with a given beat linewidth, P_C can be determined. Using the same parameters as before, P_C = -60.04dBm for a 1MHz linewidth. The total input power at the receiver, P_T, is given by:-

$$P_T = P_{DATA} + P_C \qquad \text{----[5.6B]}$$

Thus, by specifying a 1dB penalty over the theoretical minimum, P_{TH}, as before, a total power penalty, $P_T - P_{TH}$, can be determined. For PSK heterodyne detection (chapter 2), $P_{TH} = -57.2$ dBm. From [5.5B], $P_T = -54.7$dBm. Hence, the total power penalty is 2.5dB, identical to the Costas loop. Moreover, by substituting P_C into [5.3B], $\omega_{n.opt} = 148$Mrad for $\delta f_{TOT} = 1$MHz is obtained, as expected. It can therefore be seen that optimisation and performance degradation in a heterodyne PSK loop is identical to the case of the homodyne Costas loop. This is true only when the same loop filter function is used in both systems.

References.

1. D.J.Malyon 'Digital Fibre Transmission using Optical Homodyne Detection' Elect. Lett. vol 20(7) 1984. pp281-283.

2. L.G.Kazovsky 'Balanced Phase-Locked Loops for Optical Homodyne Receivers: Performance Analysis, Design Considerations, and Laser Linewidth Requirements' J. Lightwave. Technol. vol LT-4(2) 1986. pp182-194.

3. J.P.Costas 'Synchronous Communications' Proc. IRE, vol 44, 1956. pp1713-1718.

4. W.R.Leeb 'Realization of 90^{o} and 180^{o} Hybrids for Optical Frequencies' Electron. Commun. AEU. vol 37 May-June 1983. pp203-206.

5. F.M.Gardner 'Phaselock Techniques' New York: Wiley, 1979.

6. A.Blanchard 'Phase-Locked Loops. Application to Coherent Receiver Design.' New York: Wiley, 1976.

7. L.G.Kazovsky 'Decision-Driven Phase-Locked Loop for Optical Homodyne Receivers: Performance Analysis and Laser Linewidth Requirements' J. Lightwave Technol. vol LT-3(6) 1985. pp1238-1247.

8. B.Glance 'Performance of Homodyne Detection of Binary PSK Optical Signals' J. Lightwave Technol. vol LT-4(2) 1986. pp228-235.

9. T.G.Hodgkinson 'Costas Loop Analysis for Optical Receivers' Elect. Lett. vol 22(7) 1986. pp394-396.

10. T.G.Hodgkinson 'Phase-Locked Loop Analysis for Pilot Carrier Coherent Optical Receivers' Elect. Lett. vol 21 1985. pp1202-1203.

11. J.J.Stiffler 'Theory of Synchronous Communications' Prentice-Hall Inc. 1971.

12. F.G.Stremler 'Introduction to Communication Systems' 2nd Ed. Addison-Wesely Inc. 1982.

13. W.C.Lindsey 'Synchronous Systems in Communication and Control' Englewood Cliffs, NJ: Prentice Hall, 1972.

CHAPTER 6.

THE EFFECTS OF PROPAGATION DELAY ON LOOP PERFORMANCE.

Introduction.

Propagation delay is not a factor that is usually considered in phase-locked loop design. In most practical cases, the transit time around the loop is only a small fraction of its maximum response time constant ($2\pi/\omega_n$). However, optical phase-locked loops present unique problems in that the loop bandwidths specified can be considerably higher than is normally encountered in conventional microwave designs. Moreover, the permitted deviation from optimum values of ω_n, especially towards smaller values, is minimal before severe performance degradation is encountered. Indeed, as was shown in chapter 5, only marginal phase-locked conditions can be achieved with relatively large values of ω_n when beat linewidths of >10MHz are considered. It is therefore important that no further unspecified degradation occurs when attempting to implement these systems.

When values of ω_n, such as those calculated in chapter 5, are required, it is then highly likely that a practical system will have a transit time which is significant in terms of loop bandwidth. This chapter evaluates a loop model which incorporates a time delay factor. The resultant modification of the loop transfer function, H(s), is given in section 6.2, and the modulus of the appropriate terms derived in order to allow the modified noise bandwidth integrals to be evaluated. A technique for solving these integrals is discussed in section 6.3. Section 6.4 describes the application of these results to a Costas loop, incorporating the previously specified values of beat linewidth. The significance of these results is discussed in section 6.5 and conclusions drawn. A listing of the program used to evaluate the

performance of a Costas loop in the presence of non-negligible loop propagation delay is given in appendix 6A.

6.2. Phase Locked Loop Model Incorporating Time Delay.

The model of the Costas loop with non-negligible propagation delay relies, as does the original model, on the system being linear. A necessary condition for linearity has been stated as requiring $\sigma_E^2 < 0.25\text{rad}^2$[1]. With this assumption, the various components that make up the loop delay can be lumped together and modelled as one term. The Laplace transform of a time translation, τ_D, is given as:-

$$L\{ \delta(t - \tau_D) \} = e^{-s\tau_D} \qquad \text{----[6.1]}$$

where $\delta(.)$ is the Dirac delta function. Using this transformation, the loop model can be modified as shown in figure 6.1. The loop transfer function, H(s), then becomes:-

$$H(s) = \frac{K_o.K.F(s).e^{-s\tau_D}}{s + K_o.K.F(s).e^{-s\tau_D}} \qquad \text{----[6.2]}$$

Evaluation of the shot noise bandwidth integral, I_{SN}, and the phase noise bandwidth integral, I_{PN}, using [6.2] is necessary before the effect of time delay on the phase error variance of the loop [5.26] can be determined. The first step in this process is to determine $|H(s)|^2$ and $|(1 - H(s))/f|^2$. However, the inclusion of the complex exponential term rules out immediate use of the integral identity given in appendix 5A. It also makes any alternative analytical solution very difficult to find.

It is possible to substitute the first two terms of a

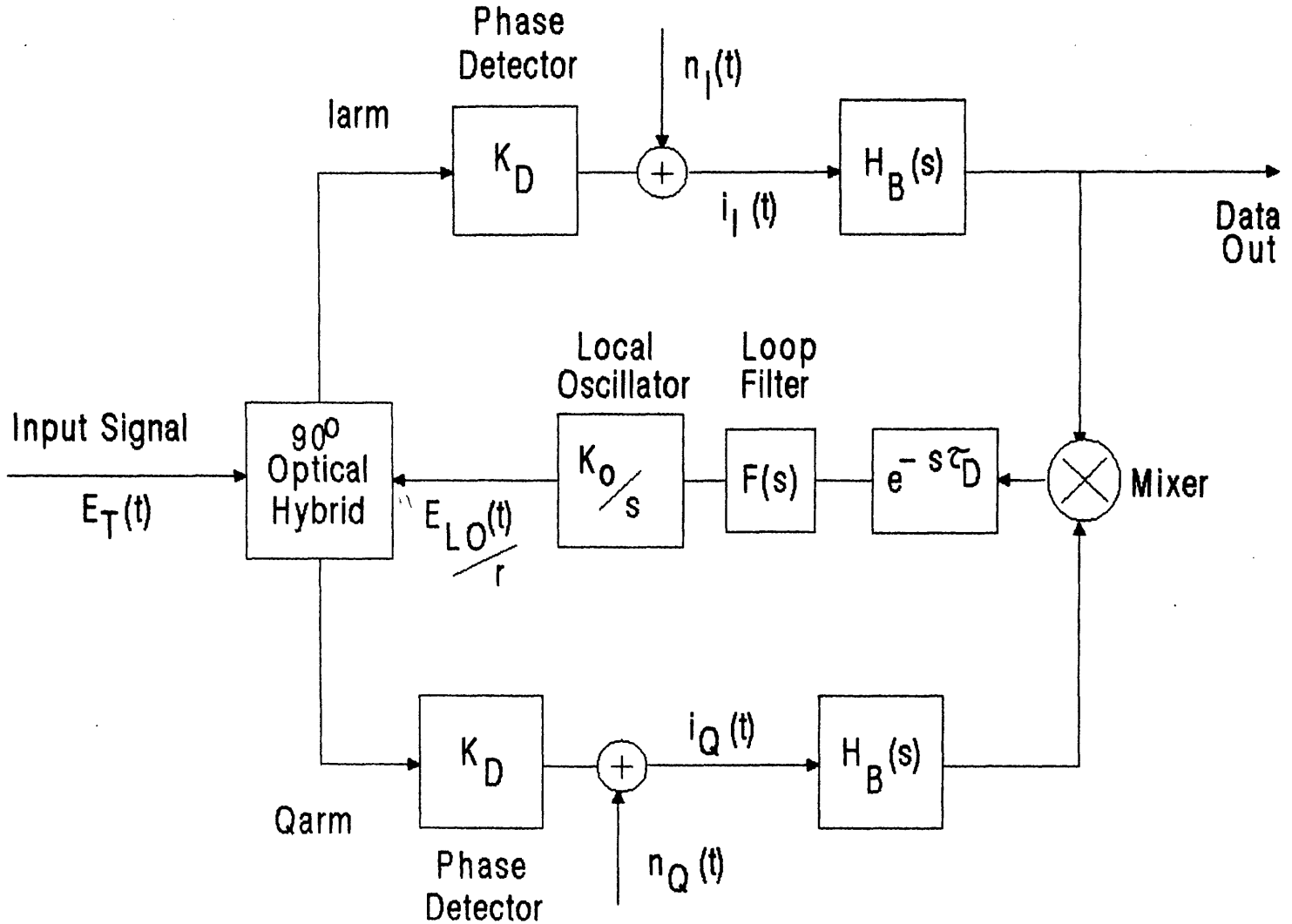


Figure 6.1. Schematic of the Optical Costas Loop with Time Delay.

r^2 = Power Reflection Coefficient of the Beam Combiner in the 90^o Hybrid

binomial expansion of e^{-x} into [6.2], thus returning H(s) to a polynomial form. However, for $s\tau_D > 0.1$ the errors associated with this approximation become significant and , since this is the area of interest in the investigation, this particular substitution was rejected. The inclusion of higher order terms of the expansion can also be considered. However, it was found that higher order solutions to the integral identity in table 5A.1 are not easily obtainable. As a result the binomial expansion approach was discarded.

An alternative to a binomial expansion of $e^{-s\tau D}$ was to expand H(s) in terms of complex parts. This method was eventually to lead to a solution. To ease discussion, a specific form of F(s) is introduced at this point and the expressions derived in terms of loop bandwidth and damping factor. As in the analytical analysis presented in chapter 5, the loop filter is chosen to be an active lead-lag filter. Hence, F(s) takes the form (table 5A.2):-

$$F(s) = \frac{1 + s\tau_2}{s\tau_1} \qquad \text{----[6.3]}$$

Substitution of [6.3] into [6.2] realises the integrand of [5.1A] as:-

$$|H(s)|^2 = \left| \frac{K_o.K.(1 + s\tau_2).e^{-s\tau_D}}{s^2\tau_1 + K_o.K.(1 + s\tau_2).e^{-s\tau_D}} \right|^2 \qquad \text{----[6.4]}$$

Substituting $s = j\omega$ into [6.4] yields:-

$$|H(j\omega)|^2 = \left| \frac{K_o.K.(1 + j\omega\tau_2)(\cos\omega\tau_D - j\sin\omega\tau_D)}{K_o.K.(1 + j\omega\tau_2)(\cos\omega\tau_D - j\sin\omega\tau_D) - \omega^2\tau_1} \right|^2$$

----[6.5]

Given that $|H(j\omega)|^2 = Re\{H(j\omega)\}^2 + Im\{H(j\omega)\}^2$, [6.5] can be simplified by separating the equation into real and imaginary parts. Thus, with some algebraic manipulation:-

$$|H(j\omega)|^2 = \frac{(K.K_o)^2(1 + \omega^2\tau_2^2)}{\omega^4\tau_1^2 + (K.K_o)^2(1 + \omega^2\tau_2^2) - 2\omega^2\tau_1.K.K_o(\cos\omega\tau_D + \omega\tau_2\sin\omega\tau_D)}$$

----[6.6]

From table 5A.2 the following substitutions can be made to express the noise bandwidth integral in terms of loop bandwidth and damping factor:-

$$\omega_n = \left[\frac{K.K_o}{\tau_1} \right]^{1/2} \qquad \text{----[6.7]}$$

and:-

$$\xi = \tau_2.\omega_n / 2 \qquad \text{----[6.8]}$$

Further, by use of the change of variable $f = \omega/2\pi$, I_{SN} can be given as:-

$$I_{SN} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\omega_n^4(1 + \omega^2\tau_2^2)}{\omega^4 + \omega_n^4(1 + \omega^2\tau_2^2) - 2\omega^2\omega_n^2(\cos\omega\tau_D + \omega\tau_2\sin\omega\tau_D)} d\omega$$

----[6.9]

The integrand of [6.9] is a monotonically decreasing function allowing the integral to be evaluated by numerical techniques. This evaluation will be described in the next section.

A similar procedure can also be followed for I_{PN}. This time $|(1 - H(s))/f|^2$ is evaluated yielding:-

$$I_{PN} = 2\pi \int_{-\infty}^{\infty} \frac{\omega^2}{\omega^4 + \omega_n^4(1 + \omega^2\tau_2^2) - 2\omega^2\omega_n^2(\cos\omega\tau_D + \omega\tau_2\sin\omega\tau_D)} d\omega \qquad \text{----[6.10]}$$

In the linear model described in chapter 5, it is readily established that H(s) is unconditionally stable. It is evident from [6.9] and [6.10] that this is not the case for a model which includes propagation delay. The conditions for stability of [6.9] and [6.10] require to be determined before the analysis can proceed further. This can be done using standard Nyquist techniques.

The open loop transfer function of the system, $G(j\omega)$, is found to be:-

$$G(j\omega) = \frac{\omega_n^2(1 + j\omega\tau_2)(\cos\omega\tau_D - j\sin\omega\tau_D)}{-\omega^2} \qquad \text{----[6.11]}$$

The argument of this function is given by:-

$$\text{Arg}[G(j\omega)] = \arctan(\omega\tau_2) - \omega\tau_D - \pi \qquad \text{----[6.12]}$$

and the magnitude by:-

$$|G(j\omega)| = \omega_n^2(1 + (\omega\tau_2)^2)^{1/2} / \omega^2 \qquad \text{----[6.13]}$$

For stability of the closed loop transfer function, it is required that for $Arg[G(j\omega)] = (2k-1)\pi$, $k = (0,1,2...\infty)$, the magnitude of the open loop function, $|G(j\omega)|$ be less than 1. The critical solution is the case of k =0, giving:-

$$\omega\tau_D = \arctan\omega\tau_2 \qquad \text{----[6.14]}$$

and

$$\omega_n^2(1 + (\omega\tau_2)^2)^{1/2} / \omega^2 < 1 \qquad \text{----[6.15]}$$

Solving these equations numerically, an upper bound on the product $\omega_n\tau_D$ can be found. Assuming $\xi = 1/2^{0.5}$, $\omega_n\tau_D$ was found to satisfy $\omega_n\tau_D <= 0.736509$ for absolute stability. Lindsey[2] has derived an analytical solution for a similar set of equations and obtained a result consistent with the above findings.

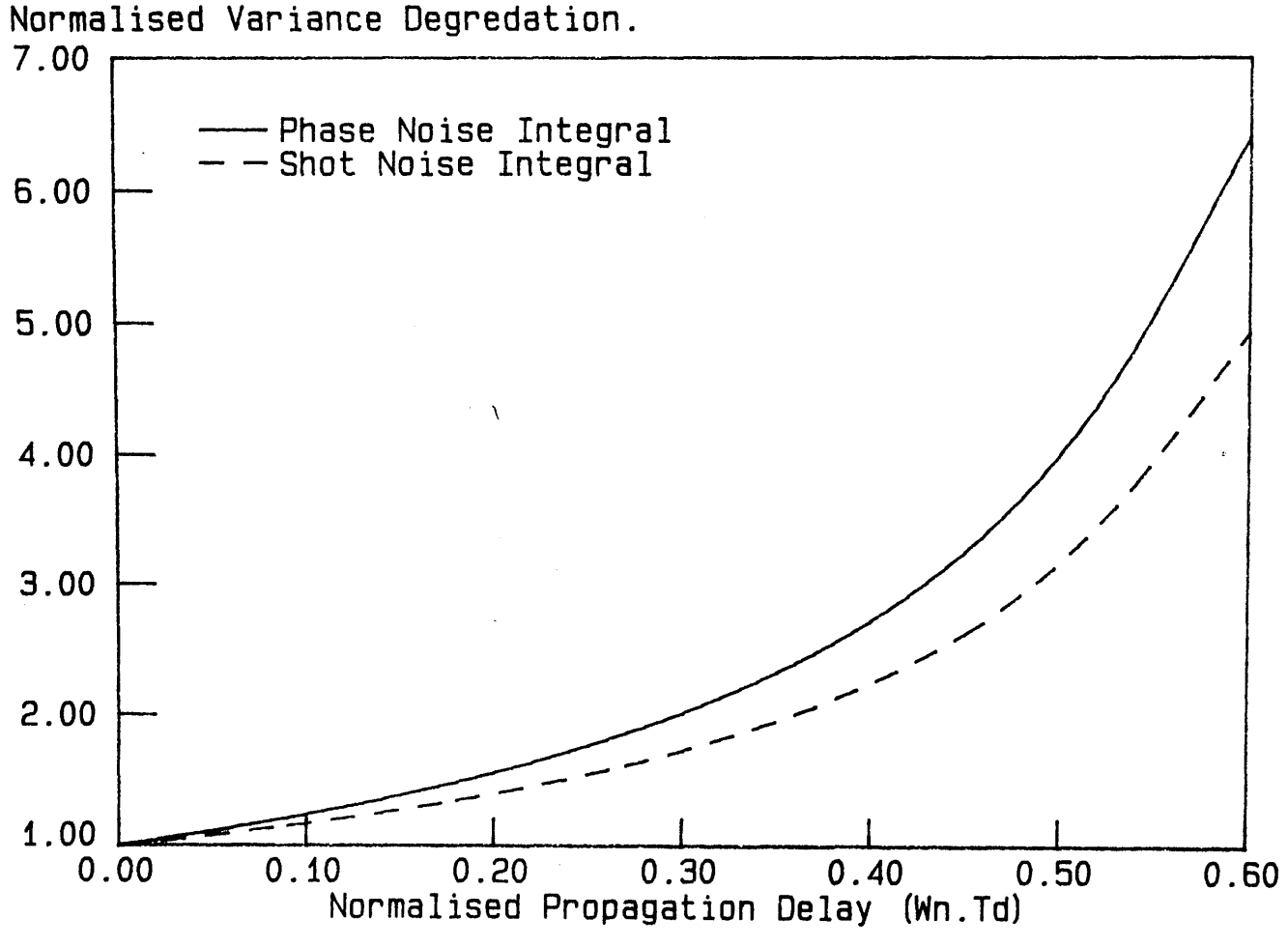
6.3. Numerical Solution for the Noise Bandwidth Integrals.

The method chosen to integrate [6.9] and [6.10] was a simple trapezoidal rule of the form:-

$$I = \left[\frac{I_1 + I_2}{2} \right] x\ \Delta\omega \qquad \text{----[6.16]}$$

where I_1 and I_2 are values of the integrand at ω_1 and ω_2, and where $\Delta\omega = \omega_2 - \omega_1$ ($\omega_2 > \omega_1$). Since only positive values of ω were used in the integration algorithm, the value of the integral is twice that calculated from [6.9] and [6.10]. The integration procedure was written in Pascal and implemented on an IBM PC AT computer which incorporated an 80287 maths co-processor for increased speed of calculation. The algorithm forms the major part

Figure 6.2. Normalised Plots of the Noise Bandwidth Integrals, Isn and Ipn.



of the program MCOSTAS.PAS, which is listed in appendix 6A. This program evaluates the total system phase error variance for a given time delay (section 6.4). The integration parameters were carefully selected in order to minimise error. The integration increment, $\Delta\omega$, was chosen as $\omega_n/100$, and the integration limit, $50\omega_n$. At this limit, the value of both functions was 10^6 down on their values at ω = 1Mrad. No significant changes in the integral values were found when smaller integration increments and larger limits were investigated. The above values were found empirically to be optimum for accuracy and computation time.

Initially, I_{PN} and I_{SN} were calculated for fixed ω_n and varying τ_D. Normalised plots of these integrals are shown in figure 6.2. It can be seen that there is a rapidly accelerating degradation in respective variance values as these functions approach instability at $\omega_n\tau_D = 0.736$. It was shown in figure 5.3 that the variance degradation need not be great before a BER of 10^{-9} will not be achieved. Given the large values of optimum loop bandwidth calculated in chapter 5, it can therefore be expected that only small values of loop delay will be tolerated before a significant degradation in performance results. The following section will investigate the application of these findings to a Costas loop system.

6.4. Application to the System Phase Error Variance.

Having established a valid integration procedure for the equivalent noise bandwidth integrals, it is now instructive to solve the complete system phase error variance expression, [5.26], and attempt to obtain a modified optimum loop bandwidth for various values of beat linewidth and time delay.

Beat linewidths of 1MHz and 15MHz were chosen for study as

before. Using the program MCOSTAS.PAS, listed in appendix 6A, plots of ω_n versus σ_E^2 were obtained for three different values of τ_D. These are shown in figures 6.3 and 6.4[3]. The parameters used in this procedure are given in table 6.1 , with values of ξ and P_T as derived in chapter 5. From these plots it can be seen that, as expected, a significant degradation in performance takes place for small values of τ_D. More specifically, it can be seen that a 10^{-9} BER will not be attained if τ_D exceeds 1.8ns for a 1MHz and 0.12ns for a 15MHz beat linewidth respectively.

Figures 6.3 and 6.4 describe absolute limits to the systems phase tracking capability. Phase tracking performance can not be improved by increasing the received power in the phase tracking arm. Therefore, the power penalties associated with this imperfect phase tracking must be compensated for in the data detection arm of the loop. The magnitude of this penalty is calculated by evaluation of the probability distribution given for non-zero σ_ϕ^2 in section 2.5[4] and equating σ_ϕ^2 to σ_E^2. By plotting $\sigma_{E.MIN}^2$ vs τ_D for each linewidth, it is also possible to observe the region in which reliable acquisition of phase-lock can be achieved. In chapter 5, it was stated that σ_E^2 should be less than 0.15 for this condition to be satisfied. This bound is shown in figures 6.5 and 6.6. From these figures it can be seen that τ_D must be <= 10ns for a 1MHz beat linewidth, and <= 0.67ns for a 15 MHz beat linewidth in order that solid phase-lock be realised.

6.5. The Effects of Loop Propagation Delay: Discussion and Conclusions.

It has been shown that inclusion of propagation delay in the noise analysis of the optical phase-locked loop can have a dramatic effect on loop performance. The system BER is likely to

Beat Linewidth δf_{TOT}	ζ	α	P_T	Integration Limit ($W_{n.max}$)	Integration Step	Loop Bandwidth Increment
1 MHz	0.707	0.29	1.69 nW	100 W_n	$1x10^6$ Mrad	$1x10^7$ Mrad
15 MHz	0.707	0.86	8.59 nW	100 W_n	$15x10^6$ Mrad	$1.7x10^8$ Mrad

Table 6.1. Table of Integral Parameters used in the Numerical Integration of Equations [6.10] and [6.11].

Figure 6.3 Loop Performance with a 1MHz Linewidth and Non-Negligible Loop Delay.

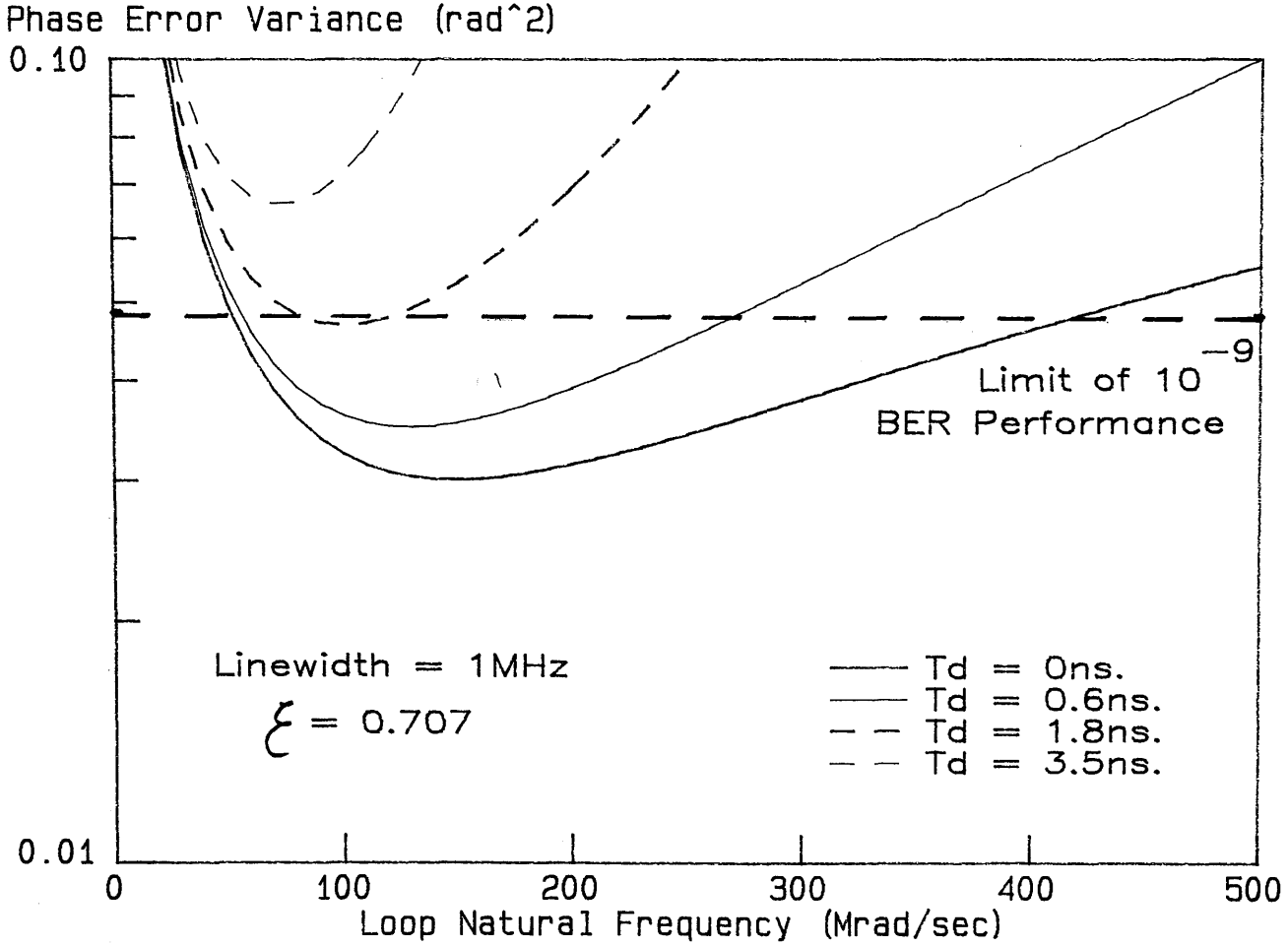


Figure 6.4 Loop Performance with a 15MHz Linewidth and Non-Negligible Loop Delay.

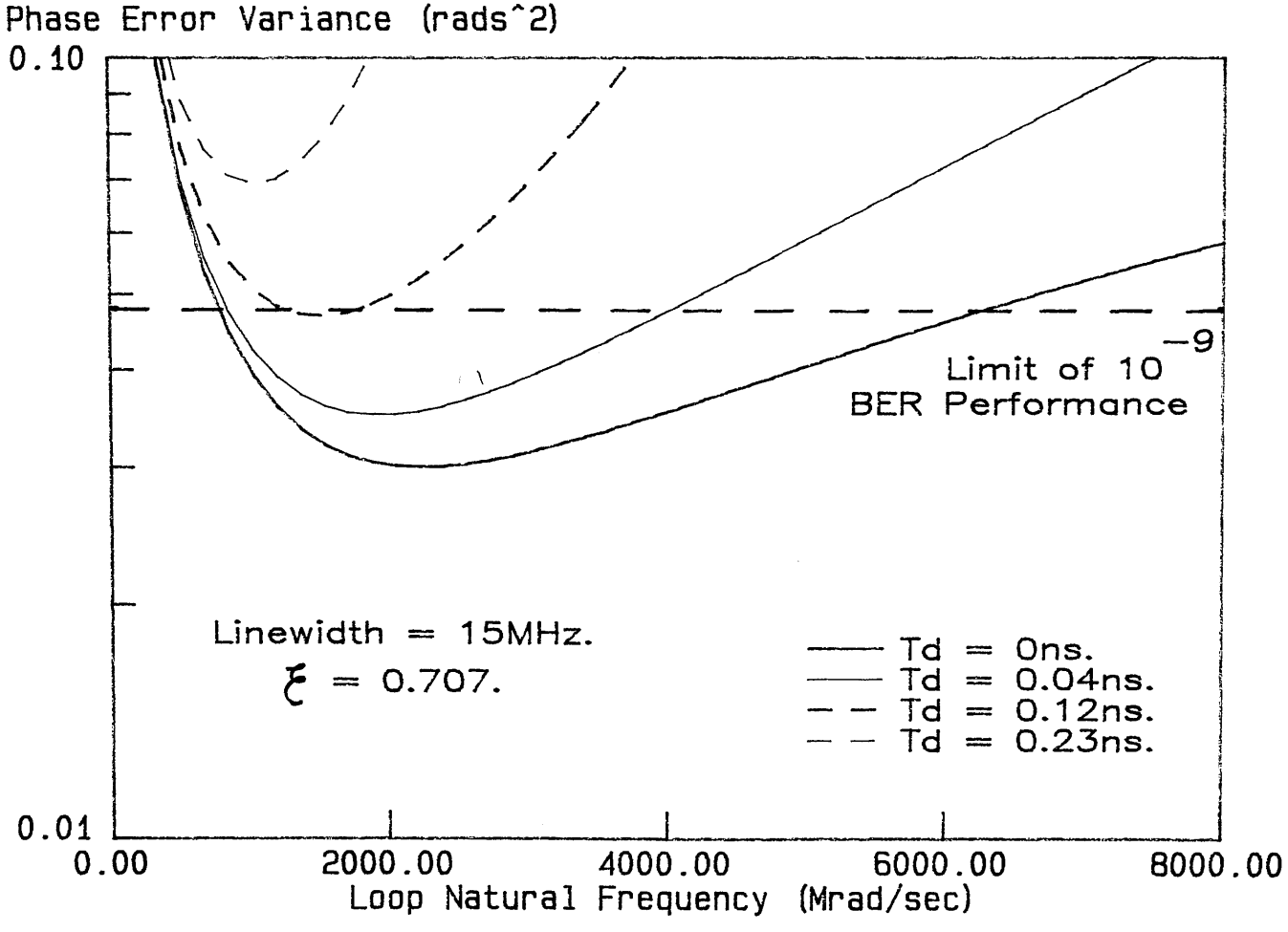


Figure 6.5. Minimum Phase Error Variance vs Loop Delay for a 1MHz Beat Linewidth

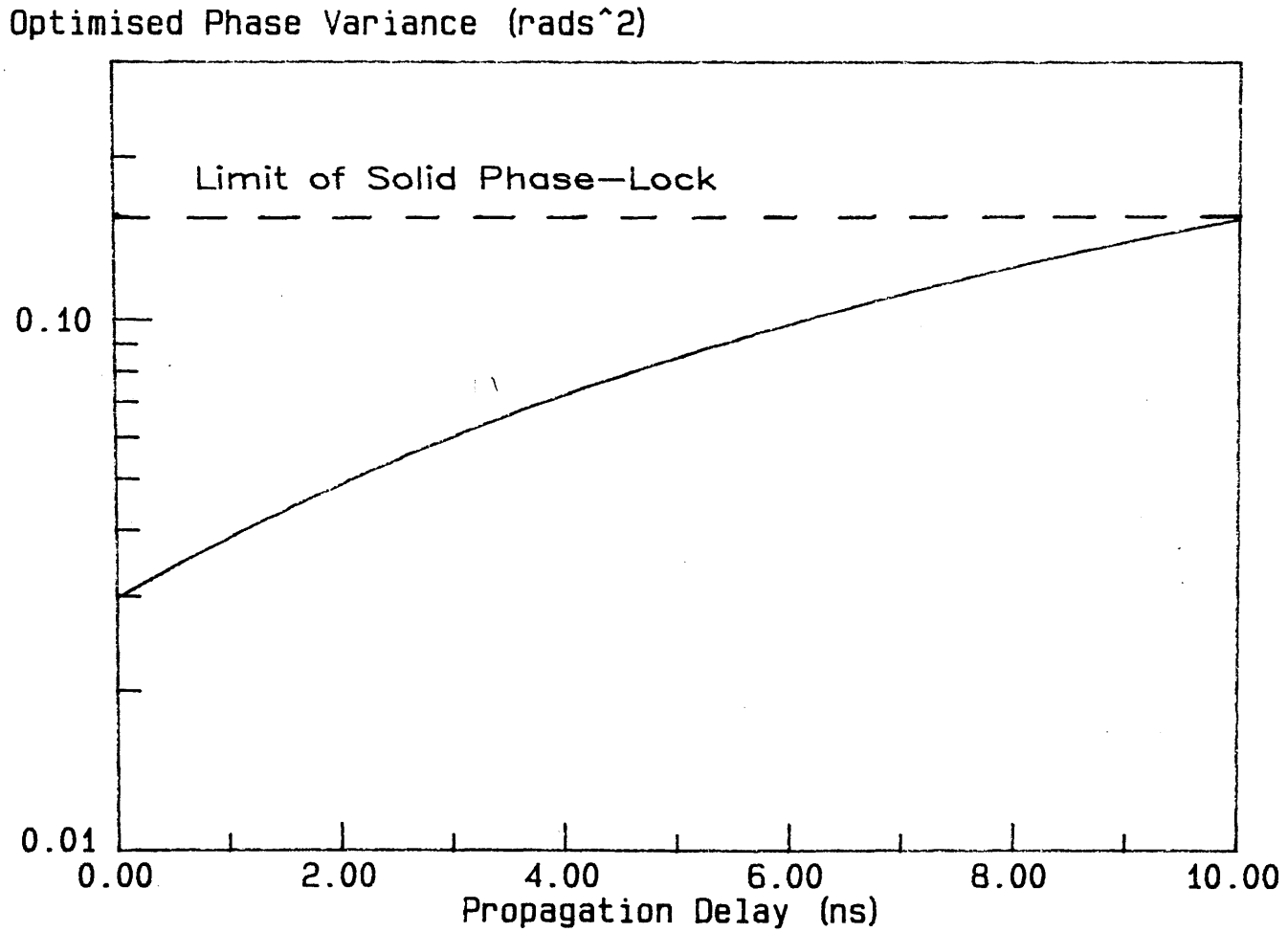
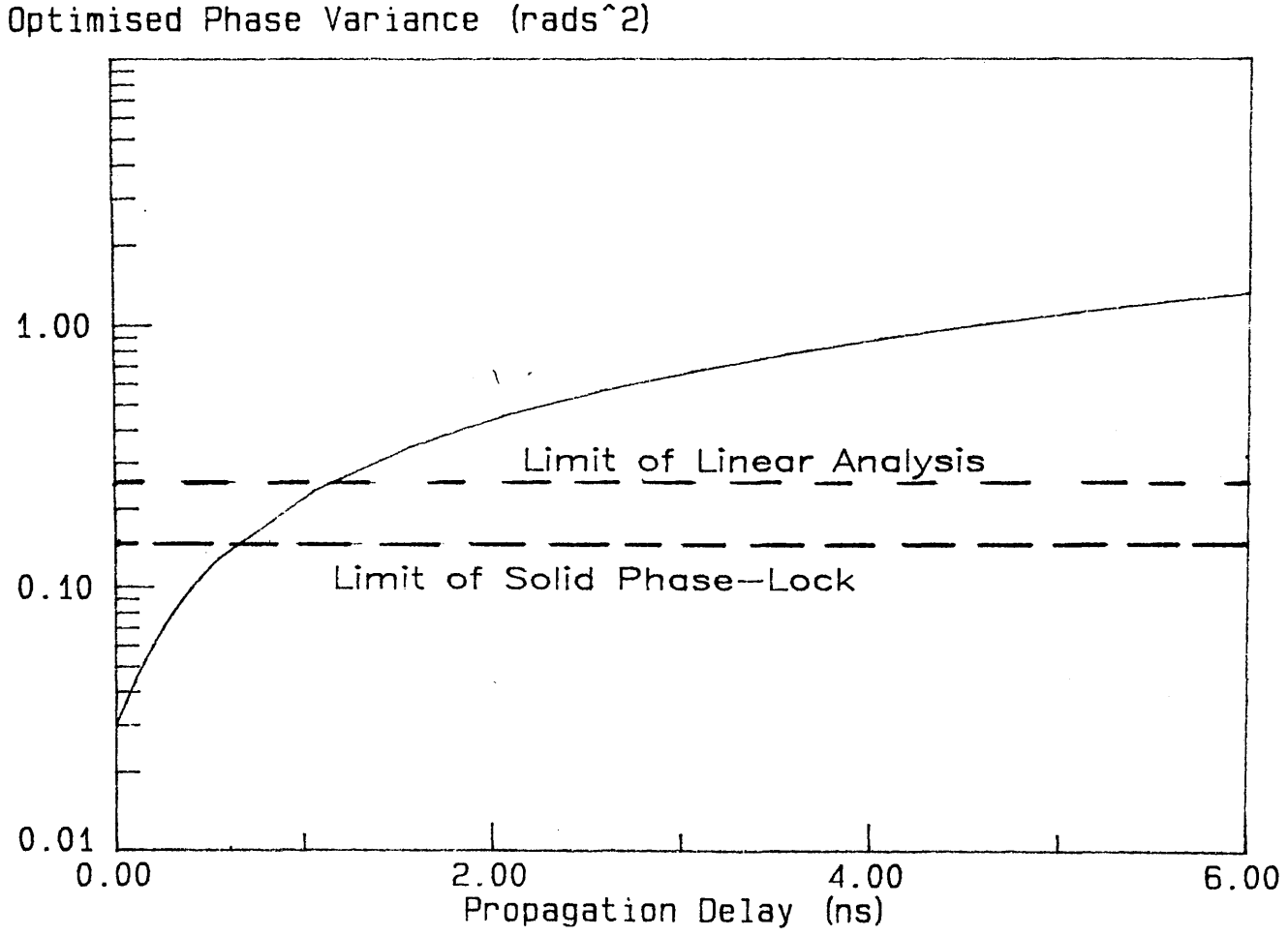


Figure 6.6. Minimum Phase Error Variance vs Loop Delay for a 15MHz Beat Linewidth



be irretrievably degraded below that which is specified. The possibility also exists, especially in the large linewidth cases (>10MHz), that a system which is stable and has achieved a specified performance when analysed without consideration of τ_D, will in reality be unable not only to achieve that performance level, but also to achieve phase-lock. If corrective action is then taken to compensate for propagation delay, there is still the possibility that solid lock will not be achieved. Note that these results and their effects are identical for other types of loop employing an active lead lag filter.

This chapter has quantified the detrimental effects of propagation delay on optical phase-locked loops. A modified loop transfer function incorporating time delay has been given. Techniques for evaluating the resultant noise bandwidth integrals have been discussed and a numerical integration solution implemented. The stability of this modified function has been examined and an upper bound on stability calculated. This condition stated that $\omega_n\tau_D$ must be less than 0.736509. Limits on the allowable system delay time in order to maintain a 10^{-9} BER have been calculated for two values of linewidth. These are τ_D < 1.8ns and τ_D < 0.12ns for 1MHz and 15 MHz beat linewidths respectively. Values for maintaining phase-lock without regard to system performance have also been calculated. These require that τ_D be less than 10ns for a 1MHz beat linewidth and 0.67ns for a 15MHz beat linewidth. The above figures will be difficult to implement in practice. For example, the figure for obtaining lock with a 15MHz beat linewidth requires that the total system path length be less than 6.9/n cm, where n is the average refractive index of the propagation medium.

It should be noted here that this analysis has assumed an

infinite loop filter bandwidth. This, of course, can never be the case in practice. However, recent, and currently unpublished work by Mr W.C.Michie has shown that, provided the cutoff frequency of the filter is greater than $4\omega_n$, the results given here will be correct to within 5%.

In conclusion, the results presented in this chapter indicate that the implementation of a Costas loop with a) a 15MHz beat spectrum is impractical and b) a 1MHz beat linewidth is difficult and is unlikely to yield a satisfactory performance.

APPENDIX 6A: LISTING OF COSTAS LOOP PERFORMANCE EVALUATION PROGRAM

The evaluation of the performance of an optical Costas loop influenced by non-negligible phase noise and loop propagation delay is carried out by the program MCOSTAS.PAS. The program is written in Pascal and implemented on an IBM PC AT micro computer.

The main body of the program begins at line 194. The desired value of beat linewidth is set, and four values of time delay (td) chosen. For each value of td, the maximum value which the loop bandwidth is required to take (Wnmax) is set. The procedure varcalc is then called to evaluate the system phase error variance for loop bandwidths, Wn, from Wnstart to Wnmax at intervals of increment.

The procedure varcalc first calculates the DC loop bandwidth from the constants set at the start of the program. A while loop is then entered until the value of phase error variance (sigmasqd[i]) is calculated for each Wn. Should td take the value zero, the integration algorithm is bypassed and an analytical solution calculated (lines 183/4). Otherwise the noise bandwidth integrals are evaluated numerically. The function DENOM(w) calculates the value of the denominator of the integrands, common to both terms, for a given w ($w \equiv \omega$). The functions SUM1(w) and SUM2(w) then evaluate I as given by [6.16] for shot noise and phase noise terms respectively in increments of freqstep. The results are then summed in tot1 and tot2 to give the values of I_{SN} and I_{PN} as derived in equations [6.9] and [6.10]. The value, sigmasqd[i] is then calculated and stored in the two dimensional array, storaged[m,n]. This array holds the set of n data points for the m^{th} td. An array containing n values of Wn corresponding to the stored data points is then calculated. The procedure

filedmp is then used to store the data to file in a form readable by the graphics package, DRGRAPH. This package is then used to display the information in the form seen in figures 6.3-6.6.

```
 1: program mcostas;
 2: {$R+}
 3: { a program to evaluate system phase error variance for the optical costas
 4: pll for various values of loop natural frequency, Wn. For non-zero values
 5: of loop propagation delay, td, the program numerically evaluates the
 6: noise bandwidth integrals for Wn = Wnstart to Wnmax.}
 7:
 8: type
 9:   axesarray = array[1..72] of real;
10:   dataarray = array[1..4] of axesarray;
11:
12: const
13:   no_elements =72; {no of elements in array for graph outout}
14:   alpha = 0.29;    {coupling ratio set to 0.29}
15:   Kamp = 1.0;      {Front end amp gain}
16:   Km = 1.0;        {Multiplier Gain}
17:   eta = 0.707;     {Damping factor}
18:   Resp = 0.85;     {Nominal manufacturers Responsivity estimate}
19:   Plo = 100.0e-6 ; {local oscillator power in uW}
20:   Pt  = 1.69e-9;   {1.69 nW signal power}
21:   e = 1.602e-19;   {charge on an electron}
22:   Ko =1.88e13;     {LO gain in rads per Amp}
23:   freqstep  = 1e6; {Increment in w for integration is set to 1Mrad/sec}
24:   increment =10e6; {Natural frequency (Wn) increment is set to 10Mrad/sec}
25:   Wnstart = 1e7;   {Starting value of Wn is set to 10Mrad/sec}
26:
27:
28:
29: var
30:    linewidth, T2, K, Kd, td                              :real;
31:    integ1a, integ2a, integ1b, integ2b, denominator       :real;
32:    tot1, tot2, Wn, Wnmax, w                              :real;
33:    omega_n, sigmasqd                                     :axesarray;
34:    I,J,countf,countd                                     :integer;
35:    storaged                                              :dataarray;
36:
37:
38: Function DENOM(w:real):real;              {EVALUATE DENOMINATOR}
39:  begin
40:       DENOM := sqr(sqr(w)) - 2*sqr(w*Wn)*(cos(w*Td) + w*T2*sin(w*Td))
41:                       + sqr(sqr(Wn))*(1 + sqr(w*T2));
42:  end;
43:
44: Function SUM1(w:real):real;              {EVALUATE mod[H(s)]^2 (SHOT
45:                                                    NOISE TERM)}
46:  begin
47:       integ1a := integ1b;
48:       integ1b := sqr(sqr(Wn))*(1+sqr(w*T2))/denominator;
49:       SUM1 := freqstep*(integ1a + integ1b)/2;
50:  end;
51:
52: Function SUM2(w:real):real;              {EVALUATE mod[1-H(s)/f]^2
53:                                                  (PHASE NOISE TERM)}
54:  begin
55:       integ2a := integ2b;
56:       integ2b := sqr(w)*sqr(2*pi)/denominator;
57:       SUM2 := freqstep*(integ2a +integ2b)/2;
58:  end;
59:
60: {$I costdmp}
```

```
 61> procedure filedmp(inparray:dataarray; axes:axesarray);
 62>
 63> {This procedure is used in the program mcostas to format a file for
 64> the DRGRAPH package and dump input data calculated in the main program,
 65> to it. It may obviously be modified to suit other programs as required.}
 66>
 67>
 68> type
 69>      filename = string[13];
 70>
 71> var
 72>     i,j         :integer;
 73>     outputfile :text;
 74>     name        :filename;
 75>     outputvar   :real;
 76>
 77>
 78> begin
 79>      {initialise the file}
 80>      clrscr;
 81>      writeln;
 82>      chdir('\mg\gsx');   {PUT OUTPUT FILE IN CORRECT SUB-DIRECTORY}
 83>      writeln('Input name of file to be created (cost****.dif).');
 84>      readln(name);
 85>      {FILE INITIALISATION}
 86>      assign(outputfile,name);
 87>      rewrite(outputfile);
 88>
 89>      {DATA INITIALISATION FOR SDF FORMAT}
 90>      write(outputfile,'TABLE');
 91>      writeln(outputfile);
 92>      write(outputfile,'0,1');
 93>      writeln(outputfile);
 94>      write(outputfile,'""');
 95>      writeln(outputfile);
 96>      write(outputfile,'DATA');
 97>      writeln(outputfile);
 98>      write(outputfile,'0,0');
 99>      writeln(outputfile);
100>      write(outputfile,'""');
101>      writeln(outputfile);
102>      {END OF SDF FILE INITIALISATION}
103>
104>      {WRITE DATA TO FILE}
105>      for i:= 1 to 4 do
106>          begin
107>               write(outputfile,'-1,0');
108>               writeln(outputfile);
109>               write(outputfile,'BOT');
110>               writeln(outputfile);
111>               for j := 1 to no_elements do
112>               begin
113>                    outputvar:=inparray[i,j];
114>                    write(outputfile,'0,',outputvar:10:5);
115>                    writeln(outputfile);
116>                    write(outputfile,'V');
117>                    writeln(outputfile);
118>               end;                    {END, curve element dump FOR loop}
119>          end;                         {END, curve number FOR loop}
120>      write(outputfile,'-1,0');
```

```
121>      writeln(outputfile);
122>      write(outputfile,'BOT');
123>      writeln(outputfile);
124>      for j:= 1 to no_elements do
125>           begin
126>                write(outputfile,'0,',axes[j]/1e6:10:5);
127>                writeln(outputfile);
128>                write(outputfile,'V');
129>                writeln(outputfile);
130>           end;                        {END, axes dump FOR loop}
131>
132>      {WRITE END OF FILE MARKER}
133>      write(outputfile,'-1,0');
134>      writeln(outputfile);
135>      write(outputfile,'EOD');
136>      writeln(outputfile);
137>
138>      {CLOSE OUTPUTFILE}
139>      close(outputfile);
140>      chdir('\mg\noise'); {RETURN TO ORIGIONAL SUB-DIRECTORY}
141>
142> end;                                 {END, PROCEDURE FILEDMP}
143:
144: procedure varcalc;                        {EVALUATE PHASE VARIANCE
145:                                                FOR GIVEN DELAY TIME}
146:  begin
147:
148:       countf := 1;
149:       Kd :=Resp*sqrt(Pt*Plo*alpha*(1-alpha));
150:       K:=2*Ko*Kd*Kd*Kamp*Kamp*Km;
151:       Wn := Wnstart;
152:       while Wn<Wnmax do
153:          begin
154:
155:            {Initialisation}
156:                 T2 := 2*eta/Wn;
157:                 Tot1 := 0.0;
158:                 Tot2 := 0.0;
159:                 Integ1a:=0.0;
160:                 Integ1b:=0.0;
161:                 Integ2a:=0.0;
162:                 Integ2b:=0.0;
163:                 w := 1e6;
164:                 write('CYCLE ',countd:1,'-',countf:2,' of 4.'
165:                                                          ,char(13));
166:
167:                 if td <>0.0e-9 then      {TEST FOR NON
168:                                                NUMERICAL SOLUTION}
169:                 begin                    {INTEGRATION LOOP}
170:                     while w < 100*Wn do  {INTEGRATE UNTIL
171:                                              INTEGRAL LIMIT REACHED}
172:                     begin
173:                            denominator := DENOM(w);
174:                            tot1 := tot1 + SUM1(w);
175:                            tot2 := tot2 + SUM2(w);
176:                            w:=w+freqstep;
177:                     end;                {END, INTEGRATION LOOP}
178:                     sigmasqd[countf] := e*tot1/(4*pi*Resp*Pt*alpha)
179:                                       + (linewidth/(2*sqr(pi)))*tot2;
180:                 end;
```

```
181:
182:                If td = 0.0 then          (ANALYTICAL SOLUTION)
183:                   sigmasqd[countf] := e*Wn*3/(16*Resp*Pt*alpha*eta)
184:                                          + pi*linewidth/(2*eta*Wn);
185:
186:                storaged[countd,countf] :=sigmasqd[countf];
187:
188:                Wn := Wn + increment;
189:                countf :=countf+1;
190:           end;                                  (END, OMEGA N LOOP)
191:      end;                                       (END, PROCEDURE VARCALC)
192:
193:
194: begin
195:      clrscr;
196:      writeln('This machine is working. DO NOT SWITCH OFF.');
197:
198:      countd:=1;
199:      linewidth := 1.0e6;  (VALUE OF BEAT LINEWIDTH)
200:      td:= 0.0e-9;         (TIME DELAY PARAMETER)
201:      Wnmax := 510e6;      (MAX VALUE OF LOOP BANDWIDTH TESTED)
202:      varcalc;
203:
204:      countd := 2;
205:      td := 0.6e-9;
206:      Wnmax := 510e6;
207:      varcalc;
208:
209:      countd :=3;
210:      td :=1.8e-9;
211:      Wnmax :=300e6;
212:      varcalc;
213:
214:      countd := 4;
215:      Wnmax := 150e6;
216:      td := 3.5e-9;
217:      varcalc;
218:
219:      Wn:= Wnstart;
220:      for i:= 1 to no_elements do
221:           omega_n[i] := wn + (i-1)*increment;
222:      filedmp(storaged, omega_n);
223: end.
```

References.

1. T.G.Hodgkinson 'Phase-Locked Loop Analysis for Pilot Carrier Coherent Optical Receiver' Elect. Lett. vol 21 1985. pp1202-1203.

2. W.C.Lindsey 'Synchronisation Systems in Communication and Control' Englewood Cliffs, N.J.. Prentice-Hall 1972.

3. M.A.Grant, W.C.Michie, M.J.Fletcher 'The Performance of Optical Phase-Locked Loops in the Presence of Non-Negligible Loop Propagation Delay' Submitted for Publication to the IEEE J. of Lightwave Technol.

4. V.K.Prabhu 'PSK Performance with Imperfect Carrier Phase Recovery' IEEE Trans. on Aero. and Elect. Syst. vol AES 12(2) 1976. pp275-286.

CHAPTER 7.

COMPONENTS FOR A COSTAS LOOP SYSTEM.

7.1. Introduction.

The optical Costas loop (Figure 7.1) has now been extensively analysed to determine optimum loop parameters under certain specified operating conditions, namely the beat linewidth, system bit rate, and the required bit error rate for the link. Given this information it is now possible to investigate whether it is feasible to construct this loop with the facilities which are available to the project at the present time. Two lines of investigation are examined. Firstly, several components unique to the Costas loop will be either acquired or developed, then tested. Secondly, as a precursor to a Costas loop, attempts will be made to construct a heterodyne PLL using the semiconductor laser sources reported on in chapter 4. As was discussed in appendix 5B, this loop forms essentially half of the Costas loop. Analytically, it is very similar to the Costas loop, and in practice, it is considerably easier to implement. Therefore, success in obtaining phase-lock with the heterodyne format is vital before consideration can be given to moving on to the more complex system.

This chapter describes the steps taken along these two lines of investigation. Section 7.2 discusses the acquisition of the phase modulator required for the encoding of the transmit signal, E_T. Section 7.3 considers a design for a 90^o optical hybrid, which will combine the receive and local oscillator signals giving in phase and quadrature products at the two output ports. The design and construction of a detector/front end amplifier module is discussed in section 7.4, while section 7.5 describes the various beat spectra which were obtained using the semiconductor laser

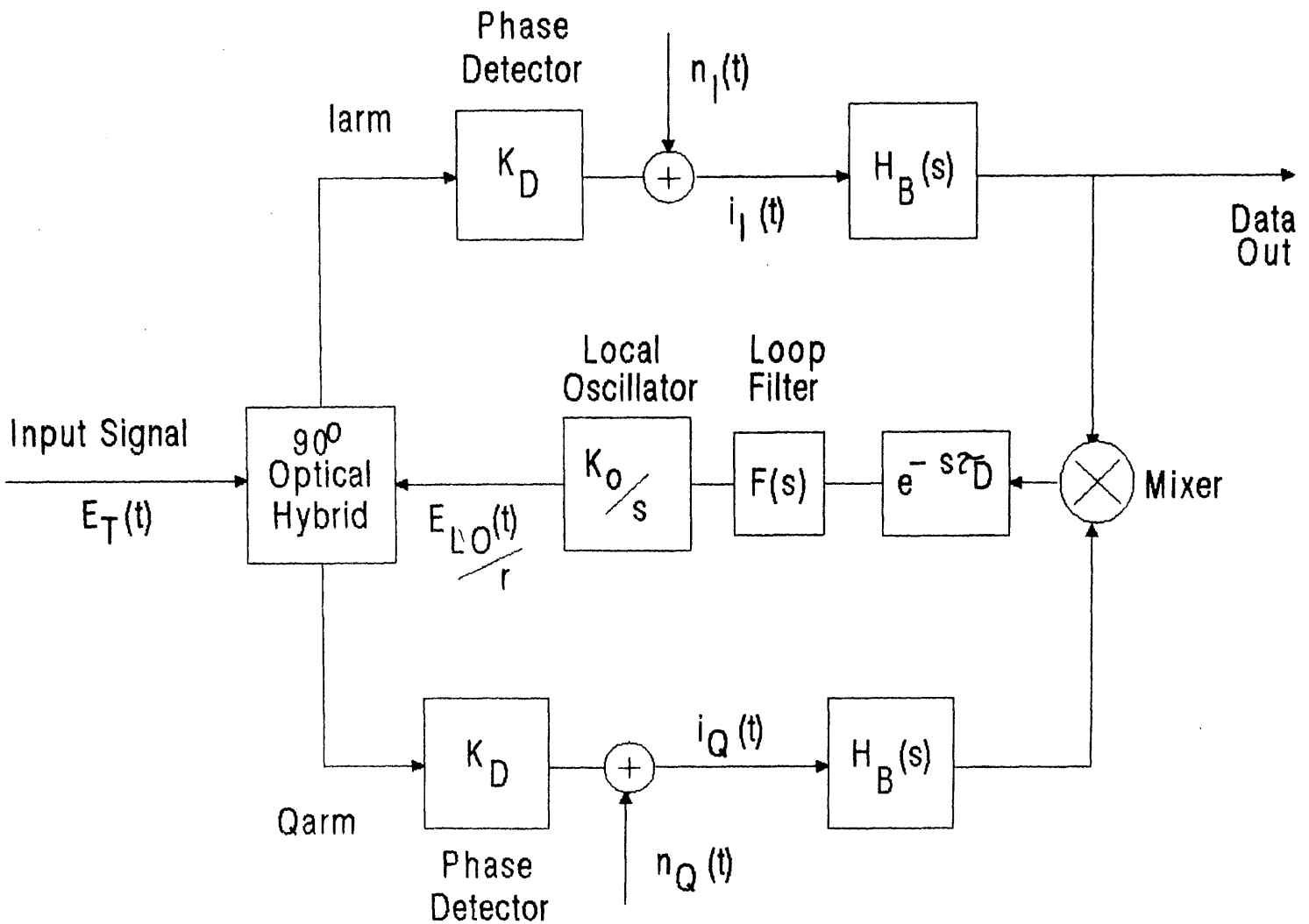


Figure 7.1. Schematic of the Optical Costas Loop with Time Delay.

r^2 = Power Reflection Coefficient of the Beam Combiner in the 90^o Hybrid

modules discussed in chapter 4. The results of experimental attempts to construct a heterodyne PLL are given in section 7.6. Unfortunately, this work was terminated prematurely by a fire in the Department. Consequently no optical PLL was in fact implemented. Section 7.7 contains discussion of the experimental results which were obtained and draws conclusions.

7.2. The Requirements and Acquisition of a Phase Modulator.

The phase modulator requirement for the Costas loop was for a simple single mode waveguide, lumped electrode structure that would give suppressed carrier modulation up to a modulating frequency of ~500MHz. Two lithium niobate ($LiNbO_3$) devices were constructed by Mr A. Faike. Each had five single guide/electrode structures of the type shown in figure 7.2. The Ti indiffused waveguides were constructed using the lift-off technique, as were the aluminium electrodes. The guides were 4μm wide and approximately 350 ±50 Ångstroms thick, while the electrodes were approximately 300nm thick. The load impedance for the signal generator was provided by a 50Ω through termination placed in line with the drive signal to the electrodes. This termination provided better than 1.2:1 VSWR up to 500MHz.

Testing of the two devices was carried out by coupling the output of the phase modulator into a scanning Fabry Perot spectrum analyser. Figures 7.3a and b show experimental results from the most efficient device. Suppressed carrier modulation is achieved when the phase shift applied by the device is $\pm\pi/2$. Figure 7.3b shows that good suppression was achieved. By varying the modulation frequency and adjusting the power output of the signal generator to maintain this condition, the frequency response of a device could be obtained. A plot of V_π (the required voltage for ±

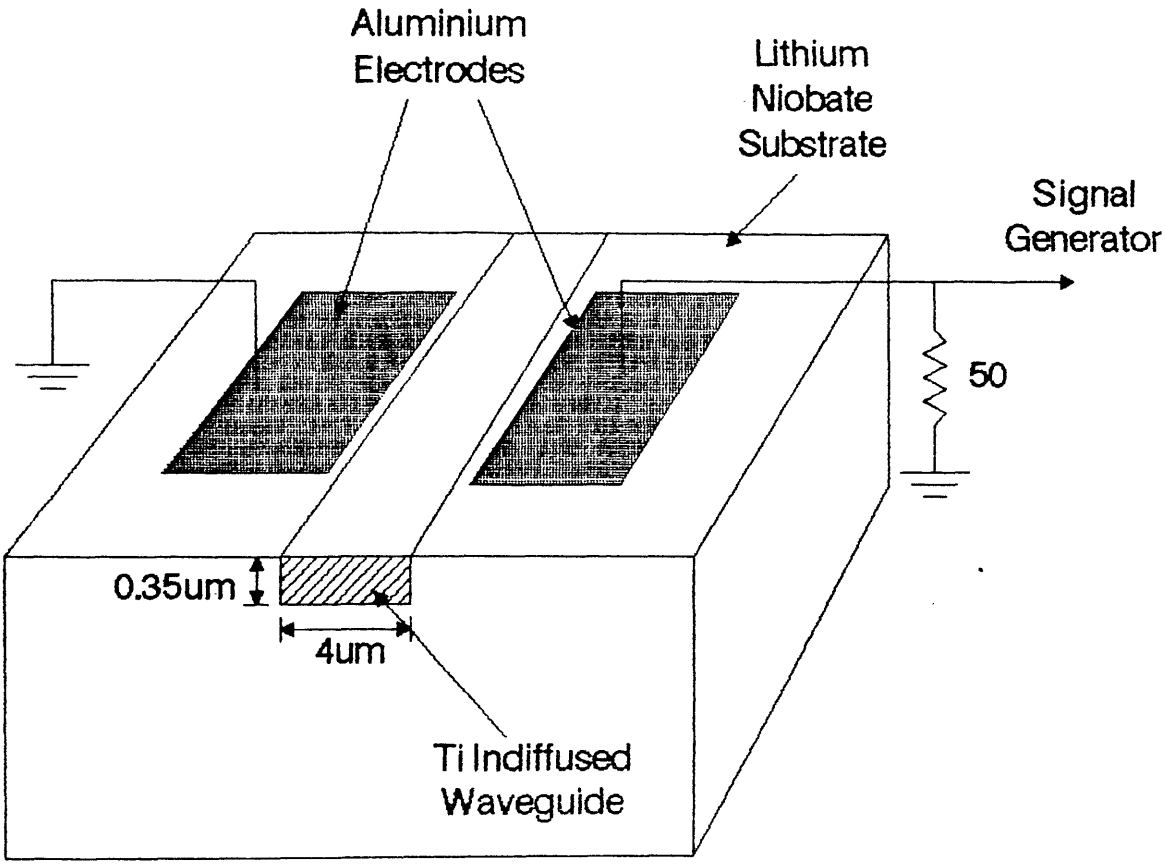


Figure 7.2. Phase Modulator Schematic Showing Lumped Electrode Structure.

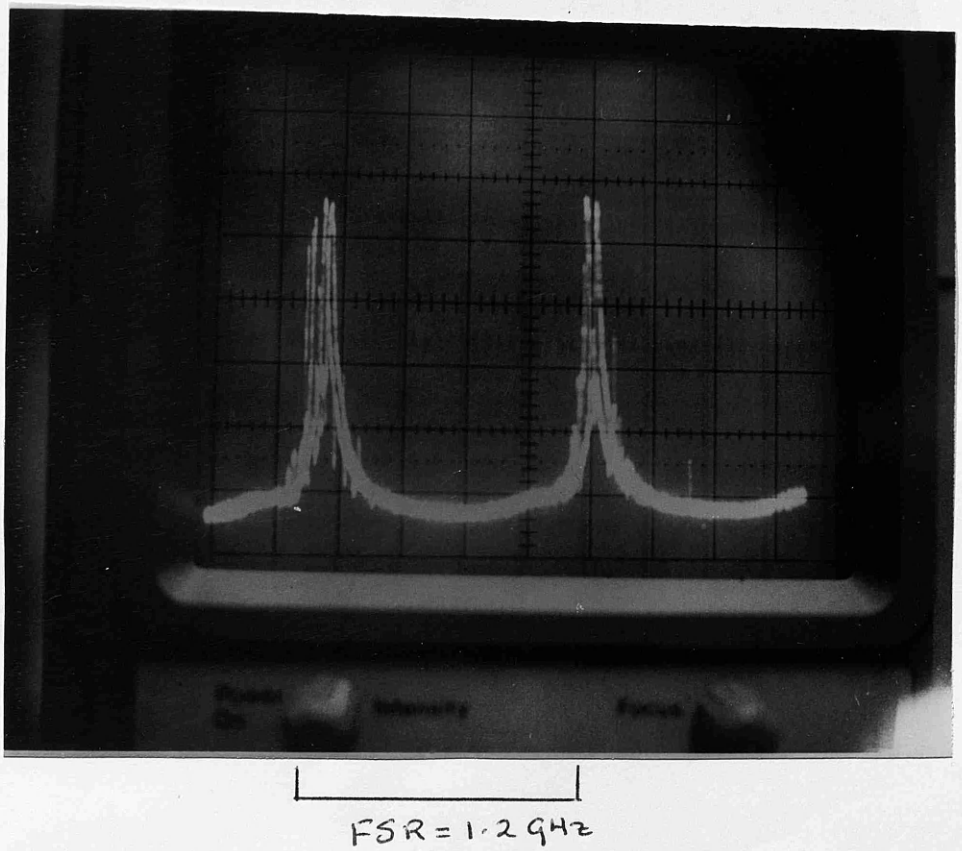


Figure 7.3a. Phase Modulator Output Without Modulation Drive Signal Applied.

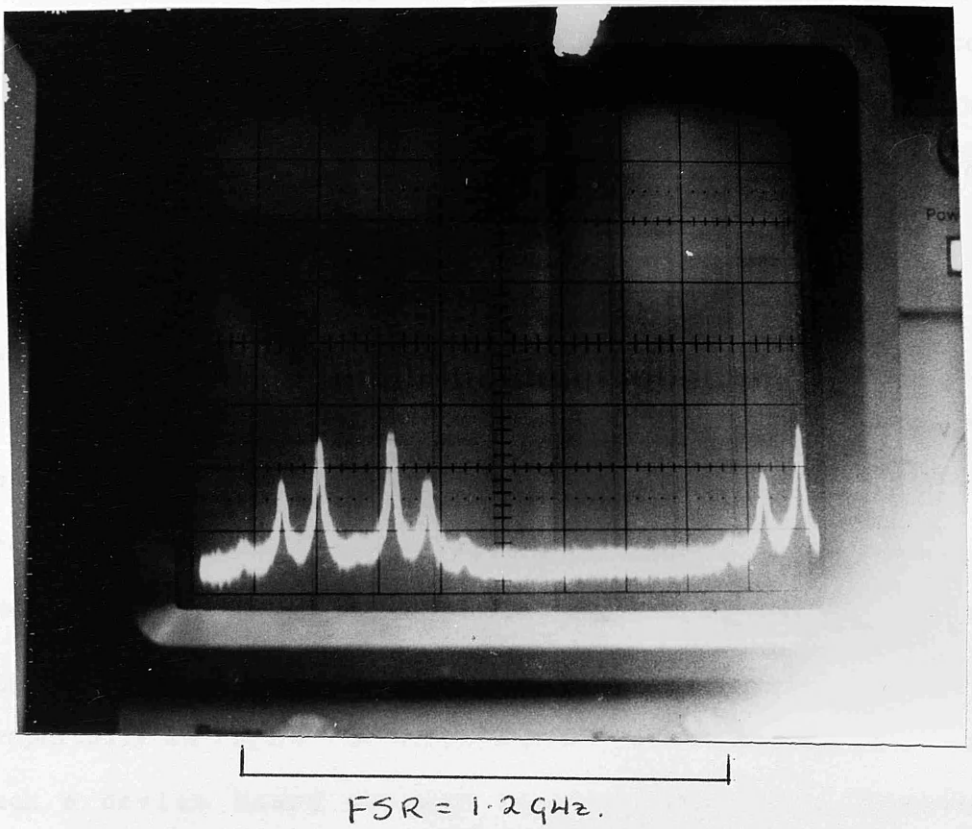


Figure 7.3b. Phase Modulator Output
With Drive Signal Applied
Showing Suppressed Carrier Modulation.

Modulation Frequency = 200MHz

Modulation Drive Voltage = 2.2V rms.

$\pi/2$ phase shift) versus modulation frequency, f_m, is shown in figure 7.4. It can be seen that the response drops off at 400MHz (note that this is an amplitude plot). However, operation at 500MHz is possible with only 8V rms across the device. The drive voltages given here are typical of Ti indiffused phase modulators in Y cut $LiNbO_3$. The second device failed to perform to this standard. It exhibited dual mode guiding as a result of variations in fabrication conditions.

7.3. 90^o Optical Hybrid.

A 90^o optical hybrid is a four port device which is required to provide at the output ports linear combinations of the signals present at the two input ports (E_T and E_{LO}). In addition, one of the output signals must exhibit a 90^o phase shift between E_T and E_{LO} relative to the other port. This requirement is presented graphically in figure 7.5. In principal, it is possible to produce such a device based on path length differences. However, tolerances of very much less than $\lambda/4$ would be required in the positioning of beam splitters and combiners in order for this technique to work. At an operational wavelength of 825nm this is clearly impractical. An alternative to this method has been proposed by Leeb[1]. His method requires manipulation of the states of polarization (SOP) of the two input beams, and then selection of the two output states using a polarizing beam splitter cube. A schematic diagram of this technique is shown in figure 7.6.

7.3.1. The Leeb Technique.

At the input to the hybrid, E_T and E_{LO} are linearly polarized in the x-plane (defined as the plane of the page). A $\lambda/4$ plate in the path of E_{LO} induces a $\pi/2$ phase shift between orthogonal

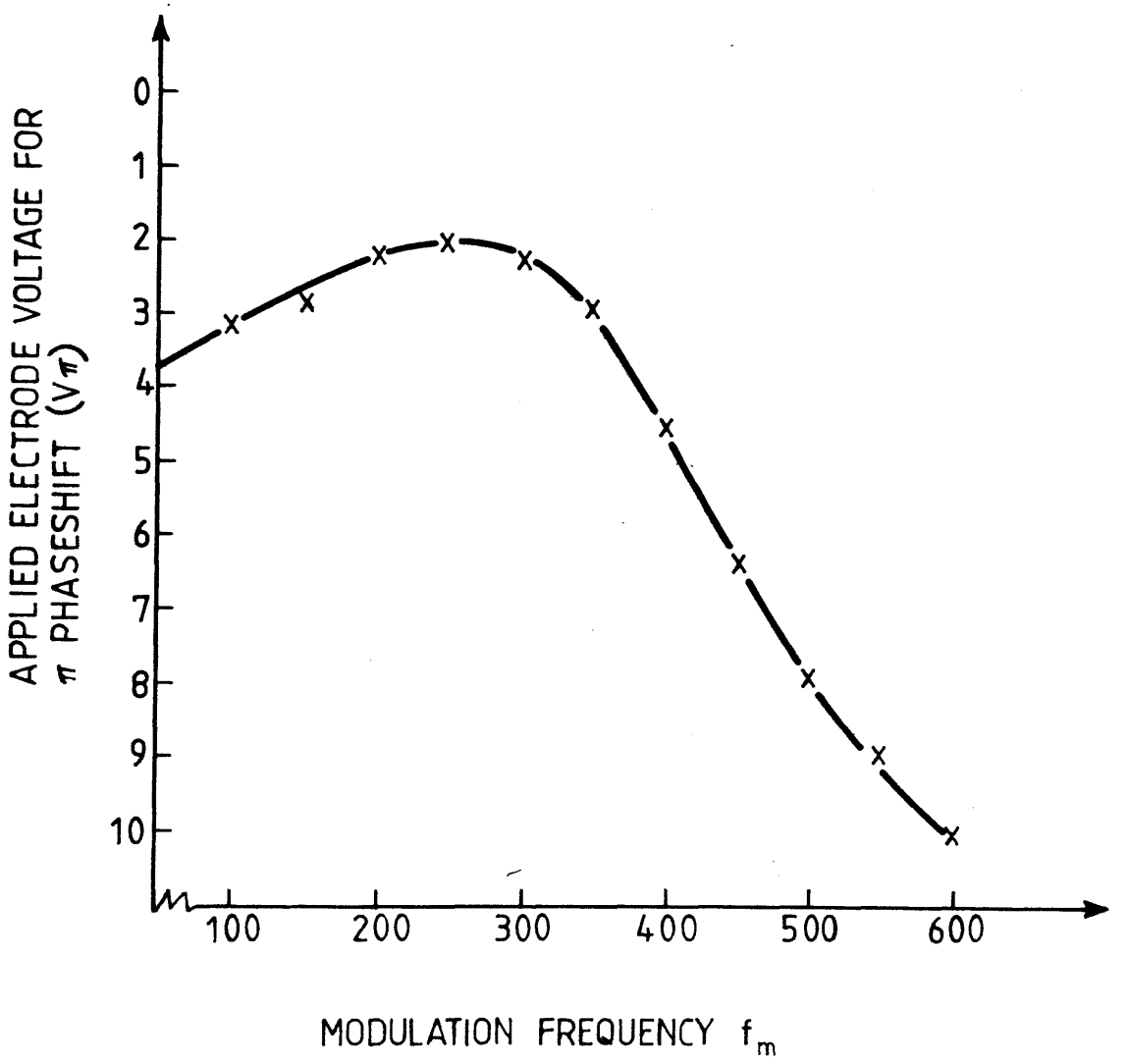


Figure 7.4. Frequency Response of Phase Modulator.

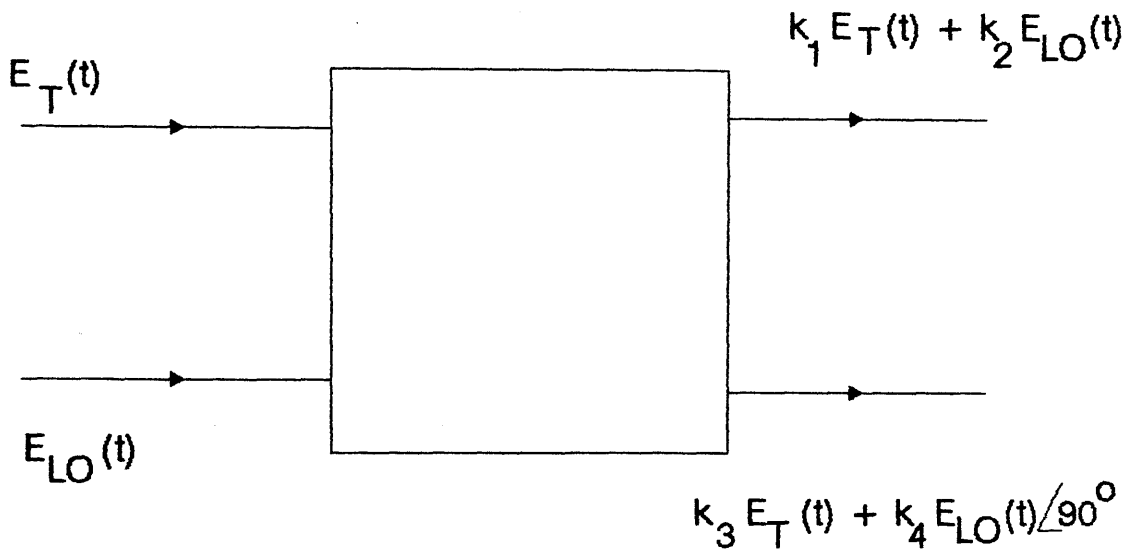


Figure 7.5. Definition of a 90^o Optical Hybrid.

K_i $\{ i = 1 - 4 \}$ are real constants.

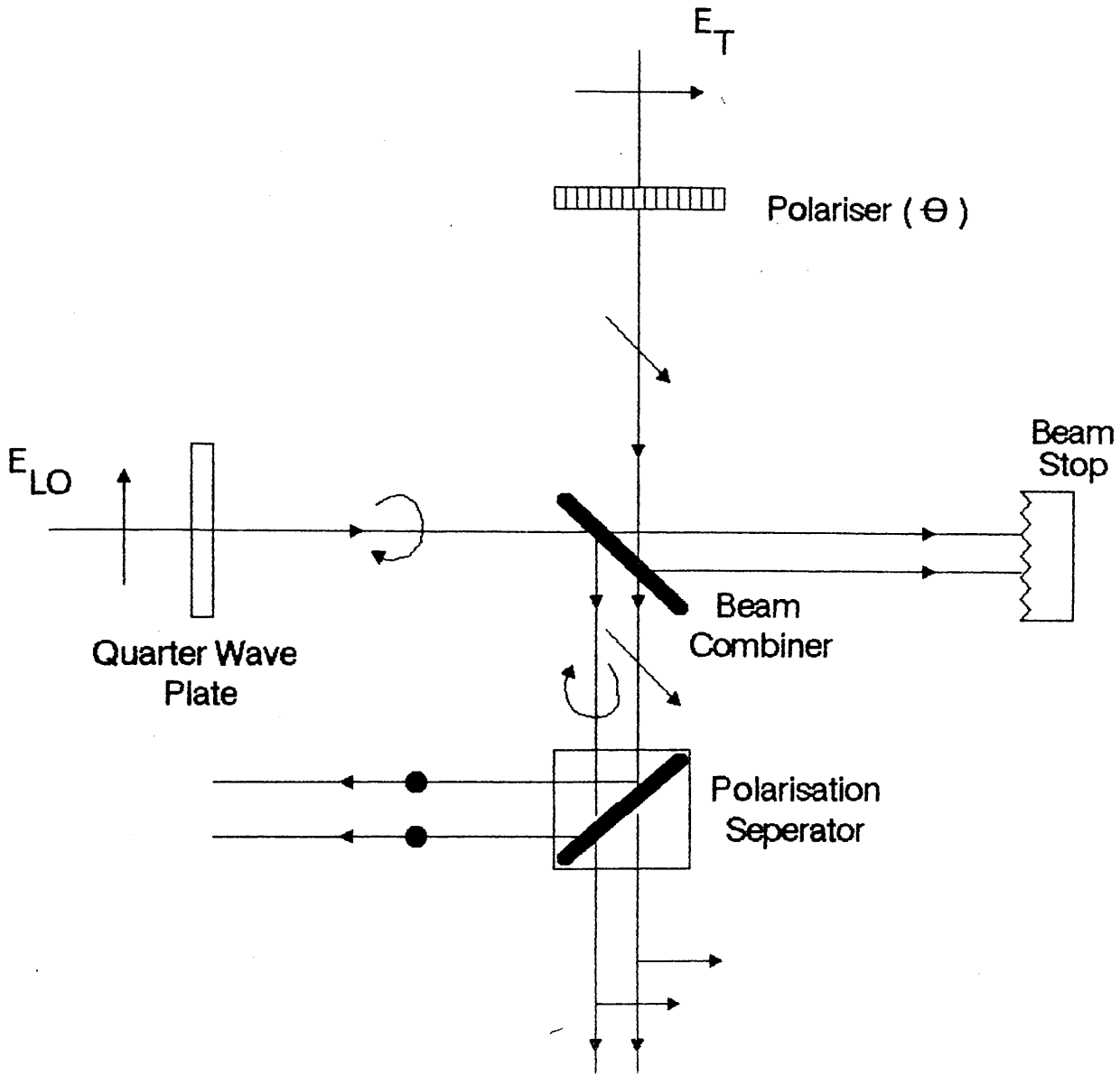


Figure 7.6. Schematic Drawing of a 90° Optical Hybrid [1].

→, ↘, ●, denote polarisation in, at a slant to, and perpendicular to the plane of the drawing.

↻ denotes circular polarisation.

polarization components, resulting in a circularly polarized beam. At the other port, E_T is passed through a polarization shifter which alters the amplitude of the x-y polarization components in proportion to the square root of α as defined in chapter 5. Assuming all other elements of the hybrid to be ideal then:-

$$\alpha = \cos^2\theta \qquad \text{----[7.1]}$$

where θ is the off x-axis angle of the polarization shifter output. These two signals are then spatially combined on the beam combiner. It is assumed that the reflection of the beam combiner, r^2, is non polarization sensitive ($r_x^2 \sim r_y^2$) and is very close to zero in order to minimise loss in the information bearing signal, E_T. If, for example $r^2 = 5\%$, then the power penalty introduced in E_T is only 0.2dB. However, this would also mean that a 10dB increase in local oscillator power would be required to restore E_{LO} to that level which would be obtained at the output of a 50/50 beam combiner. The combined signal is then passed through the polarization beam splitter cube resulting in two output signals with the desired properties.

7.3.2. Experimental Implementation.

Components for the implementation of this hybrid were acquired. At this stage, optimum BER performance was not being sought from the loop, and a large amount of signal power was available. Therefore, 50/50 and 70/30 beam combiners only were acquired from the CVI Laser Corporation, New Mexico. These elements were not ideal since they exhibited a ±20% variance in reflection coefficients in orthogonal axes. However, if the exact deviation from the nominal value is known, the polarization angle

in the transmit beam path, and the angle of the c-axis in the quarter wave plate in the LO path can be adjusted to compensate.

A polarization beam splitter cube (10BC16PC.7) was acquired from Newport Corporation, California. This had an extinction ratio of 13dB in the y-axis and 30dB in the x. Again, this is not ideal. However, by ensuring that the data port is the reflected output of the element, no significant penalty is incurred in data detection, and only 0.2dB is incurred in the phase tracking arm of the loop. Combined with the 0.2dB penalty incurred with a 5% reflectance beam combiner, it can be seen that it is practical for a hybrid of this type to exhibit losses within those allowed for in the analyses of chapters 5 and 6. Unfortunately, as a result of fire, it was not possible to confirm the operation of this hybrid experimentally. While Leeb successfully demonstrated this technique and its use in a Costas loop at $\lambda = 10\mu m$[1,2], recent work by Darcie and Glance[3] has, possibly more relevantly, demonstrated this element at 1.5μm in an image rejection mixer. This had a rejection efficiency compatible with the components described above. It should be noted that the components used in a near infra-red hybrid differ significantly from those used in a 10μm element. Hence the significance of Darcie and Glance's result.

7.4. The Detector/Front End Amplifier Module.

Detector/front end amplifier modules in a Costas loop are required to perform the following tasks: optical to electrical conversion, phase detection, and band pass filtering of the beat signal. Essentially, a photodiode, performing the optoelectronic conversion, low pass filtering, and phase detection of the optical signal, can be coupled to a wideband AC coupled amplifier to perform these functions. The characteristics of the amplifier will

be determined by the requirements of $H_B(s)$. In chapter 6 it was stated that the high frequency cut-off of the loop filter should be greater than $4\omega_n$. Hence, $H_B(s)$ should have a cut-off which at least falls outside this. Moreover, the photodiode should also be biased so that its response does not affect this loop filter pole. Taking the largest figure quoted in chapter 5 for a 15MHz beat linewidth ($\omega_{n.opt}$ = 2.22Grad/sec) then the high frequency pole of $H_B(s)$ requires to be greater than 8.88Grad/sec, or 1.41GHz. No commercial optoelectronic system which approached these specifications was available at the time. It was therefore necessary to design and construct a module 'in-house'.

7.4.1. Design and Construction.

It was decided to base the module around a hybrid wideband UHF amplifier from RS Components Ltd., device no. OM361. This gave a performance which was close to the required specification. The photodiode used was a BPW 26, an avalanche photodiode from AEG Telefunken biased to give a gain of 1. The biasing load resistance was chosen to obtain a cut-off of greater than 1.4GHz. In this mode, the thermal noise introduced by the resistor is high. However, this was not considered a problem as a) the local oscillator power available was large, and b) shot noise limited detection was not essential to the operation of the loop.

To satisfy the requirement for a 'data out' signal in the I arm, and a monitor signal output in the Q arm, a buffer amp with a BNC output was placed in circuit after the hybrid amplifier. This was a DC coupled buffer amp (device no. LH0033 CJ) again supplied by RS.

The main signal path was coupled to a wideband variable attenuator (Watkins and Johnson model WJ-G1). A 0-15v control

signal varied the through line attenuation between 25 and 2dB ± 0.5dB. When tuning the transmit and local oscillator lasers to provide a beat signal, the output powers of the two lasers are varied, thus altering $K.K_o$. The variable attenuator allows the experimenter to compensate for this variation thereby keeping the loop parameters constant.

Two of the above modules were combined to provide I and Q arm signals. The two outputs could then be fed into a double balanced mixer (Mini Circuits model ZFM-4) and the multiplier output filtered before being fed back to the local oscillator.

Tests on the mixer showed that the device required at least -20dBm of signal at one input port before it would function properly. At this level, the conversion gain, K_m, was -50dB. This was totally unacceptable for system operation. At -10dBm this figure was reduced to -20dB. However, an input signal of 0dBm realised the maximum conversion efficiency of -7dB. Whether an input signal level of 0dBm can be obtained in practice with the current module design will be discussed in section 7.5. Accommodating the introduction of a conversion gain term in the loop analysis requires simply the multiplication of the DC loop gain by K_m.

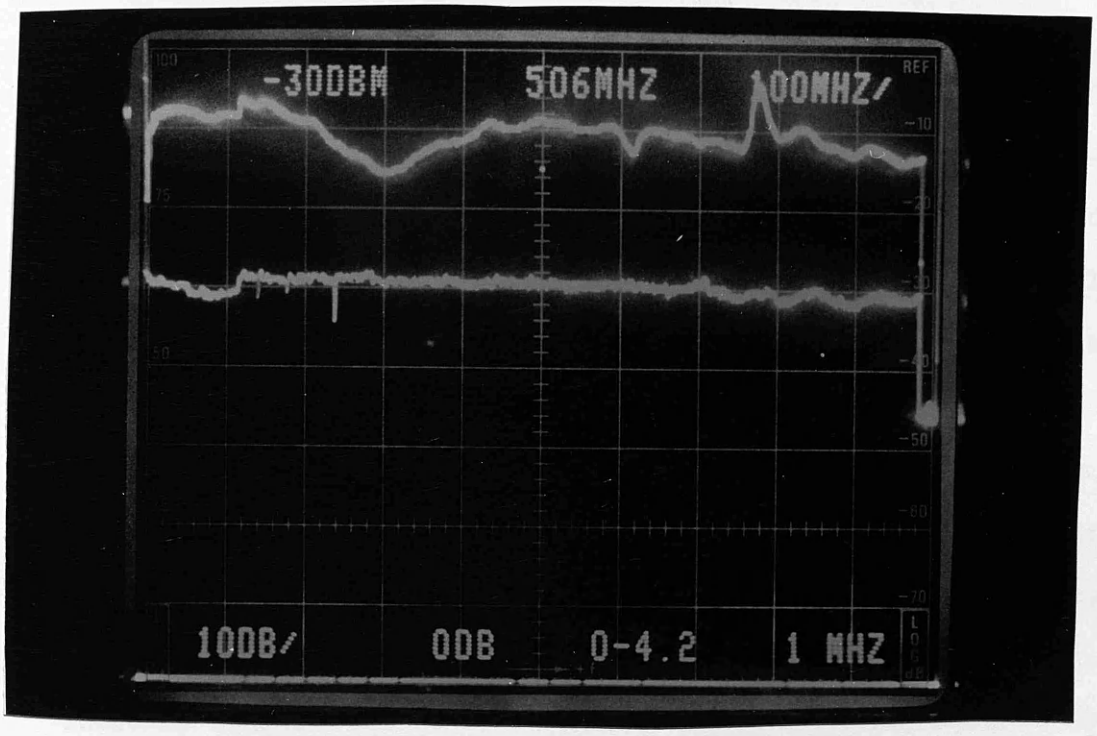
7.4.2. Module Frequency Response and Propagation Delay.

The frequency response of the amplifier/attenuator system was measured by applying the output of a signal generator to the photodiode socket. Figure 7.7 shows the combined input and output signals up to the limit of the signal generator's range, 990MHz. An average gain of 20dB (at the minimum attenuation level) is observed across the band with some fluctuations in the response noticeably around 300MHz. To check the modules optical response a

Reference Level (dBm)

Centre Frequency (MHz)

Frequency Step Per Division (MHz/div)



Vertical Display (dB/div)

Input Attenuation (dB)

Frequency Range (GHz)

Resolution Bandwidth

Figure 7.7. Frequency Response of Wideband Amplifier/ Attenuator Module.

Input Signal is Lower Trace,

Output Signal is Upper Trace.

beat signal was obtained between laser modules 5E3072 and 5E3060 and then tuned across the passband of the amplifier. The result of this experiment is shown in figure 7.8. This shows a cut-off at ~1150MHz, as well as good correlation with the RF tests. Although not ideal, the module comes close to meeting the suggested requirements of f_c = 1.4GHz.

The third important parameter to be measured was the propagation delay of the module. This was done by applying the output of a pulse generator to the system and comparing input and output pulses on a dual channel oscilloscope (Tektroniks model 2465). With the amplifier out of line, the delay introduced by the connecting lengths of coax was measured. This was found to be 0.5ns. Placing the amplifier in line, this figure rose to 2.6ns. Assuming that the response of the oscilloscope remains constant with varying amplitude/division settings, the propagation delay through the amplifier is 2.1ns. Assuming this to be the only delay inducing element in the PLL system, then, from the results of chapter 6, a maximum beat linewidth of ~5MHz would be required to obtain a system phase error variance of 0.15 rads^2. Whether a beat linewidth of this order is achievable in practice will be discussed in the next section.

7.5. Beat Spectra Experimentation.

Chapters 5 and 6 have derived the optimum value of loop bandwidth, ω_n, and maximum value of loop propagation delay, τ_D, for a given loop performance and beat linewidth. A wideband detector/amplifier module has now been obtained with a propagation delay time of 2.1ns. With the addition of a mirror cavity laser module, optical isolator, and beam combiner, the total loop delay time is likely to rise considerably. However, if loop performance

Reference Level (dBm)

Centre Frequency (MHz)

Frequency Step Per Division (MHz/div)

Vertical Display (dB/div)

Input Attenuation (dB)

Frequency Range (GHz)

Resolution Bandwidth

Figure 7.8. Frequency Response of Detector/ Wideband Amplifier/ Attenuator Module.

Response of Module to a Swept Optical Beat Signal Input.

is not a priority, i.e. the requirement is simply to phase-lock two lasers, then, with an active first order loop filter, τ_D can rise to 10ns before the beat linewidth needs to be reduced to 1MHz. Tests on the individual laser modules showed that this figure may be reached under optimum conditions.

The 'quality' of beat signal which can be obtained in practice will be discussed in this section. Results from beating two free running lasers, and one free running with one line narrowed laser will be examined and conclusions regarding locking two line narrowed lasers to obtain a 1MHz beat linewidth drawn. The power levels obtained from these spectra will also be discussed in the light of the mixer requirements outlined in section 7.4.1.

7.5.1. The Experimental System.

The experimental system used to obtain a beat spectra between two laser modules is shown in figure 7.9. All the beat spectra experiments utilised module 5E3072 as this was the most stable laser available (Figure 4.6). Of the other three new devices available in an external cavity configuration non had a modal response which fell naturally into this wavelength region. It was therefore necessary to use temperature control to 'pull' the modal response of the two lasers into an overlapping region. Unfortunately, the modal response was seen to fluctuate with temperature suggesting that this was not a linear operation. Indeed, the response was also seen to change as the nominal module temperature remained constant. This could be attributed to poor heat transfer away from the device causing localised heating, or to changes in local humidity. This postulation was reinforced by day to day results failing to be consistent when taken at

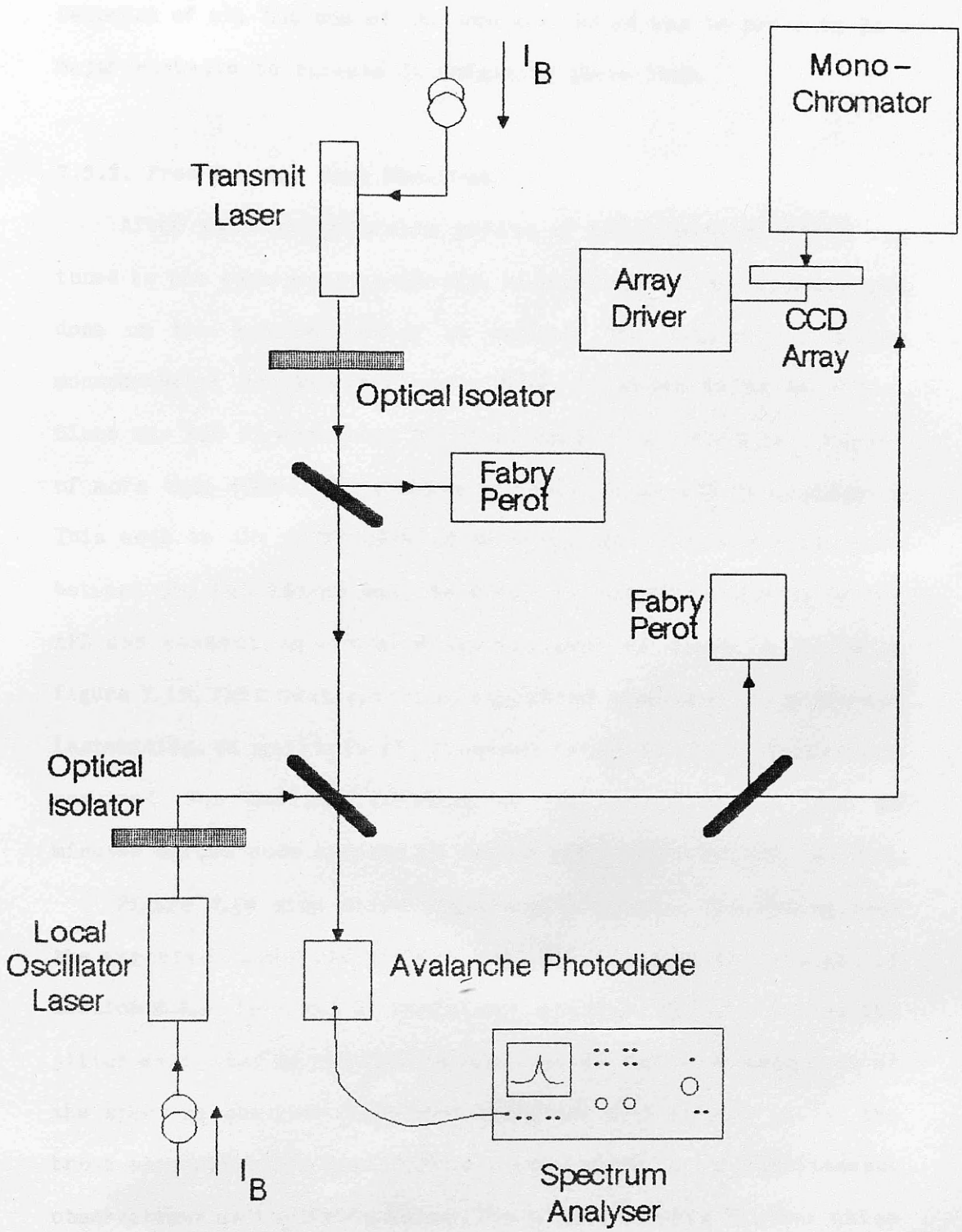


Figure 7.9. Experimental Arrangement for Obtaining Beat Spectra.

identical nominal temperatures. Inconsistency in the modal response of all but one of the devices tested was to prove to be a major obstacle to success in obtaining phase-lock.

7.5.2. Free Running Beat Spectrum.

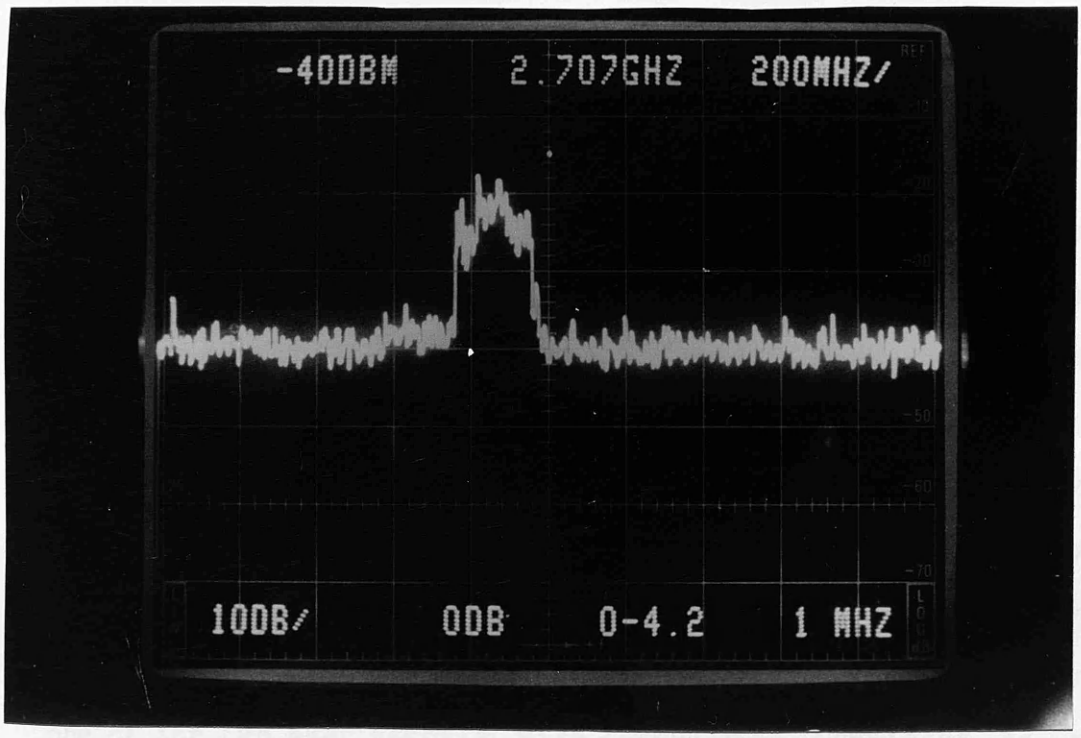
After some considerable period of time, module 5E3060 was tuned to the same mode as 5E3072. Measurement of wavelength was done on the monochromator as before. The resolution of the monochromator is, however, only 0.1Å, or about 44GHz at 825nm. Since the FSR of the Fabry Perot is only 2GHz, there is a region of more then 40GHz where the beat spectrum cannot be monitored. This adds to the difficulty of bringing the frequency difference between the two lasers down to 4GHz, the cut-off frequency of the APD and connecting cable. A typical beat spectrum is shown in figure 7.10. This beat spectrum exhibited considerable drift and instability. To maintain it, frequent retuning of the lasers was required. The maximum duration of this spectrum was some 30 minutes before mode hopping of device 5E3060 caused it to be lost.

Figure 7.10 also shows considerable spectral broadening over the expected linewidth of 20-30MHz, predicted by the results of section 4.3.3. This can be explained, however, by considering the jitter exhibited by the device (chapter 4). The slow scan rate of the spectrum analyser will then integrate this effect, giving the broad spectrum. This was confirmed experimentally by simultaneous observations on the Fabry Perot. The effect of this flicker noise has been considered by Kazovsky[4] and introduces some degradation into the system performance. However, for the system considered here, the degradation will be considered negligible relative to other noise sources.

Reference Level (dBm)

Centre Frequency (MHz)

Frequency Step Per Division (MHz/div)



Vertical Display (dB/div)

Input Attenuation (dB)

Frequency Range (GHz)

Resolution Bandwidth

Figure 7.10. Typical Beat Spectrum Between Devices 5E3072 and 5E3060.

Beat spectrum is several hundred megahertz wide due to frequency jitter exhibited by the two sources.

7.5.3. Line Narrowed Beat Spectrum.

Having obtained a free running beat spectrum attempts were made to line narrow one laser. A major problem with coupled cavity lasers, as has been stated, is that the laser will no longer exhibit continuous tuning in this mode of operation. Instead, it will resonate only on multiples of the cavity mode wavelength plus and minus some small percentage, depending on the cavity Q. Considering the difficulty with which free running beat spectra were obtained, it was undesirable to reduce the modal overlap area by more than was necessary. Thus device 5E3072 was left in its free running state and the line narrowing was performed on device 5E3060. Eventually a beat spectrum was obtained. However, as can be seen from figure 7.11, mode competition between external cavity modes produced two beat spectra spaced by the external cavity mode separation. This situation requires that the loop bandwidth of any PLL be less than this spacing in order that the loop locks to only one of the two beat signals. With larger beat linewidths, as is exhibited here, a shortening of the external cavity, increasing the cavity mode spacing, may be required to facilitate the acquisition of solid phase-lock. The duration of this spectrum was considerably shorter than the free running case. The maximum duration of the spectrum was 5-10 minutes. This was to be expected after the results of the self heterodyne measurement of the line narrowed laser, which showed the relative fragility of the coupled cavity optics.

7.5.4. Beat Spectra Power Levels.

From figure 7.10 it can be seen that the measured beat spectrum power level was -60dBm. It was established, however, that the spectrum analyser output was 8dB low, giving an actual beat

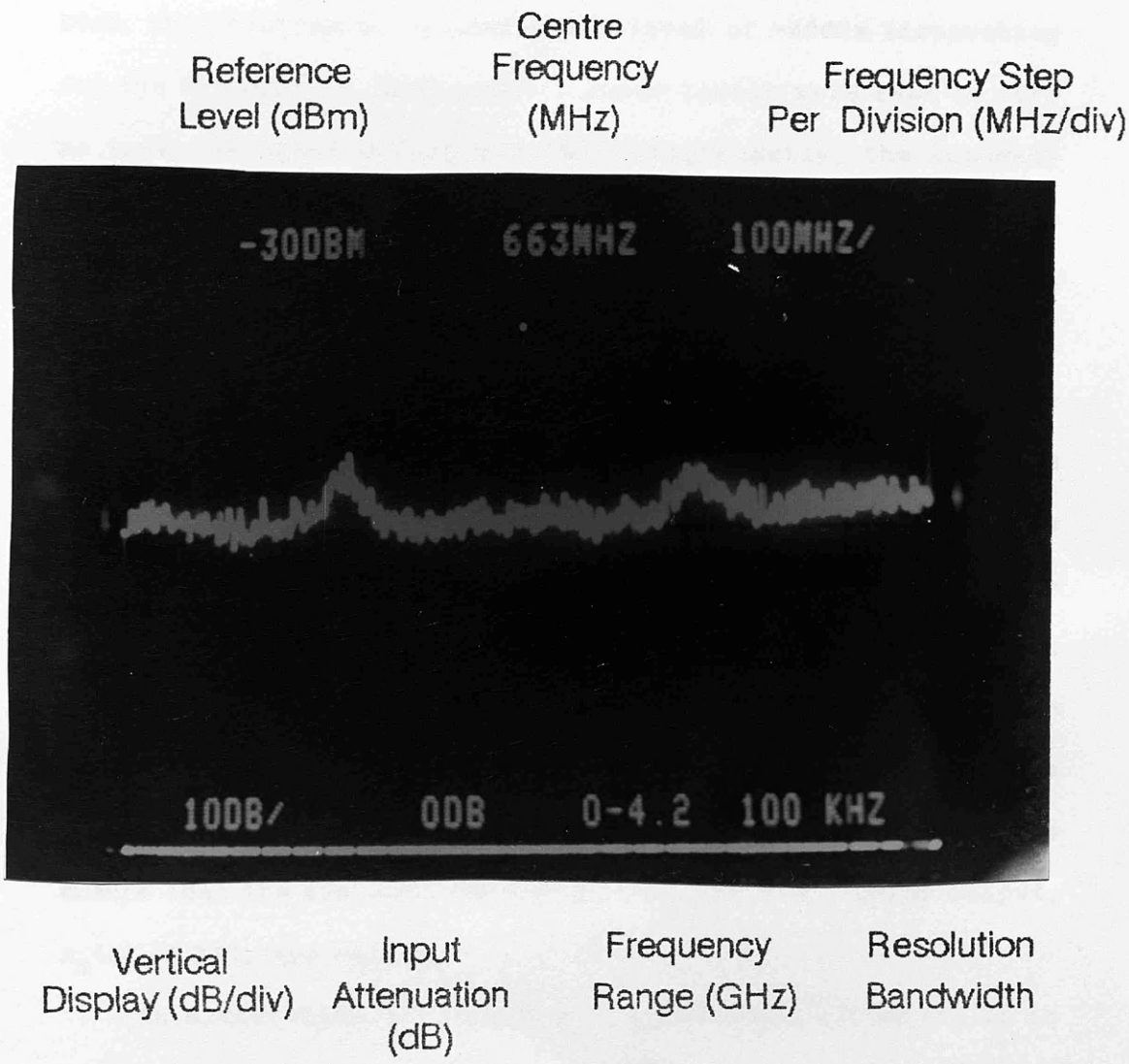


Figure 7.11. Line Narrowed Beat Spectrum Between Devices 5E3072 and 5E3060.

Only device 5E3060 is line narrowed.
Two beat spectra are observed due to coupled cavity mode competition.

signal power of -52dBm. This correlates well with the swept IF response of figure 7.8, where, with an amplifier gain of around 20dB, the spectrum has a peak power level of -40dBm (accounting for the 8dB offset). Beat spectra power levels were seen to vary as both the laser output and, more importantly, the accuracy spatial alignment and efficiency of the mixing process changed. In addition, when compared with the recorded photocurrents, the mixing efficiencies observed were poor, being estimated at ~5%.

Given the mixer requirements in section 7.4.1, there will be a need for considerable amplification in the I arm of any Costas loop which uses the above modules. It is feasible for the variable attenuator to be replaced by a cascadeable wideband amplifier (e.g. one of several devices available from Avantec Ltd.). Assuming that the delay times through the two devices are similar then the system should operate as before. In the loop analysis, a modification to K_{DI} will be required. This action would also ensure that the assumptions surrounding the mixer noise output, $n_m(t)$ [5.22], are valid.

An alternative to simply adding more amplification is to perform some decision process on the I arm signal and feed the recovered bit stream to the mixer at an appropriate level. This constitutes a decision driven PLL, as outlined by Kazovsky[4]. However, should the detection process introduce a further delay in the I arm, the Q arm signal requires to be delayed by an equal amount in order to obtain correct loop operation. This increase in loop propagation time will place further limitations on the magnitude of the allowable beat linewidth at the receiver input.

7.5.5. Discussion on Coupled Cavity Beat Spectra.

Considering the above results a number of points can be made

regarding the possibility of obtaining a homodyne beat spectrum with two line narrowed semiconductor lasers. Firstly, the modal overlap region will be further reduced from the two cases described above. This will increase the difficulty of obtaining a beat. Secondly, the stability of any beat spectrum that is obtained is proportional to half the individual module's stability. With the above devices, spectral stability, especially in the line narrowed mode of operation, is not good. A line narrowed spectrum, as has been stated, was obtainable from a single device for only a few minutes at a time. Attempts to tune this emission by altering the bias current of either laser inevitably resulted in this spectrum being lost. Therefore, the possibility of obtaining a usable line narrowed beat spectrum with two devices is remote. Attempts to do so have, so far, not succeeded. Thirdly, if a beat is obtained then trying to further tune the modules to produce a zero frequency IF will be exceptionally difficult. It can, therefore, be concluded that the two devices constructed and tested here are not suitable for use in a homodyne Costas loop.

7.6. Attempts to Implement a Heterodyne PLL.

Despite the poor results reported above it was decided to attempt to implement a heterodyne PLL. The experimental system is shown in figure 7.12. Module 5E3072 was chosen to be the local oscillator, with laser 5E3060 as the transmit source. After incorporating an optical isolator and beam combiner, the total optical path length of the loop was approximately two meters. The transit time for the local oscillator beam was therefore 6.6ns. Added to the measured front end amplifier module delay time, and including the mixer and loop filter plus connecting coaxial

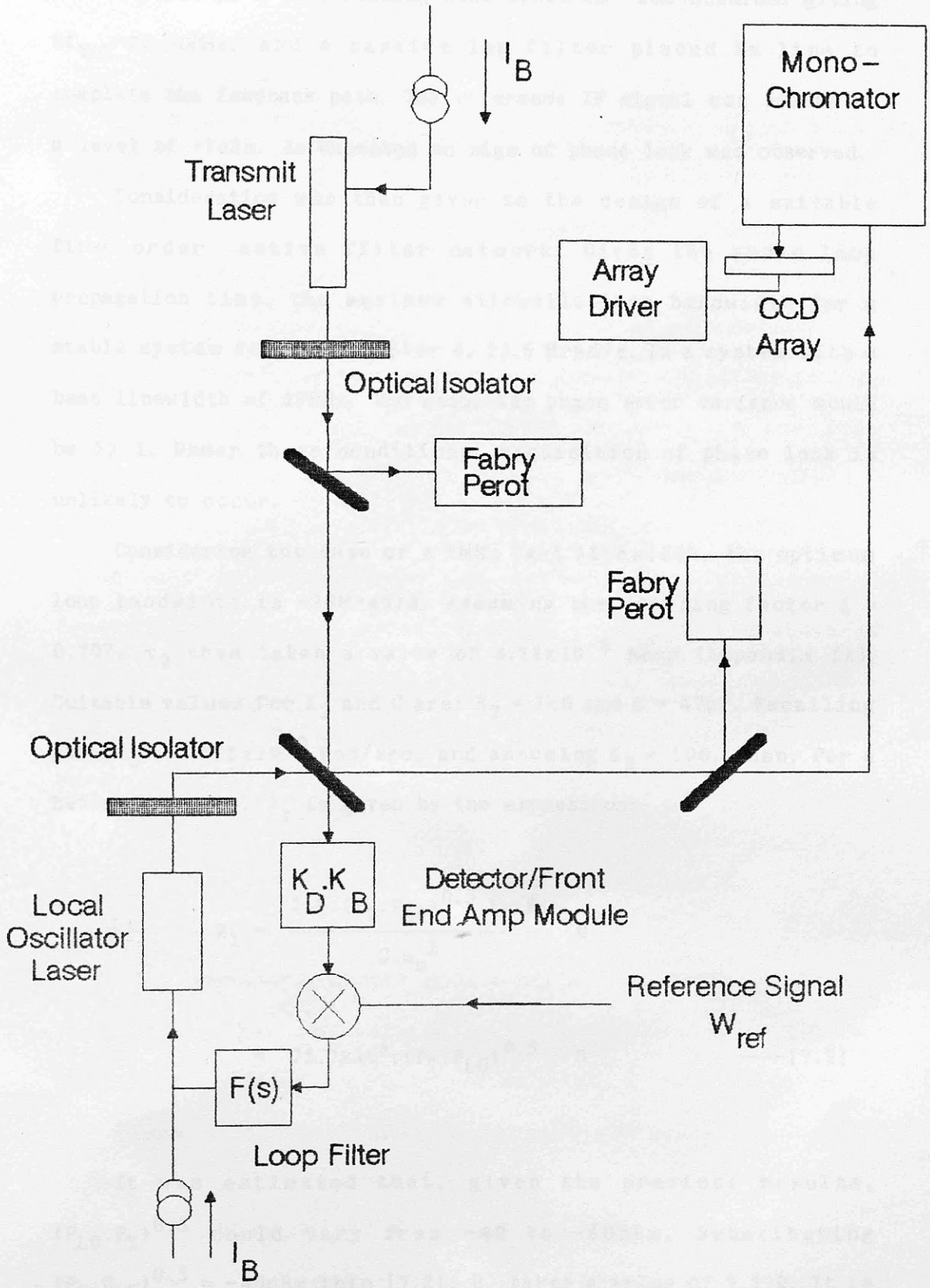


Figure 7.12. The Heterodyne PLL Experiment.

cables, the total loop delay time was approximately 10ns.

Initially, a free running beat spectrum was obtained giving δf_{TOT}~20-30MHz, and a passive lag filter placed in line to complete the feedback path. The reference IF signal was 500MHz at a level of +7dBm. As expected no sign of phase lock was observed.

Consideration was then given to the design of a suitable first order active filter network. Given the above loop propagation time, the maximum allowable loop bandwidth for a stable system is, from chapter 6, 73.6 Mrad/s. In a system with a beat linewidth of 15MHz, the resulting phase error variance would be >> 1. Under these conditions, acquisition of phase lock is unlikely to occur.

Considering the case of a 1MHz beat linewidth, the optimum loop bandwidth is ~30Mrad/s. Assuming the damping factor ξ = 0.707, τ_2 then takes a value of 4.71×10^{-8} secs (Appendix 5A). Suitable values for R_2 and C are: R_2 = 1kΩ and C = 47pF. Recalling that $K_o = 1.885 \times 10^{10}$ rad/sec, and assuming K_B = 100, then, for a heterodyne loop, R_1 is given by the expression:-

$$R_1 = \frac{2.R.(P_T.P_{LO})^{0.5}.K_B.K_o}{C.\omega_n^2} \quad \Omega$$

$$= 75.7 \times 10^6.(P_T.P_{LO})^{0.5} \quad \Omega \qquad \text{----[7.2]}$$

It was estimated that, given the previous results, $(P_{LO}.P_T)^{0.5}$ could vary from -40 to -60dBm. Substituting $(P_T.P_{LO})^{0.5}$ = -40dBm into [7.2], R_1 takes a value of 7.57Ω. It is desirable that the value of DC loop gain be sufficient to maintain a loop lock range consistant with the expected frequency deviations. The lock range of a second order loop is given by[5]:

$$\Delta f_L = K.K_o/2\pi \qquad \text{----[7.3]}$$

The values used in [7.2] will give a lock range of 51kHz. This is insufficient to maintain lock in a system subjected to frequency jitter of several tens megahertz. It is therefore necessary to increase the loop gain by >30dB.

This extra gain may be obtained through improvements in mixing efficiency, or through extra in-circuit amplification. The later, however, would result in an increase in the loop propagation delay figure, which may well be undesirable.

Unfortunately, as a result of fire, the construction of the filter was never completed and the loop never implemented. However, considering previous results, where the minimum beat linewidth obtained was 15MHz with substantial frequency jitter, success in achieving anything other than partial lock would appear to be unlikely.

7.7. The Implementation of a Costas Loop: Discussion and Conclusions.

This chapter has discussed both the acquisition of components required for the implementation of a Costas loop, and the experimental results from attempts to obtain a 1MHz line narrowed beat spectrum from the sources designed and developed in this project.

An optical phase modulator, capable of applying a $\pm\pi/2$ phase shift, was constructed. It was shown to function up to the required modulation frequency of 500MHz with an applied voltage (V_π) of 8V rms.

Following the work of Leeb[1], components for the

implementation of a 90^{o} optical hybrid were obtained. The theory behind the technique was discussed and the relationship between the coupling ratio, α, and the off X-axis angle of the input polarizer, θ, presented. it was also shown that the total power penalty relative to the input signal, E_T, could be kept below the figure specified in the analysis of chapters 5 and 6. This could be done by choosing the reflectance of the beam combiners to be less than 5%.

A wideband detector/amplifier module was designed and constructed. This had a variable gain of between -5 and 20dB, and provided a DC coupled buffered output. Using a swept beat spectrum input, the module was shown to have a cut-off frequency of 1150MHz. The propagation delay through the system was measured at 2.1ns. Beat spectra of less than 5MHz are therefore required at the input of the receiver in order to achieve a total phase error variance of $0.15 rads^2$ in a phase-locked system. Two of these modules were used in parallel to provide both I and Q arms for the Costas loop.

The properties of the double balanced mixer were also discussed. It was shown that a power level of 0dBm was required at one input port to ensure correct operation of the device. the minimum observed conversion gain, K_m, was -7dBm.

Attempts were then made to obtain the required line narrowed beat spectrum. A free running beat spectrum was obtained with some difficulty, but was maintained for up to 30 minutes with continuous monitoring and adjustment. Measurements of the beat spectra showed evidence of frequency jitter in the two lasers of ~200MHz.

In attempting to line narrow one laser, external cavity mode competition was observed. The line narrowed state was not a stable

one, with precise selection of the desired cavity mode being difficult. The beat spectrum in this case remained for a maximum of 5-10 minutes. Both of the above beat conditions exhibited inconsistency from day to day when nominal temperature and current conditions were repeated.

Attempts to obtain a beat between two line narrowed lasers did not succeed due to the instability of the two sources. Attempts to obtain phase lock with a heterodyne PLL incorporating an active filter were also terminated as a result of fire, However, it was expected that, given the propagation delay in the loop, the obtainable beat linewidth would be too large for the experiment to be successful.

Given that success in locking a heterodyne PLL is a prerequisite for a Costas loop, it is concluded that the above sources were unsuitable for use in the latter system. It is further concluded that coupled sources of this type are unlikely to be suitable for any homodyne system due to the discrete nature and limited tuning range of their emissions. Some form of continuously tunable source is recommended to maximise the possibility of success in this situation. Such a source could either encompass a Bragg grating external cavity, or some form of fine tuning of the cavity length of a coupled cavity system. This would be best effected by mounting the reflective element on a stable micro-positioning stack with piezoelectric (PZT) assisted actuators.

References.

1. W.R.Leeb 'Realization of 90^{o} and 180^{o} Hybrids for Optical Frequencies' Electron. Commun. AEU. vol 37 May-June 1983. pp203-206.

2. H.K.Phillip et al 'Costas Loop Experiments for a 10.6μm Communications Receiver' IEE Trans. Commun. vol COM-31(8) Aug. 1983. pp1000-1002.

3. J.E.Darcie, B.Glance 'Optical Heterodyne Image Rejection Mixer' Elect. Lett. vol 22(15) 1986. pp825-826.

4. L.G.Kazovsky 'Decision-Driven Phase-Locked Loop for Optical Homodyne Receivers: Performance Analysis and Laser Linewidth Requirements' J. Lightwave Technol. vol LT-3(6) dec. 1985. pp1238-1247.

5. F.M.Gardner 'Phaselock Techniques', Wiley, New York, 2nd Ed., 1979.

CHAPTER 8.

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH.

8.1. Conclusions.

The work presented in this thesis seeks to advance the knowledge base of coherent optical communication systems. In particular, it has been directed towards the analysis and implementation of a homodyne receiver for phase reversal keyed (PRK) optical transmissions. The most appropriate receiver format available for this task is the Costas loop. It is for this receiver that a set of design criteria is derived.

A theoretical set of maximum receiver sensitivities for a number of optical detection schemes is initially presented in chapter 2. The differences between the various available formats are outlined, and the performance improvements realisable from coherent detection highlighted. Expressions in terms of desired system performance, bit rate, and operational wavelength are evaluated to give a minimum theoretical received power for a given modulation format. For a 10^{-9} BER at 565MBit/s and transmitted at 830nm, a PRK homodyne detection scheme will require a minimum received power of -60.2dBm. A heterodyne PRK scheme will require an additional 3dB of power, while FSK requires a further 6dB. ASK homodyne and heterodyne detection is, on a power per bit basis, 9 and 12dB less sensitive than PRK homodyne detection.

It is also stated that, if there is a finite system phase error in the detection process, a degradation in performance will occur. A phase error variance of 0.03 rads^2 will incur a 0.5dB penalty in the received power required to attain a 10^{-9} BER. If this rises above 0.047 rads^2 then a 10^{-9} BER will not be achievable. This result is significant in terms of the analysis developed in chapters 5 and 6.

Relevant expressions describing the 3dB linewidth, or full width half maximum (FWHM), of a semiconductor laser are described in chapter 3. In the free running state, given the values of parameters reported by other workers, a FWHM of approximately 28MHz might be expected from a channel substrate planer (CSP) laser. Expressions describing linewidth reduction resulting from passive coupled cavity and external cavity techniques are also stated.

Measurements of the linewidths of free running semiconductor lasers are described in chapter 4. The experimental method used to obtain these results was the self heterodyne technique. Linewidths of 10MHz were measured in new devices, and 50MHz in older devices, suspected to be suffering from spectral ageing.

Two methods of linewidth reduction were investigated. These were the fibre coupled cavity module and the mirror coupled cavity module. A linewidth reduction of 1.82 (δf = 32MHz) was obtained from the fibre cavity module, while a reduction of 16 (δf = 750kHz) was achieved in the mirror cavity module. While a linewidth of this order was indeed sought for use in system experiments, a number of undesirable attributes were also observed in this module. These were: external cavity mode competition, severe mode instability, and inadequate tunability. Good performance in these areas is necessary for the successful implementation of a heterodyne PLL and is vital for a homodyne PLL. Reductions in cavity length will improve mode selectivity, and may improve tunability but only at the expense of increased linewidth. Improvements in mode stability require the very careful design and engineering of the laser mount to produce a stable, rigid structure. This should be considered the prime objective of any researcher considering constructing an external or coupled

cavity laser module for coherent communications purposes. In addition, vibration isolation of the optical table is also advised as an aid to mode stability. The consequences of failure to establish such a baseline for the semiconductor laser module are described in chapter 7.

An analysis of the performance of the Costas loop under the influence of non-negligible phase noise is presented in chapters 5 and 6. Expressions for the total system phase error variance are derived, and minimum performance penalties resulting from non-negligible beat linewidths calculated. It was found that minimum penalties of 2.5dB and 9.5dB would be incurred if beat linewidths of 1MHz and 15MHz respectively were encountered. The optimum values of loop bandwidth corresponding to these states were 146Mrad/sec and 2.22Grad/sec respectively. A loop model is then developed which accounts for the effects of non-negligible loop propagation delay. It was found that in order to maintain a 10^{-9} BER irrespective of power penalty, ($\sigma_E^2 = 0.047$ rads^2), the loop delay was required to be less than 1.8ns and 0.12ns respectively for each of the above cases. If that level of performance was not required, reliable acquisition of phase-lock could still be obtained with delays of 10ns and 0.67ns respectively.

It is probably not possible to obtain sub-nanosecond propagation delays in an optical phase-locked loop. The propagation of an error signal through the LO laser, the optical propagating medium (air or fibre), the detector, and the associated feedback electronics, mean that the minimum feasible time delay around the loop is probably a few nonaseconds. This being the case, there is a window of operation in an OPLL where time delay is a significant parameter but where phase-lock can be achieved. This window can be reasonably said to extend from beat

linewidths of approximately 100kHz to around 1MHz. Below 100kHz, time delay no longer becomes significant. Above a few megahertz, phase-lock is unlikely to be achieved. The identification of this window, and the description of loop performance within it, is a significant contribution, which this work makes to the understanding of optical phase-locked loops.

Attempts to realise reliable beat spectra of this order are described in chapter 7. Beat spectra of 20-30MHz were obtained from free running laser modules, although frequency jitter of several hundred megahertz, produced by the instabilities previously described, was also observed. Beat spectra of ~15MHz were obtained from one line narrowed source and one free running device. However, coupled cavity mode competition was observed and a double beat spectrum was recorded. The duration of this beat was never longer than 5-10 minutes. Attempts to obtain a beat between two-line narrowed modules did not succeed primarily as a result of the instability of the line narrowed state as discussed in chapter 4.

The construction of a suitable wideband detector/amplifier module is also described in chapter 7. The measured cut-off frequency of the module was 1150MHz, and the propagation delay time was 2.1ns. The design of an active filter network, considering beat spectrum, system lock range, and delay time requirements is also described. However, this work, plus an attempt to implement a heterodyne PLL was terminated by a fire in the Department.

8.2. Suggestions for Further Research.

The results of this work undoubtedly suggest that the major requirement for further research is the development of a stable but tunable semiconductor laser module with a linewidth of less than 500kHz. Short mirror coupled or external cavities are bulky but viable experimental solutions to the heterodyne phase-locked loop problem, providing they are carefully constructed in a single piece to a stable design. However, for a homodyne loop, the discrete, high Q, resonant modes of simple external cavities make it very unlikely that a zero beat frequency between two such modules will be obtained, no matter how well they are constructed. In addition, it should be noted that in both cases, the lack of tuning range associated with the high Q of the mode will severely limit the lock-in range of a loop. This is a highly undesirable effect when the magnitude of the frequency deviations likely to be encountered are considered. Thus, it can be seen that there are many problems associated with the use of external or coupled cavity semiconductor laser modules in coherent optical communications.

The replacement of the mirror with a Bragg grating in an external cavity format, has been shown to be the most promising experimental source yet developed for coherent communications. Properly constructed, it operates on a stable single mode with a narrow linewidth and exceptional tunability. However, the diodes used in this project would require extensive and complex modification before they would be suitable for use in this configuration.

In the longer term, considering the needs of a commercial system, the above solutions are likely to prove too bulky and unstable for serious consideration. A more plausible format may be

the combination of a fibre cavity with an etched fibre grating, incorporating compactness and stability with a frequency selective tuning element. This in itself would be a suitable subject for further study.

Ultimately, however, developments in DFB and DBR lasers, with their inherent single mode emission, mode stability, and tunability, should eventually provide a commercial coherent communications source. Indeed, it is envisaged that in the next few years, beat linewidths of the order of 1MHz will be realised with DFB lasers. This being the case, the work presented here will be particularly relevant for optical phase-locked loop design.

For the reception of PSK transmissions, the Costas loop is an optimum solution. Theoretically, its performance is equivalent to that of a pilot carrier system. However, in practice, the pilot carrier scheme will suffer additional degradation due to data-to-phase-lock crosstalk. For the Costas loop, there is a need to develop a fibre implementation of the 90^{o} optical hybrid, discussed in chapter 7. In this respect, a system based on the technique proposed by Leeb is a feasible proposition, and, indeed, is under investigation at this time by the same group in Vienna.

In conclusion, a considerable research effort is still required before coherent optical communications can be considered to be a mature and commercially viable technology. As with radio, that effort will undoubtedly continue up to and beyond the point where commercial viability is reached. It is hoped that the work presented here has aided this process, and contributed in some measure to the understanding of the field.