

Devices and Networks for Optical Switching

Michael George Taylor

A thesis submitted to the University of London
for the degree of Doctor of Philosophy (Ph.D.)

Department of Electronic and Electrical Engineering,
University College London.

October 1990

ProQuest Number: 10611083

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10611083

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

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

ABSTRACT

This thesis is concerned with some aspects of the application of optics to switching and computing. Two areas are dealt with: the design of switching networks which use optical interconnects, and the development and application of the t-SEED optical logic device.

The work on optical interconnects looks at the multistage interconnection network which has been proposed as a hybrid switch using both electronics and optics. It is shown that the architecture can be mapped from one dimensional to two dimensional format, so that the machine makes full use of the space available to the optics. Other mapping rules are described which allow the network to make optimum use of the optical interconnects, and the endpoint is a hybrid optical-electronic machine which should be able to outperform an all-electronic equivalent.

The development of the t-SEED optical logic device is described, which is the integration of a phototransistor with a multiple quantum well optical modulator. It is found to be important to have the modulator underneath rather than on top of the transistor to avoid unwanted thyristor action. In order for the transistor to have a high gain the collector must have a low doping level, the exit window in the substrate must be etched all the way to the emitter layer, and the etch must not damage the emitter-base junction. A real optical gain of 1.6 has been obtained, which is higher than has ever been reached before but is not as high as should be possible. Improvements to the device are suggested.

A new model of the Fabry-Perot cavity is introduced which helps considerably in the interpretation of experimental measurements made on the quantum well modulators. Also a method of improving the contrast of the multiple quantum well modulator by grading the well widths is proposed which may find application in long wavelength transmission modulators.

Some systems which make use of the t-SEED are considered. It is shown that the t-SEED device has the right characteristics for use as a neuron element in the optical implementation of a neural network. A new image processing network for clutter removal in binary images is introduced which uses the t-SEED, and a brief performance analysis suggests that the network may be superior to an all-electronic machine.

CONTENTS

ABSTRACT	2
CONTENTS	3
ACKNOWLEDGEMENTS	6
CHAPTER 1. Introduction	7
1.1 Structure of this thesis	11
1.2 References	12
CHAPTER 2. Switching Networks which employ Optical Interconnects	15
2.1.1 Optical interconnects	15
2.1.2 Multistage interconnection networks	16
2.2 Two dimensional networks	18
2.2.1 Perfect shuffle networks	18
2.2.2 Mapping rules for general interconnection patterns	23
2.2.3 Other 2D networks	24
2.2.3.1 Reverse perfect shuffles and granular perfect shuffles	24
2.2.3.2 Butterfly	25
2.2.3.3 Crossover	27
2.3 Controlling the electronic island size	27
2.4 Further comments on mapped 2D networks	28
2.5 Evaluation of a practical hybrid network	29
2.6 Conclusions	30
2.7 References	30
CHAPTER 3. Multiple Quantum Well Modulators	34
3.1 History of MQW modulator development	34
3.2 Operation of MQW Stark effect modulators	35
3.3 Fabry-Perot cavity effects	39
3.3.1 Five wave model of the Fabry-Perot cavity	43
3.3.2 Relationship between absorption and transmission	43
3.3.2.1 MQW absorption measurement in transverse modulators	45
3.3.2.2 Other applications of the five wave model	46
3.4 Graded well width MQW modulators	47

3.5 Conclusions	52
3.6 References	53
CHAPTER 4. The Integrated t-SEED	56
4.1 SEED development	56
4.2 Principle of operation of the t-SEED	57
4.3 Material growth and processing	61
4.4 Practical considerations in investigating devices	61
4.5.1 Device MV247	62
4.5.2 Device CB27	67
4.5.3 Device CB245	70
4.6 Dynamic response of the integrated t-SEED	74
4.7 Further improvements to the t-SEED	77
4.8 Conclusions	77
4.9 References	79
CHAPTER 5. Systems employing t-SEEDs	82
5.1 Requirements on t-SEED characteristic	82
5.2 t-SEED in a neural processor	84
5.3 Applications of t-SEED arrays as logic gates	88
5.3.1 An image processing network	91
5.4 Conclusions	96
5.5 References	96
CHAPTER 6. Conclusions	98
APPENDIX A. Proof of the standard Fabry-Perot equations starting from the five wave model	101
APPENDIX B. Epitaxial layer structures	105
B.1 MV247	105
B.2 CB27	106
B.3 CB245	107
Papers written in connection with this thesis	108
Patent applied for	108

ACKNOWLEDGEMENTS

I would like to thank John Midwinter my supervisor, Tony Overbury for his invaluable assistance with the laser system, Tony Rivers for the considerable work he has done in device processing, and Gareth Parry for his help and support throughout the project.

My own achievements during this Ph.D. would have been much less without the efforts of colleagues. In particular I need to thank Peter Stevens for teaching me most of what I know about quantum wells, Mark Abbott for hours of coaching in computer programming, and Craig Tombling for the work he has put into the experimental arrangements and elsewhere.

I have thoroughly enjoyed my three years at UCL, which is not in a small part due to the contributions of others. These people know who they are.

CHAPTER ONE

Introduction

All the work presented in this thesis can be said to fall into the category of optical computing, and as is noted later this also covers much of optical switching. Each chapter introduces its own topic, and this introduction will try to cover the ideas behind optical computing. How is optics made use of in a computer? The discussion of early work follows the excellent review by Goodman [1].

The first 'optical computers' utilised the Fourier transforming properties of a simple lens. The earliest machine was for processing synthetic aperture radar data. An aeroplane emits radar pulses as it flies, and records the resulting reflections from the ground on a length of photographic film via a cathode ray tube. Cutrona et al. [2] describe a coherent optical system which processes the coded information essentially by acting as a scaled down version of the radar system. As the film is wound through the input plane, a high resolution picture of the surveyed terrain is generated at the output. A later modification of this machine [3] can probably be considered to be the most successful optical computer yet, doing the equivalent of a huge number of computations.

The optical matched filter pattern recognition method of Vander Lugt [4] similarly uses the natural properties of optics for an otherwise difficult computation. By placing in the Fourier plane of a lens a mask whose transmission function is the complex conjugate of the Fourier transform of a search object, a spot, which is a correlation spike, appears in the image plane of the lens wherever the search object is found in the object plane. Obviously the method permits very rapid searching of a plane of data for some test pattern. The approach is reckoned by Goodman not to have been as useful as one might expect, perhaps because it is not very tolerant to distortions. If the search object is present but distorted by any two of translation, rotation and magnification, then the system cannot be made to pick it out.

Time varying signals can be processed using acousto-optic modulators. When an rf signal is input into the modulator it generates a travelling spatial variation in refractive index which can be explored with a coherent optical system. By using two acousto-optic modulators in a variety of configurations it is possible to correlate or convolve two signals [5-7]. Such systems can

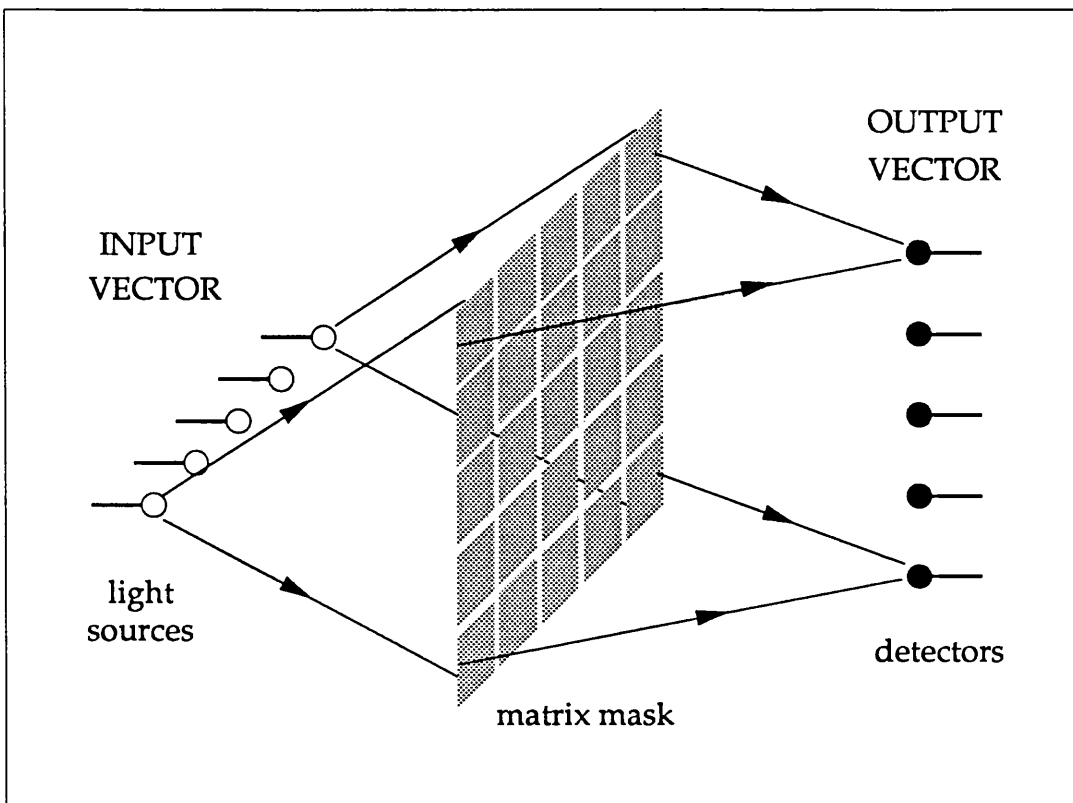


Figure 1.1 Stanford matrix-vector multiplier.

process in real time and can have very high time-bandwidth products, that is effective number of resolvable points. Whole systems can be integrated onto a chip using surface acoustic wave technology [8].

Clearly then optics is useful for computations which are naturally suited to optical imaging phenomena. Following this success, from the mid-1970s there were efforts to apply optics to more general computing problems where parallelism is involved. An example is the Stanford matrix-vector multiplier [9] shown in figure 1.1. Light is expanded vertically from a row of incoherent sources to fill a mask, and then condensed horizontally onto a column of detectors. The input vector is coded in the source array, and the matrix is the mask transmission. The problem here is in finding good applications for this fast but analogue system. There are algorithms which involve sending the output of the computation back into the input, but this becomes limited by noise after several iterations. Reference 10 describes similar mask-based systems for integrating over continuous spatial functions rather than discrete vectors, and reference 11 gives a proposal for a matrix-matrix multiplier which uses serially addressed acousto-optic modulators rather than 2D masks. It has also been noted that the matrix-vector multiplier is functionally the same as a crosspoint switch in telecommunications.

DMAC (digital multiplication by analogue convolution) is an attempt to

use optics for numeric processing. The central idea is that there is a similarity between multiplication and convolution operations, as is illustrated below, and convolution is one of the operations performed by the acousto-optic modulators.

MULTIPLICATION	CONVOLUTION
$\begin{array}{r} 1011 \\ \times 1110 \\ \hline 0000 \\ 1011 \\ 1011 \\ 1011 \\ \hline 10011010 \end{array}$	$\begin{array}{r} 1011 \\ \otimes 1110 \\ \hline 0000 \\ 1011 \\ 1011 \\ 1011 \\ \hline 1122210 \end{array}$

The overflow binary digits in the convolution are converted into true binary format by electronics afterwards, but hopefully the total computational task is reduced. DMAC has been proposed in a complicated matrix-vector multiplier arrangement [12] using serial input to arrays of two types of Bragg cell, fast and slow. The system requires fast A/D converters, and so the limit to the throughput comes from the speed of the electronics again.

Residue arithmetic is another example of numeric processing. A number can be expressed solely as the remainders when it is divided by some small bases; for example any integer between 1 and 100 is uniquely identified by its residues when divided by 3, 7 and 8. This number system has the interesting property that addition, subtraction and multiplication require no carry; the results of the individual digit operations are simply written down as residues. Residue arithmetic can therefore be done in parallel, and has been suggested for optical processors [13], and implemented using lookup tables involving an array of laser diodes [14]. To this writer it is questionable whether there is really any feature of optics which makes it suited to residue arithmetic. In any case the results have to be converted out of residue format, and so the most difficult part of the computation has been delegated to electronics.

A quite different type of system which is receiving a great deal of attention at the moment is the neural processor, which is based around the action of the brain [15]. The optical implementations of these processors are quite similar to the matrix-vector multiplier of figure 1.1. The output is fed back into the input via an array of neuron elements, which fix each pixel to a binary rather than a continuous value. A thick hologram can be used instead of the planar mask. The machine acts as an associative memory, and provides a way of exploiting the huge memory capacity of holographic

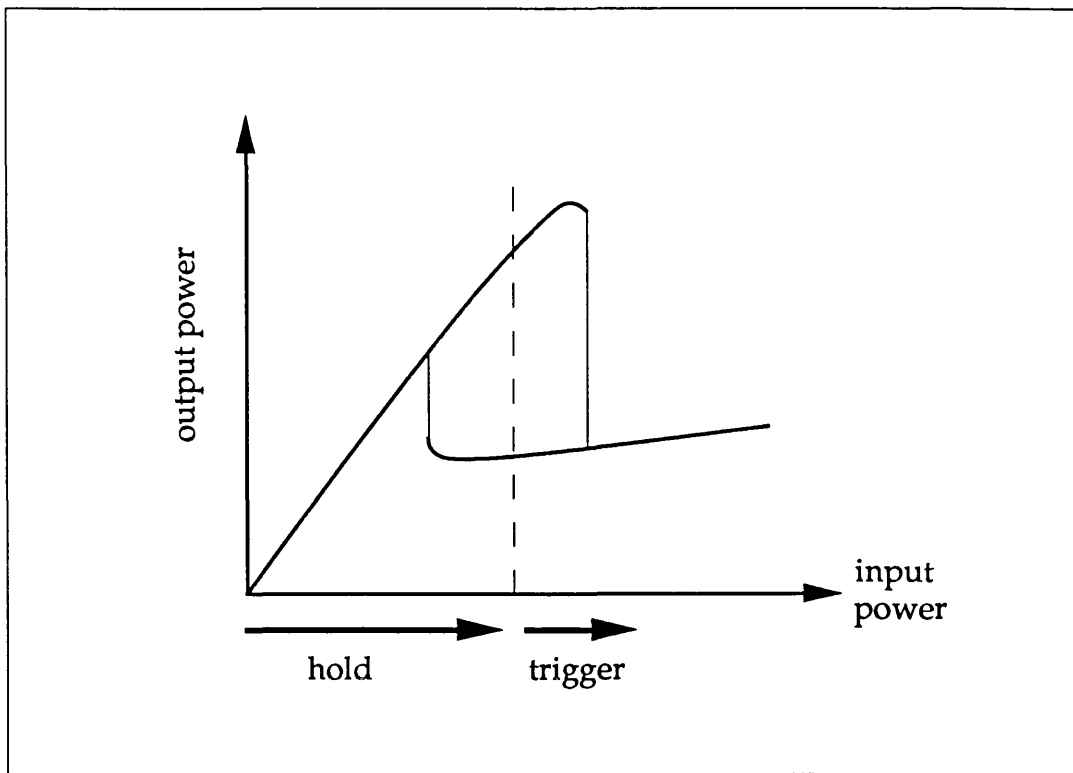


Figure 1.2 Input-output characteristic of a bistable device showing how the hysteresis loop can be used to do logic operations.

media. It is also tolerant to errors and nonuniformities to some extent, which is obviously desirable with optics. There is more on neural networks in section 5.2.

On the whole then optics has been applied to digital processors less successfully than to the specialist analogue ones. A new idea which has become widely accepted is the application of optics in a computer as an interconnect [16]. What limits the performance of conventional computers is not gate speeds but communication delays within the machine, and one would use optics not to do processing but only to transfer data. Optics is therefore used to augment electronic technology rather than to supplant it. Optical interconnects have been studied in free space and guided configurations for different levels of interconnect in a computer.

Recently the concept of an all optical digital computer became popular, following the discovery of bistable optical devices [17]. Here the logic gates which make up the processors have optical input and output, and the idea of an optical device which can remember its state seemed to be a great breakthrough. Smith et al. demonstrated logic operations with a switch comprising a Fabry-Perot etalon filled with nonlinear material [18]. Their approach is illustrated in figure 1.2. A constant power holding beam biases the device in the middle of the hysteresis loop in its input-output

characteristic, and a small triggering beam is added on top. If the trigger is large enough it will switch the device into the other state, allowing logic operations, and the output power of the device is larger than the triggering beam, so that the device has fanout. The culmination of this work was an all optical computer which clocked three binary digits round in a circle at about 10kHz, consuming 100mW of laser power [19]. The invention of the SEED [20], which employs a multiple quantum well optical modulator, made the all optical bistable devices look poor. Miller [21] showed that a SEED can emulate a nonlinear optical device, but with a lower switching energy and with much greater flexibility in trading off switching speed with power. Wheatley [22] said that a bistable characteristic was not necessary for optical logic, and proposed the t-SEED device as a logic element [23]. The project which this thesis describes follows on from Wheatley's work, and the development of the t-SEED is discussed in chapter 4. At the present time work is underway on telecommunications switching networks using SEED arrays [24], and array systems using liquid crystal light valves [25] and LED based VSTEP devices [26] have been built. Architectures are being studied which might be able to make efficient use of arrays of optical switching devices [27-29].

At the start of this project, then, the ideas of optical interconnects, optical neural networks and the optoelectronic SEED switches were of great interest. There were also a variety of schemes mixing these ideas with some of the older optical computing proposals - references 30 and 31 offer several examples. It is interesting to note that in the future there may be little distinction between optical computing and optics applied to telecommunications switching. The trend in telecommunications is for optical communication to be economic over shorter distances, and the computing optical interconnect can be considered as an extrapolation of that trend. The place where optical technologies might be applied in computing is in the interconnection fabric between parallel processing elements, as is mentioned in chapter 2, and there is little difference between this application and a switch in a high speed local area network. Conversely, as telephone traffic is packetised what one uses for routing is a computer as much as a telecommunications switch.

1.1 Structure of this thesis

The next chapter will discuss a class of multistage switching networks for telecommunications which use free space optical interconnects. The work on developing a monolithically integrated t-SEED involved considerable work

on the multiple quantum well modulator, which is a component part of the t-SEED. Chapter 3 is about the multiple quantum well modulator, and chapter 4 goes on to deal with the t-SEED. Chapter 5 looks at the role the t-SEED can play in an optical computing system, and tries to evaluate what performance advantage it might bring. Chapter 6 is a brief conclusion.

1.2 References

- [1] J. W. Goodman, "A short history of the field of optical computing," in *Optical Computing*, B. S. Wherrett & F. A. P. Tooley, eds., Proc. 34th SUSSP, 1988.
- [2] L. J. Cutrona, E. N. Leith, L. J. Porcello & W. E. Vivian, "On the application of coherent optical processing techniques to synthetic-aperture radar," *Proc. IEEE*, vol. 54, pp. 1026-1032, 1966.
- [3] A. Kozma, E. N. Leith & N. G. Massey, "Tilted-plane optical processor," *Appl. Opt.*, vol. 11, pp. 1766-1777, 1972.
- [4] A. Vander Lugt, "Signal detection by complex spatial filtering," *IEEE Trans.*, vol. IT-10, pp. 139-145, 1964.
- [5] L. Slobodin, "Optical correlation technique," *Proc. IEEE*, vol. 51, p. 1782, 1963.
- [6] R. A. Sprague & C. L. Koliopoulos, "Time integrating acoustooptic correlator," *Appl. Opt.*, vol. 15, pp. 89-92, 1976.
- [7] W. T. Rhodes, "Acousto-optic signal processing: convolution and correlation," *Proc. IEEE*, vol. 69, pp. 65-79, 1981.
- [8] A. P. Goutzoulis & I. J. Abramovitz, "Digital electronics meets its match," *IEEE Spectrum*, vol. 25, no. 8, pp. 21-25, August 1988.
- [9] J. W. Goodman, A. R. Dias & L. M. Woody, "Fully parallel, high speed incoherent optical method for performing discrete Fourier transforms," *Opt. Lett.*, vol. 2, pp. 1-3, 1978.
- [10] M. A. Monahan, K. Bromley & R. P. Bocker, "Incoherent optical correlators," *Proc. IEEE*, vol. 65, pp. 121-129, 1977.
- [11] R. A. Athale & W. C. Collins, "Optical matrix-matrix multiplier based on outer product decomposition," *Appl. Opt.*, vol. 21, pp. 2089-2090, 1982.
- [12] W. T. Rhodes & P. S. Guilfoyle, "Acoustooptic algebraic processing architectures," *Proc. IEEE*, vol. 72, pp. 820-830, 1984.
- [13] A. Huang, Y. Tsunoda, J. W. Goodman & S. Ishihara, "Optical computing using residue arithmetic," *Appl. Opt.*, vol. 18, pp. 149-162, 1979.
- [14] A. P. Goutzoulis, E. C. Malarkey, D. K. Davies, J. C. Bradley & P. R. Beaudet, "Optical processing with residue LED/LD lookup tables," *Appl. Opt.*, vol. 27, pp. 1674-1681, 1988.

- [15] Y. S. Abu-Mostafa & D. Psaltis, "Optical neural computers," *Scientific American*, pp. 66-73, March 1987.
- [16] J. W. Goodman, F. J. Leonberger, S.-Y. Kung & R. A. Athale, "Optical interconnections for VLSI systems," *Proc. IEEE*, vol. 72, pp. 850-866, 1984.
- [17] H. M. Gibbs, S. L. McCall, T. N. C. Venkatesan, A. C. Gossard, A. Passner & W. Wiegmann, "Optical bistability in semiconductors," *Appl. Phys. Lett.*, vol. 35, pp. 451-453, 1979.
- [18] S. D. Smith, I. Janossy, H. A. MacKenzie, J. G. H. Matthew, J. J. E. Reid, M. R. Taghizadeh, F. A. P. Tooley & A. C. Walker, "Nonlinear optical circuit elements as logic gates for optical computers: the first digital optical circuits," *Opt. Eng.*, vol. 24, pp. 569-574, 1985.
- [19] S. D. Smith, A. C. Walker, F. A. P. Tooley & B. S. Wherrett, "The demonstration of restoring digital optical logic," *Nature*, vol. 325, pp. 27-31, 1985.
- [20] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood & C. A. Burrus, "Novel hybrid optically bistable switch: The quantum well self-electrooptic effect device," *Appl. Phys. Lett.*, vol. 45, pp. 13-15, 1984.
- [21] D. A. B. Miller, "Quantum-well self-electro-optic effect devices," *Opt. Quantum Electron.*, vol. 22, S61-S98, 1990.
- [22] P. Wheatley, "Devices for optical logic," Ph.D. Thesis, University of London, 1987.
- [23] P. Wheatley, P. J. Bradley, M. Whitehead, G. Parry, J. E. Midwinter, P. Mistry, M. A. Pate & J. S. Roberts, "Novel nonresonant optoelectronic logic device," *Electron. Lett.*, vol. 23, pp. 92-93, 1987.
- [24] T. J. Cloonan, M. J. Herron, F. A. P. Tooley, G. W. Richards, F. B. McCormick, E. Kerbis, J. L. Brubaker & A. L. Lentine, "An all-optical implementation of a 3-D crossover switching network," *IEEE Photonics Tech. Lett.*, vol. 2, pp. 438-440, 1990.
- [25] B. Clymer & S. A. Collins, "Optical computer switching network," *Opt. Eng.*, vol. 24, pp. 74-81, 1985.
- [26] I. Ogura, Y. Tashiro, S. Kawai, K. Yamada & M. Sugimoto, "Reconfigurable optical interconnection using a two-dimensional vertical to surface transmission electrophotonic device array," *Appl. Phys. Lett.*, vol. 57, pp. 540-542.
- [27] K.-H. Brenner, A. Huang & N. Streibl, "Digital optical computing with symbolic substitution," *Appl. Opt.*, vol. 25, pp. 3054-3060, 1986.
- [28] T. J. Cloonan, "Performance analysis of symbolic substitution," *Appl. Opt.*, vol. 27, pp. 1701-1707, 1988.
- [29] J. Tanida, M. Fukui & Y. Ichioka, "Programming of optical array logic. 2: Numerical data processing based on pattern logic," *Appl. Opt.*, vol. 27, pp. 2931-2939, 1988.

- [30] Optical Computing, 1989 Technical Digest Series, vol. 9, (Optical Society of America, Washington D.C. 1989).
- [31] Optical Computing, J. W. Goodman, P. Chavel & G. Roblin, Eds., SPIE Proc., vol. 963, 1988.

Switching Networks which employ Optical Interconnects

The networks described in this chapter are intended to be examples of very good optical computing machines, which use optics only as an interconnect. By sticking closely to the basic ideas of when an optical interconnect is appropriate, it is the aim of this chapter to arrive at a system which can perform better than an all-electronic equivalent, so providing support to the arguments in favour of using optics in computing.

2.1.1 Optical interconnects

The seminal work in this area was by Goodman et al. [1] who noted that what limits electronics is communication within the system rather than speed of processing, and by using optics for transmission of data one can remove the bottleneck. Although the technologies available were not as advanced as they are now (multiple quantum well modulators, which are the subject of the remaining chapters of this thesis, were undeveloped), the paper put forward several suggestions as to how to incorporate optical data links into a computer. The simplest application is in clock distribution, where timing information arrives in synchrony at many sites in a computer system via light links. Without thinking about the interface with electronics, one notes that a free space interconnect made with glass elements or with holograms can link an array of channels in parallel, with no crosstalk and with negligible time skew between channels, and the transmission medium has an effectively infinite bandwidth. Optics has great potential for array processing.

By comparing optical and electrical interconnects one can find the regime where the former alternative is better. An optical interconnect comprises an emitter, which could be a semiconductor laser or a multiple quantum well modulator, and a detector, both of which are integrated with the electronics. An electrical interconnect is basically a metal track. Often the limit to how much circuitry will fit on a chip comes from the number of pin outs around the periphery, and so gaining access to the third dimension using optics is a large advantage. Optical interconnects have been modelled [2-4] to enable a quantitative comparison with electrical interconnects, though it must be noted that the electrical link is treated as a lumped capacitance and not as a

transmission line. One finds that on a power consumption versus speed basis optics wins over the slower inter-chip electrical links, but it is more marginal for communication within a chip. There is a break even length over which the photonic link is the superior one; its exact value depends on the analysis, a distance of 10mm at data rate of 1GHz being an example. The fact that there is a break even line length suggests that an optimal system might have electronic islands, within which all communication is electrical, which are linked to one another by optical interconnects [5].

It is only really at high speeds, then, that the application of optics becomes appropriate, and routine substitution of electrical connections with photonic ones will not yield a better machine. To evaluate the advantages of optical interconnects a whole system must be considered, and that is what the remainder of this chapter will be dealing with.

2.1.2 Multistage interconnection networks

The best example of a machine which benefits from the inclusion of optical interconnects is the multistage interconnection network (MIN). The task performed by a MIN is to interconnect two sets of terminals, like a crossbar switch. Figure 2.1 is an example of a MIN. There are many different kinds which have a common structure, but the details of operation need not be of concern here. References 6 to 10 explain how the networks operate, and references 11 to 18 give more recent work on some MIN variations and practical applications. Obviously these networks are of interest to the telecommunications community, but they also have applications in supercomputers [19] where one requires fast communication between parallel processing units and parallel memories.

In any MIN the communication lines are processed by a pipeline of stages, each stage comprising an interconnection pattern (the network of figure 2.1 uses a perfect shuffle pattern) where the lines are permuted in some way, followed by a layer of switching units which operate on adjacent pairs of data lines. In non-broadcasting networks these switches either exchange or bypass their inputs. For an N channel network it is possible to achieve full connectivity (an input can talk to any output) in $\log N$ stages, full rearrangeability (any combination of inputs talking to outputs) in $2\log N - 1$ stages, and a self-routing rearrangeable network in $\frac{1}{2}\log N(\log N + 1)$ stages (all logarithms to base 2). There are also parallel processing networks, such as fast Fourier transforming [7,20] and sorting networks [7,21], which have the same structure as the interconnection networks, but whose modules execute some processing task on their inputs. The results of this chapter

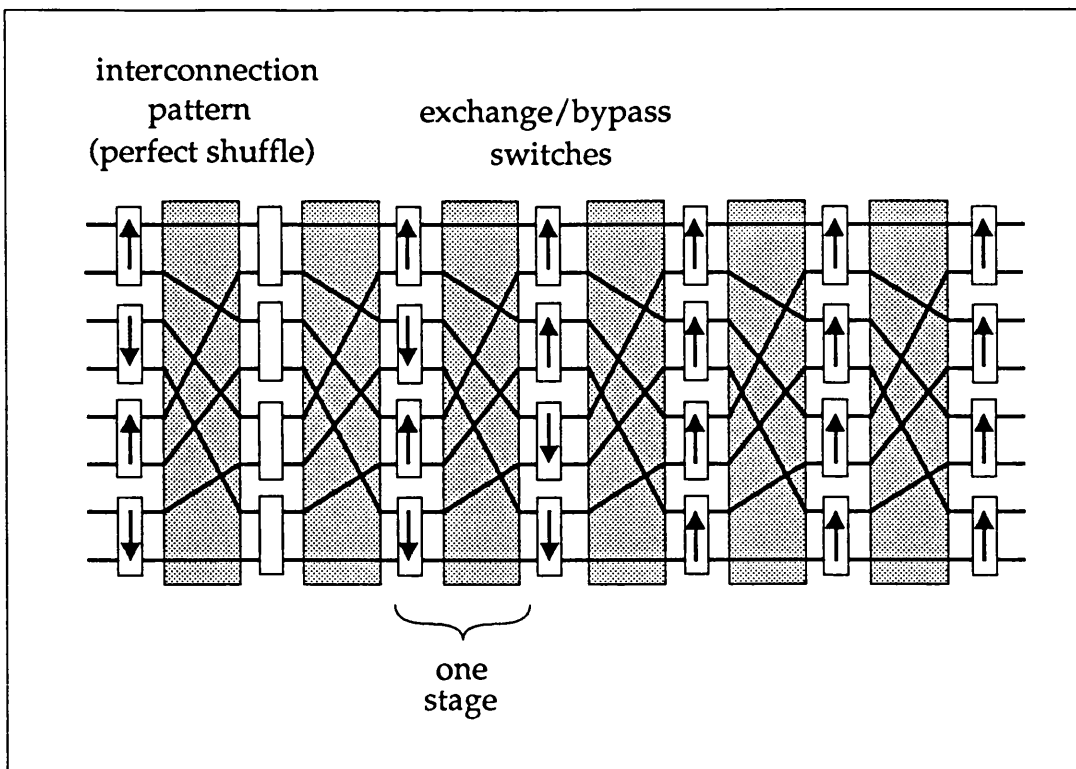


Figure 2.1 A self-routing multistage interconnection network based on bitonic sorting. A mixture of destination addresses is input at the left, and boxes always move the channel with the higher destination in the direction of the arrow. The channels at the right are sorted.

apply to these parallel processing machines as much as to the interconnection networks.

The place where optics can be applied to the multistage networks is at the interconnection pattern between stages, which is the most difficult part of the system to get right with electronics, but which becomes quite simple using optics. The interconnection patterns like the perfect shuffle have already been demonstrated with conventional and with holographic optical components [22-24]. A very efficient packet switch has been proposed [25] by using a self-routing network where the control decisions are made at the switching nodes. If a mixture of destination numbers is input at the left hand side of the network shown in figure 2.1, and each module exchanges or bypasses so that the higher destination emerges from the port indicated by the arrow (empty modules always bypass), then the destinations at the output will be in the correct order. Packet format data could be properly routed without any central control unit, and the limit to the throughput rate comes from the speed at which the individual islands can clock data in and out.

The hybrid MIN implementation with its free space parallel optical

interconnect already falls into the optimal category of electronic islands, the switching modules, linked together with optical interconnects, but it is not ideal. The contribution made by this writer is to say how to adapt the multistage networks so that they have the optimum electronic island size, and they make the best use of the optical imaging system by operating on two dimensions.

2.2 Two dimensional networks

Optics works for two dimensional images by nature, but the optical MINs which are direct analogues of the originals [26] operate on one dimensional data arrays and so fail to exploit fully the parallelism available. The advantage of moving to two dimensions can be quantified thus: there will be a limit to the number of parallel channels that a hybrid network can accommodate, which could come from either the maximum width of the optoelectronic chip or from the distance across which the optics can image faithfully, and if this limit is N for a planar network, it becomes N^2 for a two dimensional network. Equivalently, a 2D network is much more compact than the flat network having the same number of lines.

Other workers have noted the requirement for 2D processing. The responses to the problem include the 2D TIADM blocking network of Cloonan et al. [27], a 2D Clos network [28] and new two dimensional perfect shuffle interconnection patterns [22,29]. A two dimensional analogue of one of the ordinary MIN configurations, which could possess the property of rearrangeability, would have the form of the 2D perfect shuffle network of figure 2.2. The processing part is expanded to two dimensions so that the switching modules have four inputs and four outputs. It has been noted [30] that while such a network appears very promising, no algorithm exists for generating an arbitrary permutation of the input array. In this work this control problem is tackled by starting from an established planar multistage network and mapping it into a 2D MIN. The idea of mapping one network into another has been used before to demonstrate isomorphisms between different 1D networks [13,31,32], but this is the first work where it is used to show the equivalence of 1D and 2D networks.

2.2.1 Perfect shuffle networks

The simplest interconnection to deal with is the perfect shuffle. A perfect shuffle [7] consists of splitting a stack of elements into two halves and interleaving them, as shown in figure 2.3a. For $N=2^n$ elements the transformation can be expressed as the barrel rolling of each element's

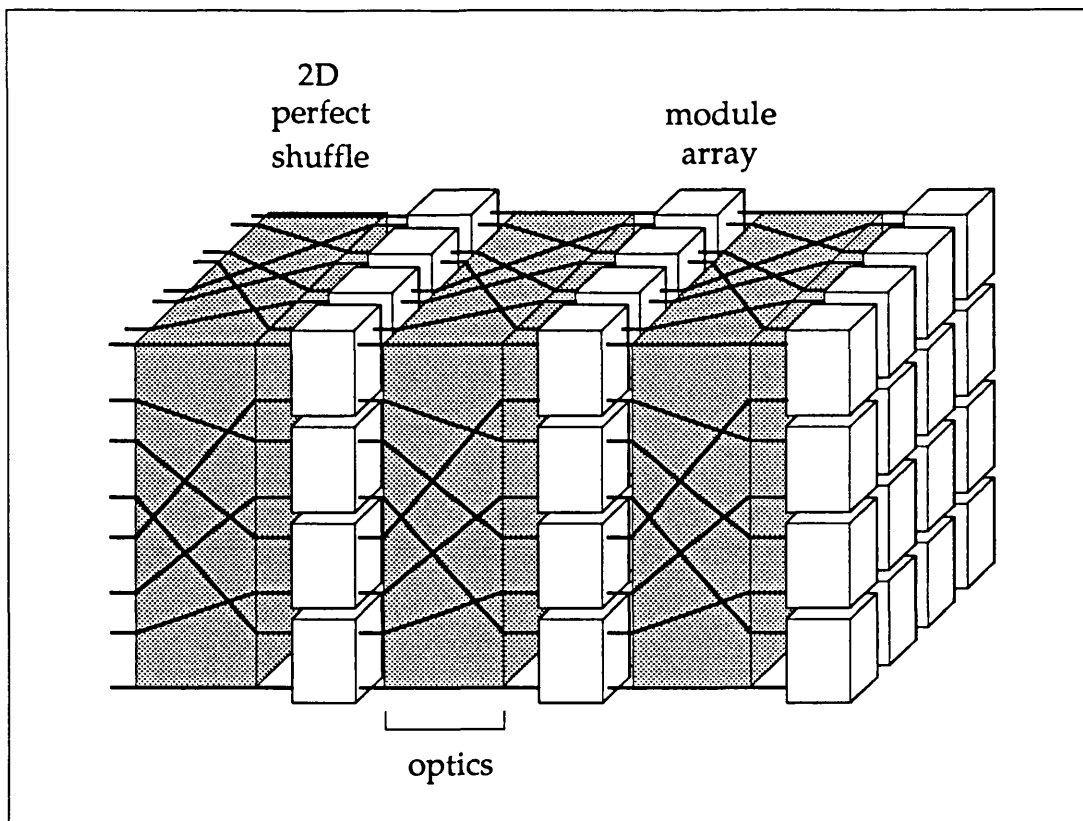


Figure 2.2 The form of a two dimensional network based on 2D perfect shuffles.

address written in binary - the most significant bit is moved to the right hand side of the address and the remaining bits are shunted along one space. For 64 ports

$$abcdef \rightarrow bcdefa$$

where the a to f are 0s and 1s. The perfect shuffle is the interconnection in many MINs and related parallel processing networks [19] and has been extensively studied because of its unique mathematical simplicity.

The two dimensional perfect shuffle invented by Lohmann [22] operating on a square data array is a horizontal followed by a vertical perfect shuffle. Like the 1D counterpart it can easily be generated optically - an example of the optical system required is shown in figure 2.4. Mathematically the 2D PS is represented as barrel rotations of both the column and row parts of the binary address of each array element. For an 8x8 array the transformation is written as

$$(jkl, pqr) \rightarrow (klj, qrp)$$

where once again the j to r are binary digits. The four ports which are linked by one of the the 2D modules are those which differ only in the final bit of the column and row parts of the addresses, that is bits l and r in (jkl, pqr) .

We will now show how to map a planar network, all of whose

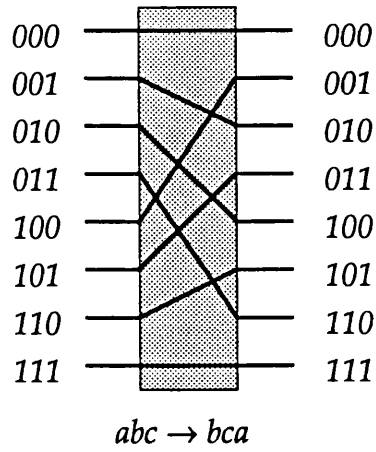


Figure 2.3a

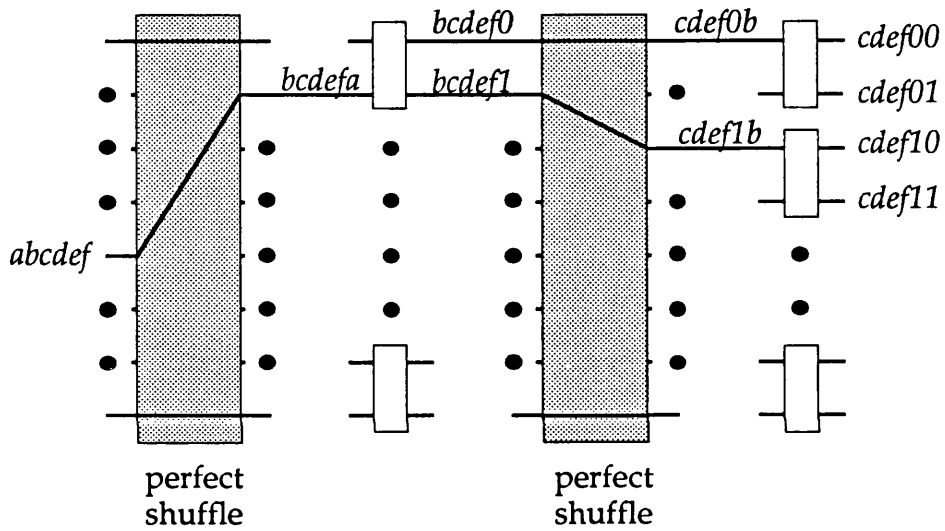


Figure 2.3b

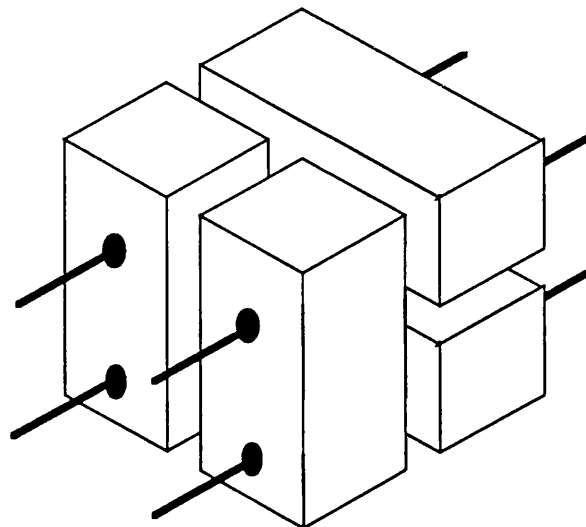


Figure 2.3c

Figure 2.3a A perfect shuffle of eight elements.
b The routing effects of two perfect shuffle-exchange stages in sequence.
c How a 2D exchange module breaks down into four 1D modules.

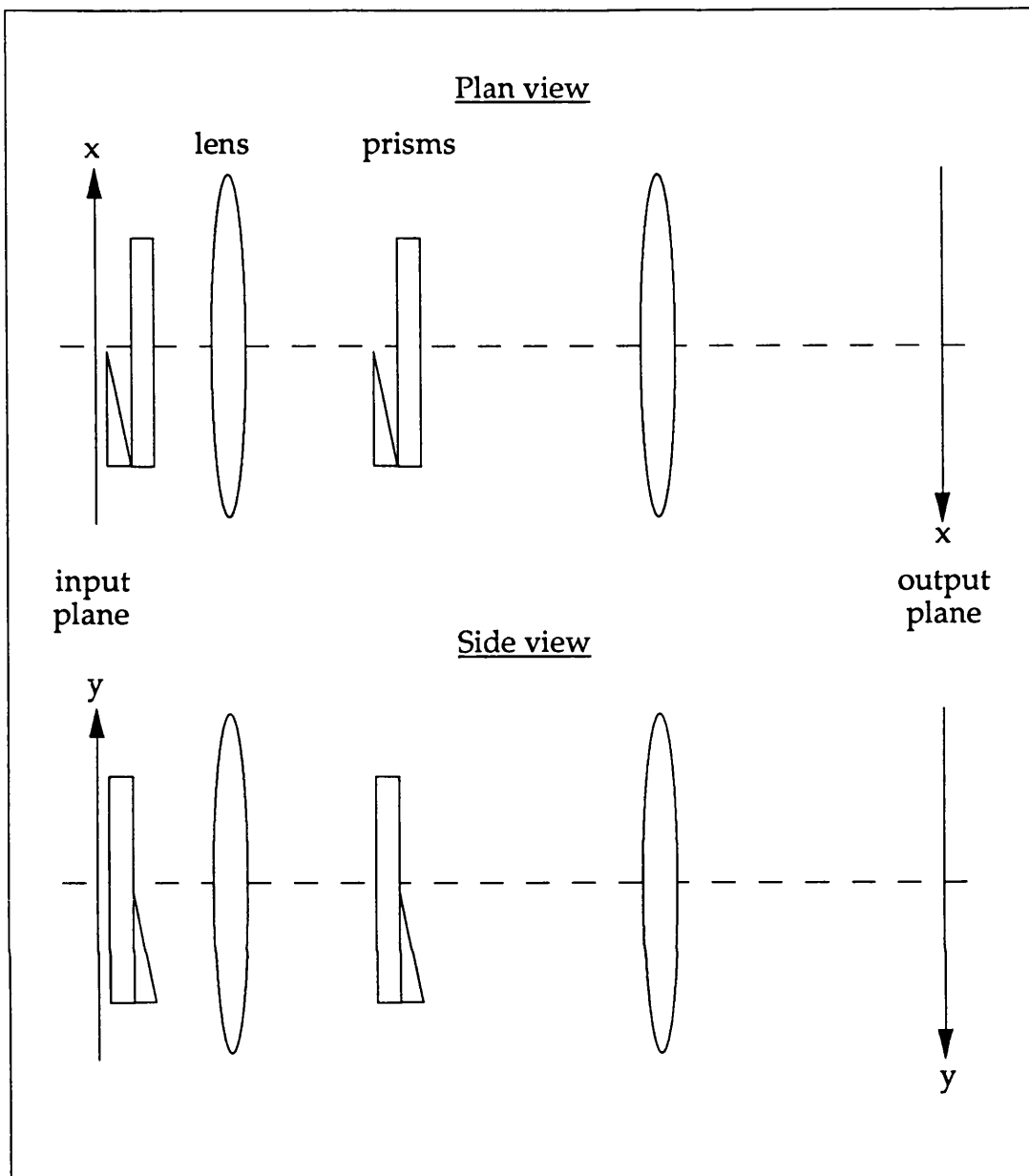


Figure 2.4 Optical arrangement to generate a two dimensional perfect shuffle, using prisms and cylindrical lenses. After Lohmann [22].

interconnections are full perfect shuffles, into a 2D perfect shuffle network as in figure 2.2. The number of stages is halved. It is sufficient to demonstrate that one stage of the new 2D net has the same routing capability as two consecutive stages of the planar network.

Consider two such stages of a 64 channel network as drawn in figure 2.3b. Port $abcdef$ is shuffled to $bcdefa$, and at the exchange module it has a choice of two outputs $bcdef0$ or $bcdef1$. We can say that it is switched to $bcdefA$, where A is 0 or 1. After another perfect shuffle and an exchange module it is sent to $cdefAB$. In a similar manner, one stage of a 2D PS network has the routing effect

$$(jkl, pqr) \rightarrow (klJ, qrP)$$

where bits J and P are 0 or 1 depending on the switch settings. Now let us pick up the lines in the 1D network and move them in space according to the mapping rule

$$abcdef \rightarrow (bdf, ace)$$

which can be considered as the reverse perfect shuffle of the bits of the address. The lines on the right hand side of figure 2.3b, that is those which have been processed by two MIN stages are mapped thus

$$cdefAB \rightarrow (dfB, ceA)$$

Hence one stage of the 2D network does the same routing as two stages of the 1D net.

One of the 2D modules does the same work as four of the old 1D modules when they are arranged in the topology of figure 2.3c. The mapping rule obviously defines to which of the original binary modules each sub-module corresponds, and so the function of the 2D module is completely specified. It has been shown that a 2D perfect shuffle MIN is equivalent to a 1D MIN through the mapping transformation we introduced. There has been no assumption about the modules in the 1D MIN, so our result applies to any network incorporating perfect shuffles.

The breakdown of the four port module into two layers of pairs of binary modules is common to all the 2D networks which will be discussed in this chapter. Another feature of the 2D networks which must be taken into account is the apparent scrambling of output destinations. If the outputs of a 64 channel 1D network are numbered in order, then they are mapped onto the grid of figure 2.5. Because this thorough redistribution of the input and output arrays during mapping is done in the same way every time, it is of no consequence when the purpose of the network is simply to interconnect two sets of terminals. If some formal ordering of the matrix of figure 2.5 is required, then the destinations which are fed into the control algorithm can be modified to achieve this. If it is intended to generate some special class of permutations [9] using a short MIN which does not possess the rearrangeable property, then the mapping into two dimensions may have punitive consequences. Kumagai and Ikegawa [33] have shown that a two dimensional omega network can produce orthogonal rotations of a two dimensional array, in analogy with the ordinary omega net, though their 2D network strictly does not belong to our class because their four port switches cannot be broken down into the four sub-module configuration of figure 2.3c.

1	2	5	6	17	18	21	22
3	4	7	8	19	20	23	24
9	10	13	14	25	26	29	30
11	12	15	16	27	28	31	32
33	34	37	38	49	50	53	54
35	36	39	40	51	52	55	56
41	42	45	46	57	58	61	62
43	44	47	48	59	60	63	64

Figure 2.5 The destinations to which channels would be sorted using a self-routing algorithm mapped from 1D to 2D.

2.2.2 Mapping rules for general interconnection patterns

The perfect shuffle maps easily into two dimensions, and so was discussed first. Can any 1D interconnection pattern be mapped into two dimensions?

Figure 2.6 is a version of figure 2.3b showing two consecutive stages of a network employing some general interconnection pattern. Any input channel, say the broad line, has a choice of four final outputs, and a mapping rule can always be found to bring these four outputs together in a square. In order for the network to be mappable there must be three other channels (shown as the fine lines in figure 2.6) which share the four output alternatives. The first stage brings the four inputs into two pairs, and there is a very specific condition on the second of the two interconnection patterns: the outputs of one of the left hand modules, say A, fan out to two nodes in the right hand layer, B and C; the other inputs to B and C must emerge from a common module D in the left hand layer. This is the 'buddy property' of Agrawal [34], who also noted that every interconnection pattern which has been proposed in a MIN possesses the symmetry.

Strictly then, every multistage network whose alternate interconnection patterns possess the buddy property is mappable into two dimensions, but one must add a further condition. There will in general be many possible mapping rules, but it is required that the 2D interconnection pattern

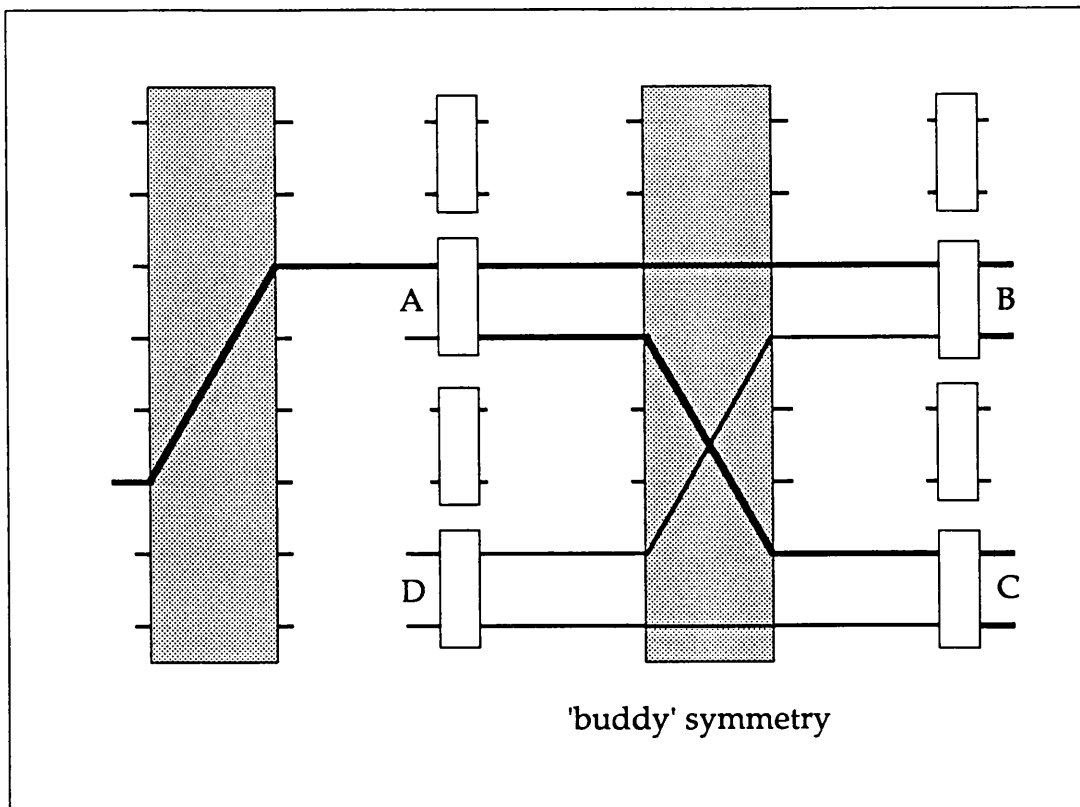


Figure 2.6 Two stages of a general multistage network. In order for the stages to be mappable into two dimensions the second interconnection pattern must possess the 'buddy' property.

generated be amenable to a simple optical implementation. Whether this second condition can be satisfied is a matter of trial and error with the mapping rules. It will be shown in the next section that good mappings can be found for several different interconnection patterns.

2.2.3 Other 2D networks

The perfect shuffle is the simplest interconnection pattern mathematically, but there are others with useful attributes. When free space optical interconnects are used there are definite advantages in using one pattern over another. Because of the magnification step the perfect shuffle inherently involves a 3dB loss in space-bandwidth product; it has been noted that the butterfly [31] and the crossover [35] interconnections are better choices for optical interconnects in MINs. In this section we will develop mapping rules to generate a number of different 2D interconnection patterns.

2.2.3.1 Reverse perfect shuffles and granular perfect shuffles

A reverse perfect shuffle, as put to use in references 12-13, can be mapped into a 2D reverse perfect shuffle using the same mapping rule as before.

Some networks employ 'granular' forward and reverse shuffles [13], by which we mean that the transformation operates separately on several neighbouring subsections of the parallel data lines rather than on the whole column. The granular shuffle is represented as the barrel rotation of only some of the bits in the address, for example a 2-granular shuffle of 64 elements is represented by the barrel rotation of the 5 least significant bits

$$abcdef \rightarrow acdefb$$

If the usual mapping rule is to be applied to two consecutive shuffle stages having different granularity, say the 2-granular shuffle above followed by a full shuffle, then a small adjustment is required. The routing performed by the two stages is

$$abcdef \rightarrow cdefBA$$

which is mapped to

$$(bdf, ace) \rightarrow (dfA, ceB)$$

so that bits *a* and *b* have changed dimension and the 2D interconnection pattern no longer reduces to two orthogonal transformations. The Lohmann 2D shuffle can be used, however, if the 2D exchange module first swaps channels *cdef01* and *cdef10* - a diagonal exchange - and then proceeds with the other processing.

$$(bdf, ace) \xrightarrow{\text{shuffle}} (dfb, cea) \xrightarrow{\text{module}} (dfa, ceb) \xrightarrow{\text{module}} (dfA, ceB)$$

2.2.3.2 Butterfly

The interconnection of the left hand side of figure 2.7a is a full butterfly. Unlike the perfect shuffle, when it is employed in a real network such as the banyan [36] of figure 2.7a, the butterfly interconnections at each stage must necessarily have a different granularity. The full butterfly of $N=2^n$ elements is expressed as the exchange of the most and least significant bits in the binary address, which is

$$abcdefgh \rightarrow hbcdefga$$

for 256 elements. Successively higher degrees of granularity correspond to exchanging the lsb with a bit further to the right in the address than the msb. The first module-interconnection stage of the banyan of figure 2.7a fixes the least significant bit, and subsequent stages fix the digits from the msb onwards.

$$abcdefgh \rightarrow Hbcdefga \rightarrow HACdefgb \rightarrow HABdefgc \rightarrow \dots \rightarrow HABCDEFG$$

The mapping rule we apply to the *r*th stage of the banyan is equal to the mapping applied to the perfect shuffle plus a barrel rotation of the $1/2(n-r)$ least significant bits in the row part of the array address. For example, the

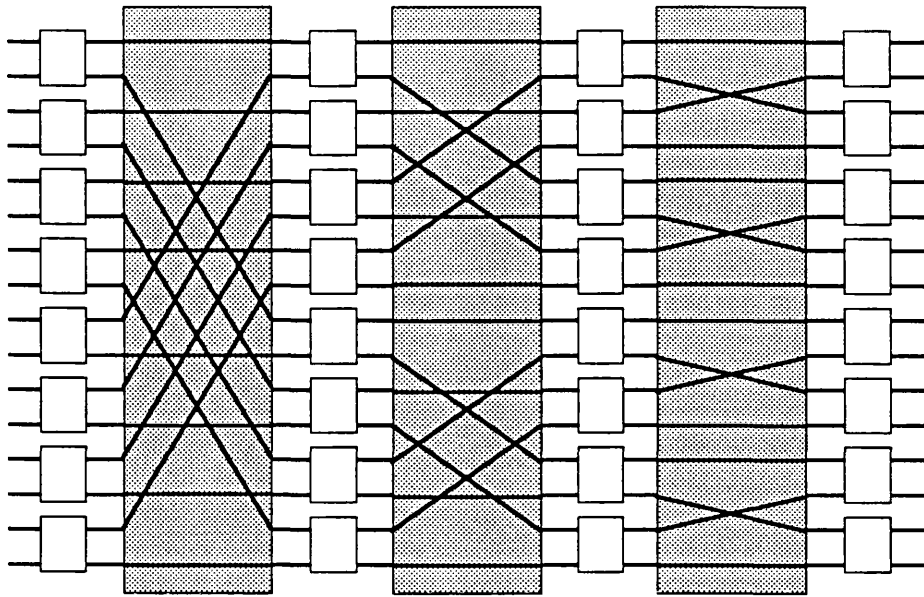


Figure 2.7a

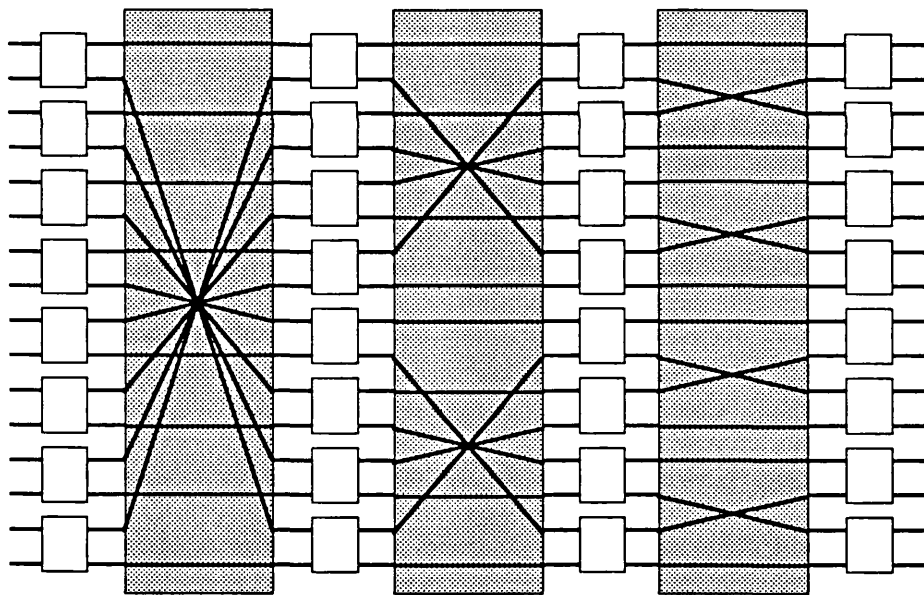


Figure 2.7b

Figure 2.7 Banyan networks based on a the butterfly interconnection pattern and b the crossover pattern.

2nd stage of a 256 port ($n=8$) network is mapped thus

$$abcdefgh \rightarrow (bdfh, aegc)$$

The new 2D interconnection pattern which the mapping produces is a horizontal butterfly accompanied by a vertical butterfly. To verify this consider the interconnection between the 2nd and 4th stages of the original banyan network. In terms of the addresses at the start, the mapping of stage 2 is

$$HAcdefgb \rightarrow (Adfb, Hegc)$$

and that of stage 4 is

$$HABCefgd \rightarrow (ACfd, HBge)$$

so that the 2D transformation between the two stages is

$$(Adfb, Hegc) \rightarrow (ACfd, HBge)$$

The two bits which are to be fixed, b and c , are the right hand bits of the column and row addresses as is required. The 2D exchange module can diagonally swap b and c between column and row dimensions of the address, so that the 2D interconnect left is indeed a horizontal together with a vertical butterfly. All pairs of stages of the planar banyan network are successfully mapped to 2D butterflies of differing granularity, and we conclude that a useful mapping rule has been found.

2.2.3.3 Crossover

The crossover pattern [32,35] was invented specifically because of its suitability for free space optical interconnects. A version of the banyan using crossovers is shown in figure 2.7b; it has the same structure in terms of granularity as the butterfly banyan (figure 2.7a). There is a planar mapping to transpose a butterfly into a crossover pattern [32], and if this 1D to 1D mapping is applied separately to the horizontal and vertical components of the 2D butterfly created above, a new 2D network based on a 2D crossover interconnection will be generated, where the 2D crossover comprises horizontal and vertical crossover interconnections.

2.3 Controlling the electronic island size

We have discussed how to make a two dimensional network. The other improvement suggested was to fix the electronic island to be at its optimum size.

There has been a common basis to all the mapping procedures discussed so far - two stages of a known multistage net are contracted into one stage of the 2D net, and the larger switching unit which is generated does horizontal and vertical exchange operations. What we want is a shorter MIN with larger

switching units. Simply collecting together several of the four input modules and labelling them as one electronic island does not solve the problem, since there are just as many of the expensive optical links as before. We require a mapping scheme which collects together m stages of a binary MIN into one stage of a new network whose switches each have 2^m inputs and 2^m outputs. The new network would have fewer stages than the original by a factor of m .

There are indeed mapping rules to do this to do this which are only slight modifications of the ones we have been dealing with. Consider an ordinary perfect shuffle network of 2^n channels. A slightly different notation will be used in this case; a channel is represented by its binary address

$$a_1a_2..a_n$$

with each $a_i = 0$ or 1 . The routing effect of m perfect shuffle stages is written as

$$a_1a_2..a_n \rightarrow a_{m+1}..a_nA_1..A_m$$

where the A_1 to A_m are the new values of bits a_1 to a_m as set by the individual binary switches. Now we apply the same 1D to 2D mapping as used previously with the perfect shuffle interconnection pattern. The m stages become

$$(a_2a_4..a_n, a_1a_3..a_{n-1}) \rightarrow (a_{m+2}..a_nA_2..A_m, a_{m+1}..a_{n-1}A_1..A_{m-1})$$

Because the digits to be fixed A_i are the least significant bits of the column and row components of the right hand side, the channels which are to be exchanged are all grouped together and non-overlapping square switching units can do all the rerouting required. The 2D interconnection left over is

$$(a_2a_4..a_n, a_1a_3..a_{n-1}) \rightarrow (a_{m+2}..a_n a_2..a_m, a_{m+1}..a_{n-1} a_1..a_{m-1})$$

that is barrel rotations by $1/2m$ places of the individual address components. One of these barrel rotations corresponds to a general perfect shuffle [37] or $1/2m$ -shuffle [14]. The transformation has already been demonstrated in one dimension using optical components [38], and a 2D transformation comprising independent operations on the columns and rows should be quite feasible.

We can say, then, that there are mappings which generate a network whose electronic island size is that desired by the circuit designer.

2.4 Further comments on mapped 2D networks

Any of the new networks derived here is isomorphic to its parent. This means that the new network executes an identical set of switching operations, but in a physically different place. Any property of the original net which does not relate to the actual locations of the channels is retained by the derived network; a few examples are discussed in this section.

In the class of rearrangeably non-blocking switches, the network which requires the minimum switching plant, evaluated as number of crosspoints [6] or number of binary decisions [8], is the MIN with the smallest switching units, that is the Benes MIN with 2x2 switching modules. One might suspect that a 2D network with switches having more than two inputs and outputs does not possess this minimum plant property, but the isomorphism rule means that it does. The apparent paradox is resolved when one realises that the new switching module cannot produce as many permutations as a full crossbar of the same size.

The four sub-module configuration of figure 2.3c can sit in any of $2^4 = 16$ states, while a full 4x4 crossbar can produce $4! = 24$ permutations. One might expect the 8 unavailable permutations to compromise the performance of a mapped 2D network. One of the advantages of using large switches is that fewer stages are required to achieve rearrangeability. Let us compare the shortest known network having 4x4 switches with the one generated by our mapping from a binary network. Benes [6] says that a multistage network of 4x4 crossbars can achieve any permutation on $N=2^n$ lines in $(n-1)$ stages. Another of his results is that a network having $(2n-4)$ 2x2 switches and a single central stage of 4x4 crossbars is rearrangeable, and this network maps into a 2D version with $(n-1)$ stages. Curiously then, the mapped network also has the minimum number of stages, even though all but the one central stage makes use of only 16 local permutations. Moreover, unlike the binary Benes net there is no algorithm to control a multistage network of 4x4 crossbars - until this work provided one!

The problem of designing a self routing rearrangeable network, where the routing decision made at any node is based on the destinations of the channels arriving at that node, is the same as the problem of sorting N numbers. The Batcher bitonic sorting network [21] is the network which requires the minimum number of decisions, and once again the 2D version of the bitonic sorter can also lay claim to using the fastest possible algorithm. We note that the various optimal properties of binary networks are not lost in mapping them into networks with larger switching modules.

2.5 Evaluation of a practical hybrid network

We can now form a picture of an optically interconnected multistage switching network. It would have optical systems like the one in figure 2.4, with optoelectronic circuits as the input and output planes. The design issues for the optics are being investigated [39, 40], and experiments are underway into Banyan networks based on MQW modulator arrays [41].

What performance can we expect? The limit to the data rate comes from the electronic circuitry and the speed at which the optical modulators or sources can work. One would hope that each island should be able to operate at 200 Mbit/s. The optics determines the number of channels - a figure of 1024 (a 32x32 array) should be feasible. This gives a throughput of 200 Gbit/s - a very impressive number.

Trying to predict the capabilities of electronic technology is obviously very difficult. Reference 42 lays out the present capability of both CMOS and GaAs crosspoint circuits, and the figure above for our optoelectronic switch is clearly ahead of these technologies, in terms of both data throughput and number of channels.

It seems likely that very high throughput switches will be required in the future, and compact size may also be demanded. Purely as an example, in the next two decades trunk telecommunications data may be switched at satellites in space. It would be mandatory to use low mass equipment in this scenario, even if that involves going to new technologies.

2.6 Conclusions

It has been established that optical interconnects are useful for parallel array processing, and for replacing long range electrical interconnects. While the multistage network makes excellent use of the positive features of optics, it can be improved upon. It has been shown by a series of examples that conventional MINs can be mapped into even more suitable forms for optical implementations: in particular they can be made to be two dimensional, and to have switching units which can be fixed at the size of the optimum electronic island. What we have arrived at, then, is a blueprint for a system which makes the best use of optical interconnects, and a brief evaluation indicated that the system could be superior to an all-electronic one.

2.7 References

- [1] J. W. Goodman, F. I. Leonberger, S. Y. Kung, and R. A. Athale, "Optical interconnections for VLSI systems," Proc. IEEE, vol. 72, pp. 850-866, 1984.
- [2] R. K. Kostuk, J. W. Goodman & L. Hesselink, "Optical imaging applied to microelectronic chip-to-chip interconnections," Appl. Opt., vol. 24, pp. 2851-2858, 1985.
- [3] M. R. Feldman, S. C. Esener, C. C. Guest, and S. H. Lee, "Comparison between optical and electrical interconnects based on power and speed considerations," Appl. Opt., vol. 27, pp. 1742-1751, 1988.

- [4] D. A. B. Miller, "Optics for low energy communication inside digital processors: quantum detectors, sources and modulators as efficient impedance converters," *Opt. Lett.*, vol. 14, pp. 146-148, 1989.
- [5] J. E. Midwinter, "Digital optics, smart interconnect or optical logic, Part 1," *Phys. Technol.*, vol. 19, pp. 101-108, 1988.
J. E. Midwinter, "Digital optics, smart interconnect or optical logic, Part 2," *Phys. Technol.*, vol. 19, pp. 153-165, 1988.
- [6] V. E. Benes, "Mathematical theory of interconnecting networks and telephone traffic," New York: Academic, 1965.
- [7] H. S. Stone, "Parallel processing and the perfect shuffle," *IEEE Trans. Comput.*, vol. C-20, pp. 153-161, 1971.
- [8] D. C. Opferman and N. T. Tsao-Wu, "On a class of rearrangeable switching networks, Parts 1 & 2," *Bell Syst. Tech. J.*, vol. 50, pp. 1579-1618, 1971.
- [9] D. H. Lawrie, "Access and alignment of data in an array processor," *IEEE Trans. Comput.*, vol. C-24, pp. 1145-1155, 1975.
- [10] C.-L. Wu and T.-Y. Feng, "The universality of the shuffle-exchange network," *IEEE Trans. Comput.*, vol. C-30, pp. 324-332, 1981.
- [11] D. S. Parker, "Notes on shuffle/exchange type switching networks," *IEEE Trans. Comput.*, vol. C-29, pp. 213-222, 1980.
- [12] C.-L. Wu and T.-Y. Feng, "On a class of multistage interconnection networks," *IEEE Trans. Comput.*, vol. C-29, pp. 694-702, 1980.
- [13] C.-L. Wu and T.-Y. Feng, "The reverse exchange interconnection network," *IEEE Trans. Comput.*, vol. C-29, pp. 801-811, 1980.
- [14] J. H. Patel, "Performance of processor-memory interconnections for multiprocessors," *IEEE Trans. Comput.*, vol. C-30, pp. 771-780, 1981.
- [15] S.-T. Huang and S. K. Tripathi, "Finite state analysis tools for permutation networks: new analysis tools for permutation networks," *IEEE Trans. Comput.*, vol. C-35, pp. 591-601, 1986.
- [16] T. H. Szymanski and V. C. Hamacher, "On the permutation capability of multistage interconnection networks," *IEEE Trans. Comput.*, vol. C-36, pp. 810-822, 1987.
- [17] W.-T. Chen and J.-P. Sheu, "Performance analysis of multistage interconnection network with hierarchical requesting model," *IEEE Trans. Comput.*, vol. 37, pp. 1438-1442, 1988.
- [18] A. Huang & S. Knauer, "Starlite: a wideband digital switch," *Proc. of CLOBECOM '84*, Nov. 1984.
- [19] H. S. Stone, "High performance computer architecture," ch. 4, Addison-Wesley, 1987.
- [20] M. C. Pease, "An adaptation of the fast Fourier transform for parallel processing," *J. ACM*, vol. 15, pp. 297-301, 1965.

- [21] D. E. Knuth, "Sorting and searching," Addison-Wesley, 1973.
- [22] A. W. Lohmann, "What classical optics can do for the digital optical computer," *Appl. Opt.*, vol. 25, pp. 1543-1549, 1986.
- [23] K. H. Brenner and A. Huang, "Optical implementations of the perfect shuffle interconnection," *Appl. Opt.*, vol. 27, pp. 135-137, 1988.
- [24] Y. Sheng and H. H. Arsenault, "Light effective perfect shuffle using Fresnel mirrors," *Optical Computing, 1989 Technical Digest Series*, vol. 9, (OSA 1989) pp. 154-157.
- [25] J. E. Midwinter, "Novel approach to the design of optically activated wideband switching matrices," *IEE Proc. J*, vol. 134, pp. 261-268, 1987.
- [26] G. E. Lohman and A. W. Lohmann, "Optical interconnection network utilising diffraction gratings," *Opt. Eng.*, vol. 27, pp. 893-900, 1988.
- [27] T. J. Cloonan and M. J. Herron, "Optical implementation and performance of one-dimensional and two-dimensional trimmed inverse augmented data manipulator networks for multiprocessor computer systems," *Opt. Eng.*, vol. 28, pp. 305-314, 1989.
- [28] S.-H. Lin, T. F. Krile, and J. F. Walkup, "Two-dimensional optical Clos network and its uses," *Appl. Opt.*, vol. 27, pp. 1734-1741, 1988.
- [29] C. W. Stirk and R. A. Athale, "Folded perfect shuffle optical processor," *Appl. Opt.*, vol. 27, pp. 202-203, 1988.
- [30] S.-H. Lin, T. F. Krile, and J. F. Walkup, "2-D optical multistage interconnection networks," *SPIE*, vol. 752, pp. 209-216, 1987.
- [31] M. J. Murdocca, A. Huang, J. Jahns, and N. Streibl, "Optical design of programmable logic arrays," *Appl. Opt.*, vol. 27, pp. 1651-1660, 1988.
- [32] T. J. Cloonan, "Topological equivalence of optical crossover networks and modified data manipulator networks," *Appl. Opt.*, vol. 28, pp. 2494-2498, 1989.
- [33] T. Kumagai and K. Ikegaya, "The two dimensional inverse omega network," *Int. Conf. Parallel Processing 1985*, pp. 325-327.
T. Kumagai and K. Ikegaya, "Organization of the two-dimensional omega network," *Systems and Computers in Japan*, vol. 17, pp. 1-10, 1986.
- [34] D. P. Agrawal, "Graph theoretical analysis and design of multistage interconnection networks," *IEEE Trans. Comput.*, vol. C-32, pp. 637-648, 1983.
- [35] J. Jahns and M. J. Murdocca, "Crossover networks and their optical implementation," *Appl. Opt.*, vol. 27, pp. 3155-3160, 1988.
- [36] L. R. Goke and G. J. Lipovski, "Banyan networks for partitioning multiprocessor systems," in *Proc. 1st Int. Symp. Comput. Arch.*, FL, Dec. 1973, pp. 21-28.
- [37] L. N. Bhuyan and D.P. Agrawal, "Design and performance of generalized interconnection networks," *IEEE Trans. Comput.*, vol. C-32,

pp. 1081-1090, 1983.

- [38] G. Eichmann and Y. Li, "Compact optical generalized perfect shuffle," *Appl. Opt.*, vol. 26, pp. 1167-1169, 1987.
- [39] M. E. Prise, N. Streibl & M. M. Downs, "Optical considerations in the design of digital optical computers," *Opt. Quantum Electron.*, vol. 20, pp. 49-77, 1988.
- [40] F. B. McCormick & M. E. Prise, "Optical circuitry for free-space interconnections," *Appl. Opt.*, vol. 29, pp. 2013-2018, 1990.
- [41] T. J. Cloonan, M. J. Herron, F. A. P. Tooley, G. W. Richards, F. B. McCormick, E. Kerbis, J. L. Brubaker & A. L. Lentine, "An all-optical implementation of a 3-D crossover switching network," *IEEE Phot. Tech. Lett.*, vol. 2, pp. 438-440, 1990.
- [42] J. E. Berthold, "Broadband electronic switching," *IEEE Lightwave Comm. Systems Magazine*, May 1990, pp. 35-39.

Multiple Quantum Well Modulators

This chapter will introduce the multiple quantum well modulator, and discuss some of this writer's ideas on how to make measurements of the quantum well absorption and on modulator optimisation. The next chapter will discuss SEED devices.

3.1 History of MQW modulator development

The idea of a quantum well is as old as the Schrödinger equation. It has only been quite recently that semiconductor technology has been able to demonstrate the confined energy levels predicted by the textbook problem.

The advanced semiconductor growth techniques of metalorganic vapour phase epitaxy (MOVPE, also known as MOCVD), molecular beam epitaxy (MBE) or gas source MBE are necessary to make the thin ($\approx 100\text{\AA}$) layers in which quantum confinement effects are observed, and all the earliest work was in the GaAs:AlGaAs system. The first observation of the stepped absorption spectrum associated with quantum wells was by Dingle et al. [1] in 1974 at liquid helium temperatures. The first devices to be developed using the technology were quantum well lasers [2], where concentrating the density of states function at the bandgap energy means a lower laser threshold current density. In 1981 excitonic features were observed at room temperature by Ishibashi et al. [3]. The following year Mendez et al. [4] saw a shift to lower energies of exciton peaks in the low temperature photoluminescence spectrum of quantum wells with a transverse electric field. They explained the observations as being due to the reorganisation of the electron and hole wavefunctions; their mechanism for applying the field was with a Schottky contact on the surface of the sample. Chemla et al. then observed the effect of electric field on the absorption spectrum of an MQW sample using a twin electrode arrangement to apply field [5].

The first people to incorporate the quantum wells into the intrinsic region of a pin diode in the now standard MQW electroabsorption modulator structure were Wood et al. [6]. They obtained a contrast of 2:1 with 8V bias in a transverse modulator, and later a contrast of 10:1 from a waveguide modulator [7]. The switching time of 131ps [8] was shown to be limited by a capacitive time constant. Miller gave a thorough explanation

of the modulation mechanism [9] and called it the quantum confined Stark effect (QCSE) [10]. The first person to model the MQW absorption spectrum taking all the exciton broadening mechanisms into account was Stevens [11], who was able to match real spectra to theoretical, and so predict the limits to the performance of transverse modulators [12]. Boyd et al. demonstrated a double pass reflection modulator with a multilayer stack underneath the quantum wells [13], and Whitehead showed that very high contrasts are possible with a similar structure operated as an absorption switched Fabry-Perot cavity [14].

Other material systems than the GaAs:AlGaAs system have been investigated, the motivation being to operate at communications wavelengths and to do away with the need for reflector stacks and substrate removal by growing on a transparent substrate. Recently modulators have been demonstrated in the InGaAs(well):InAlAs(barrier) system [15,16], and the InGaAs:InP system [17,18], both grown on indium phosphide. It seems that the absorption coefficients are lower for the long wavelength materials, and so good low voltage modulation may not be possible [19]. There has also been interest recently in growing GaAs:AlGaAs quantum wells on silicon substrates [20,21], so that the modulators can be linked directly with silicon circuitry.

3.2 Operation of MQW Stark effect modulators

A semiconductor quantum well is a thin layer of low bandgap semiconductor between two layers having a wider bandgap. The easiest materials to work with are gallium arsenide wells and aluminium gallium arsenide barrier layers, because these solids have almost the same lattice constant over the whole range of aluminium concentration. The carriers in the well see the higher band energy of the surrounding material as 'walls' and are confined in the same way as the potential well problem of quantum mechanics (figure 3.1a). In the GaAs:AlGaAs system both electrons and holes are confined. There are a series of energy levels, and the effective bandgap of the quantum well is larger than that of the bulk well material. The carriers are free to move in directions parallel to the well, and the density of states function is modified from the parabola associated with bulk material to a stepped function, as shown in figure 3.1b. A new step rises at each energy level of the well.

The optical absorption spectrum follows the density of states function, and so has a steplike structure. The quantum well material thus has a steeper band edge than bulk material. In addition there are peaks in the absorption

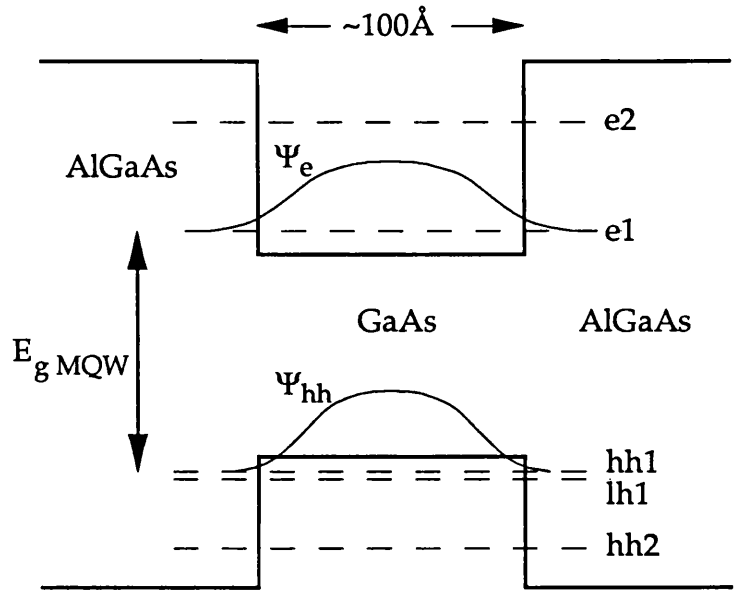


Figure 3.1a

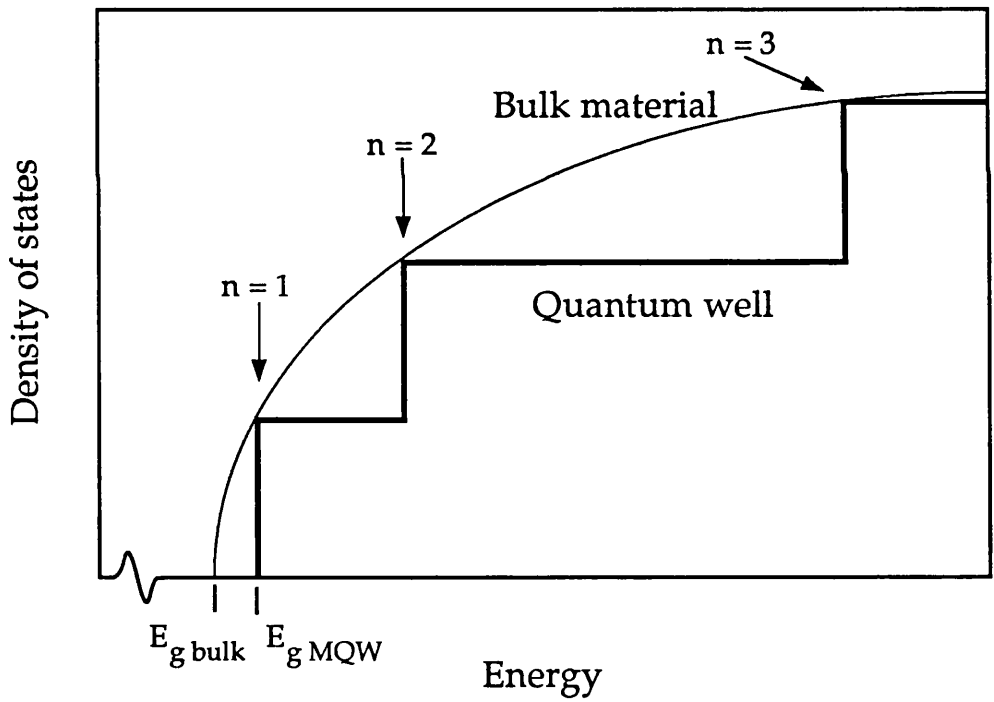


Figure 3.1b

Figure 3.1a Band diagram of a semiconductor quantum well. Confined energy levels are shown for electron (e), heavy hole (hh) and light hole (lh), together with the first level wavefunctions for the electron (e1) and heavy hole (hh1).

b Density of states function for a quantum well drawn with that of bulk material. The curve for the quantum well has a stepped form.

spectrum due to the formation of excitons, bound electron-hole pairs, even at room temperatures. Ordinarily exciton peaks are seen only at cryogenic temperatures, but the two dimensional excitons formed in quantum wells have a much larger binding energy than 3D excitons. There are two peaks associated with light and heavy holes (the degeneracy between the two types of hole is lifted by the confinement potential) located just off the band edge, and they act to further enhance the height of the band edge.

When an electric field is applied to a quantum well the energy bands are tilted (figure 3.2b) and the solutions of the Schrödinger equation are modified. The first eigenstate of the tilted well has a lower energy than the zero field well. This means that the bandgap energy of the well material is reduced on the application of an electric field. The carrier wavefunctions are concentrated towards the walls of the well, electrons at one side and holes at the other. The height of the exciton resonance (or strictly the area underneath the peak) is proportional to the overlap of the appropriate electron and hole wavefunctions, and so the heights of the peaks are reduced under an applied field. Also, the wavefunctions are no longer symmetric about the mid-well, which means that previously forbidden transitions become allowed and new exciton peaks appear in the absorption spectrum. Figure 3.2a illustrates these effects with absorption spectra at different electric fields for 80Å wells. The red shift of the band edge is known as the quantum confined Stark effect, and is thoroughly discussed in reference 22. It is found that narrow wells retain their exciton peak height better when shifting, but require a larger field than wide wells for the same shift [23]. By operating the device of figure 3.2a at a wavelength of 858nm a large absorption can be switched on and off with a modest voltage change. Typically, a contrast of 3dB is possible with an insertion loss of about 2dB using 8V applied bias in a single pass through 1µm of active region.

The device structure used to apply an electric field to the quantum wells is a pin diode, with the wells in the intrinsic region. Many well layers are grown to increase the amount of active material, and they are as close as possible without any interaction of the wavefunctions of adjacent wells; 60Å of $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ is found to be adequate as a barrier. If the doping of the multiple quantum well region is not very low then modulation is degraded - this issue is discussed in section 3.4. The p and n regions of the pin diode are transparent $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$, and the GaAs substrate must be removed to let the modulated light out. A thin highly doped capping layer is usually grown on the top to enable a good Ohmic contact to be made. Several modulators are isolated from one another on the same chip by etching mesas. Figure 3.3

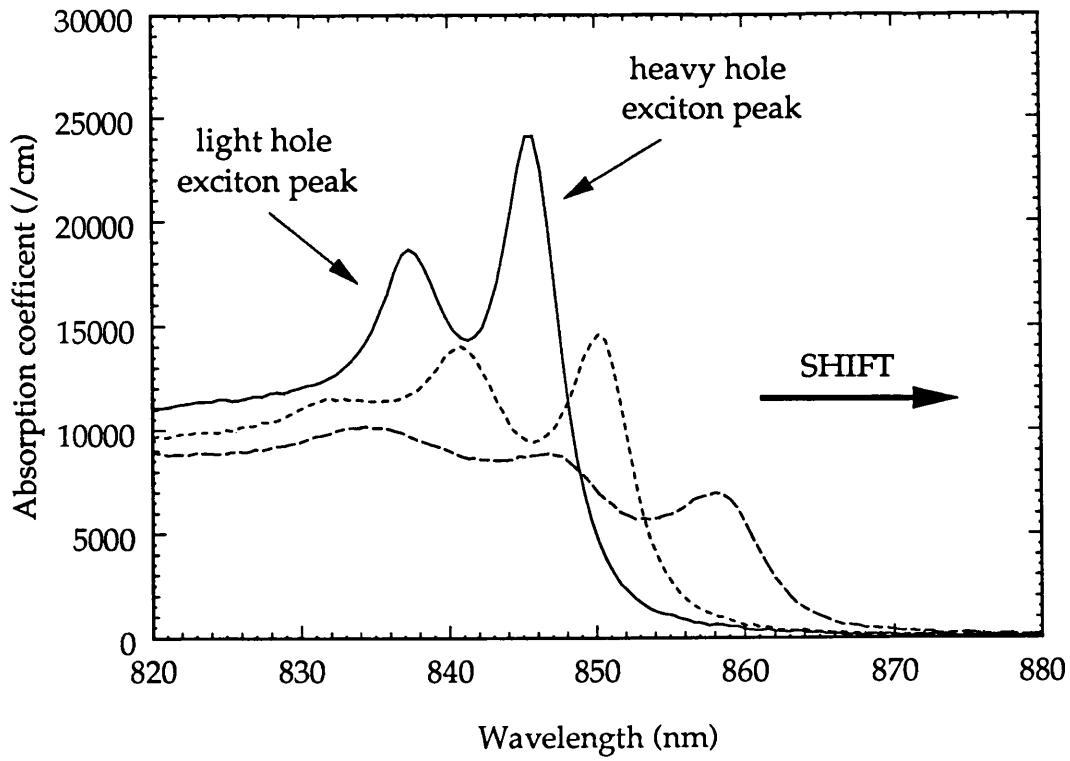


Figure 3.2a

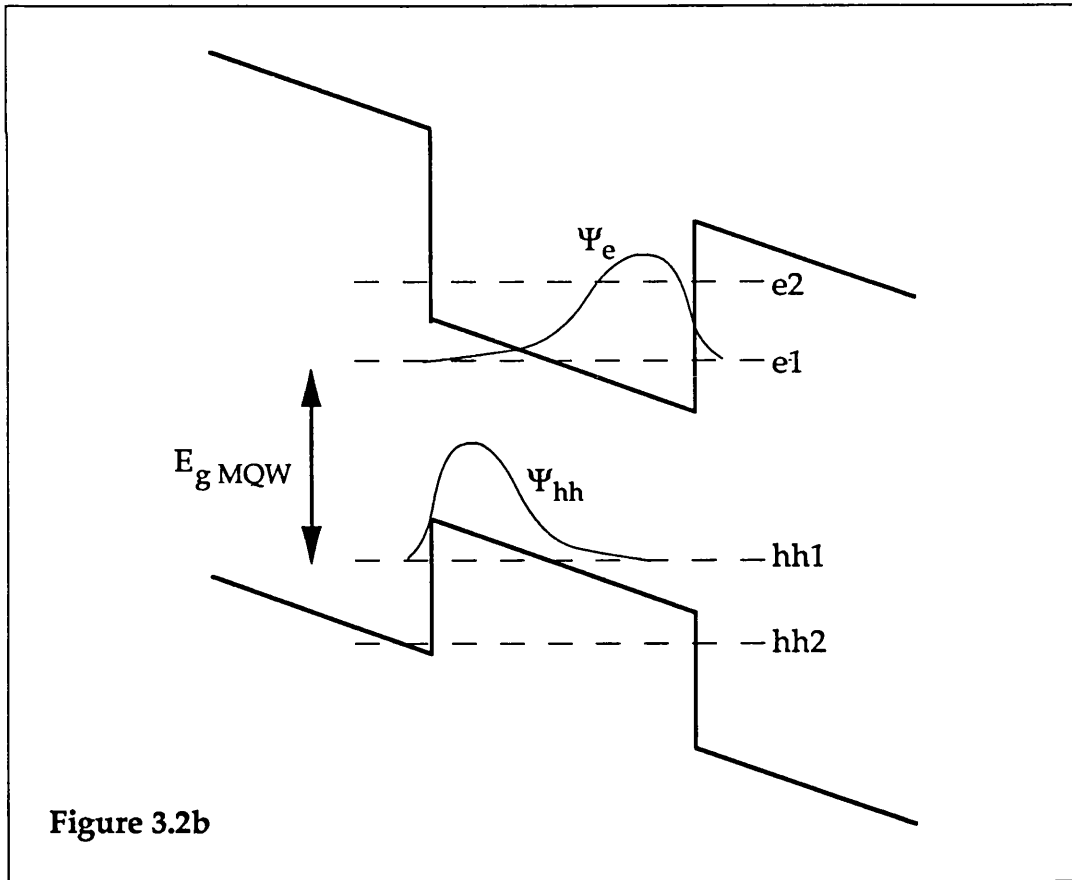


Figure 3.2b

Figure 3.2a Absorption spectra of an 80 Å well at fields of 0 (solid), 50 kV/cm (dashed) and 100 kV/cm (broken line). Data taken by M. Whitehead.
b Band diagram of a quantum well under an electric field.

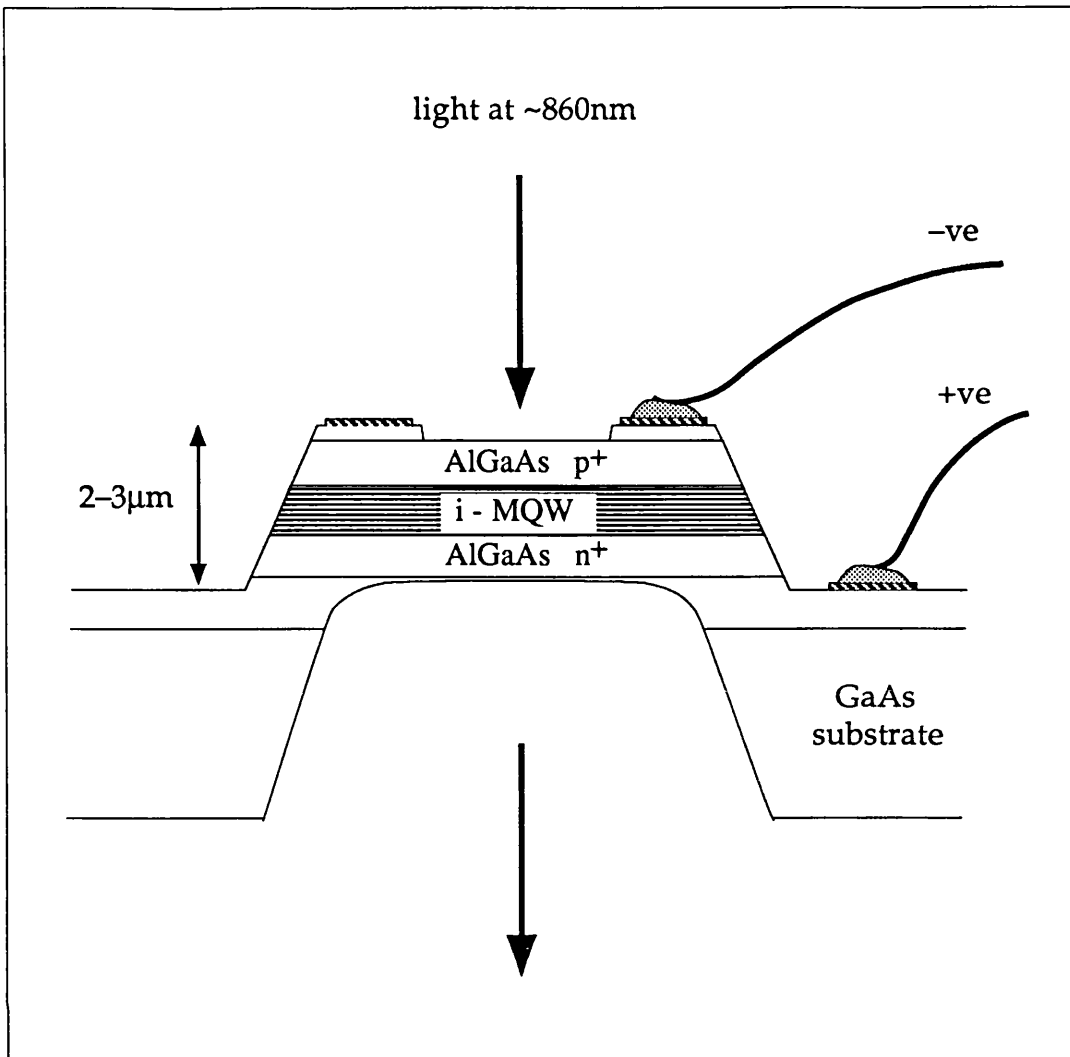


Figure 3.3 An MQW transmission modulator.

shows an MQW transmission modulator. The thin film left behind by windowing the substrate makes the device very fragile and suitable only for experimental investigations. No substrate removal is necessary if a mirror is grown underneath the modulator; a Bragg reflector with 95% reflectivity is formed from 12 periods of quarter-wave $\text{Al}_{0.1}\text{Ga}_{0.9}\text{As}$ and AlAs layers.

3.3 Fabry-Perot cavity effects

The large refractive index mismatch between semiconductor and air means that the interface between the two has a high reflection coefficient. An MQW transmission modulator with the substrate removed is a parallel film which acts as a fairly high finesse Fabry-Perot (FP) cavity. The reflection modulator has a multilayer reflector stack underneath the quantum wells, and so forms a stronger Fabry-Perot cavity. The cavity effects can dominate device spectra; figures 3.4a and b show photocurrent and transmission spectra for a typical MQW modulator where the FP resonances at 820 and 876nm make it difficult

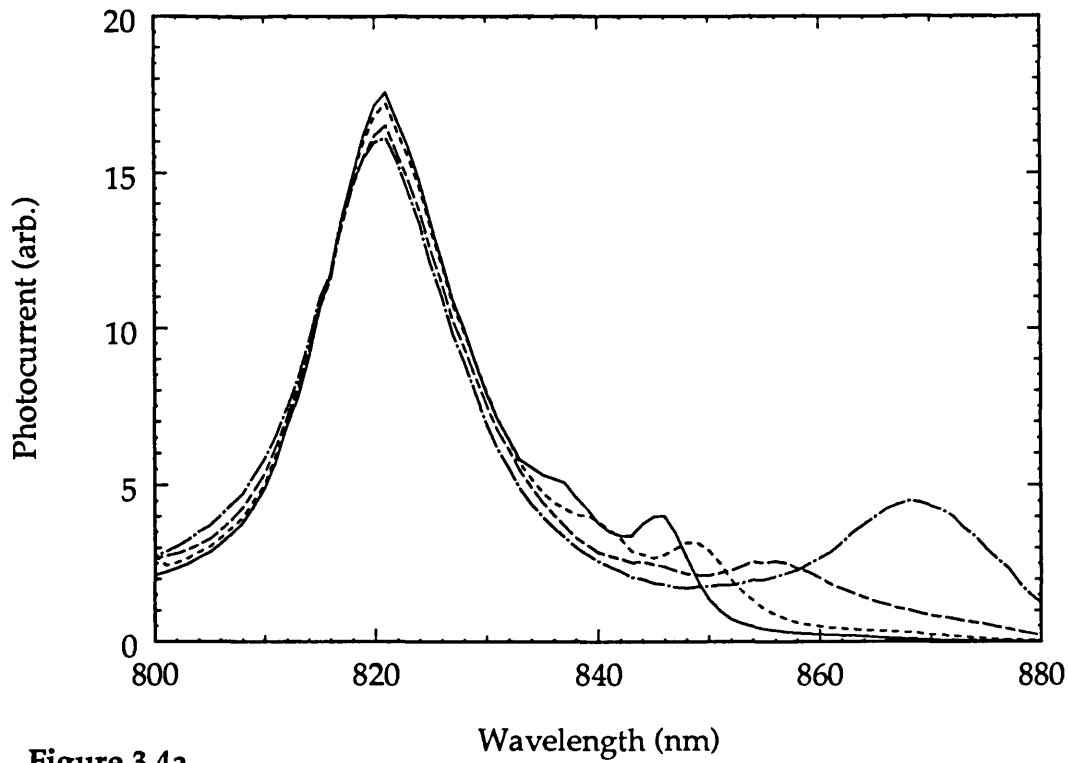


Figure 3.4a

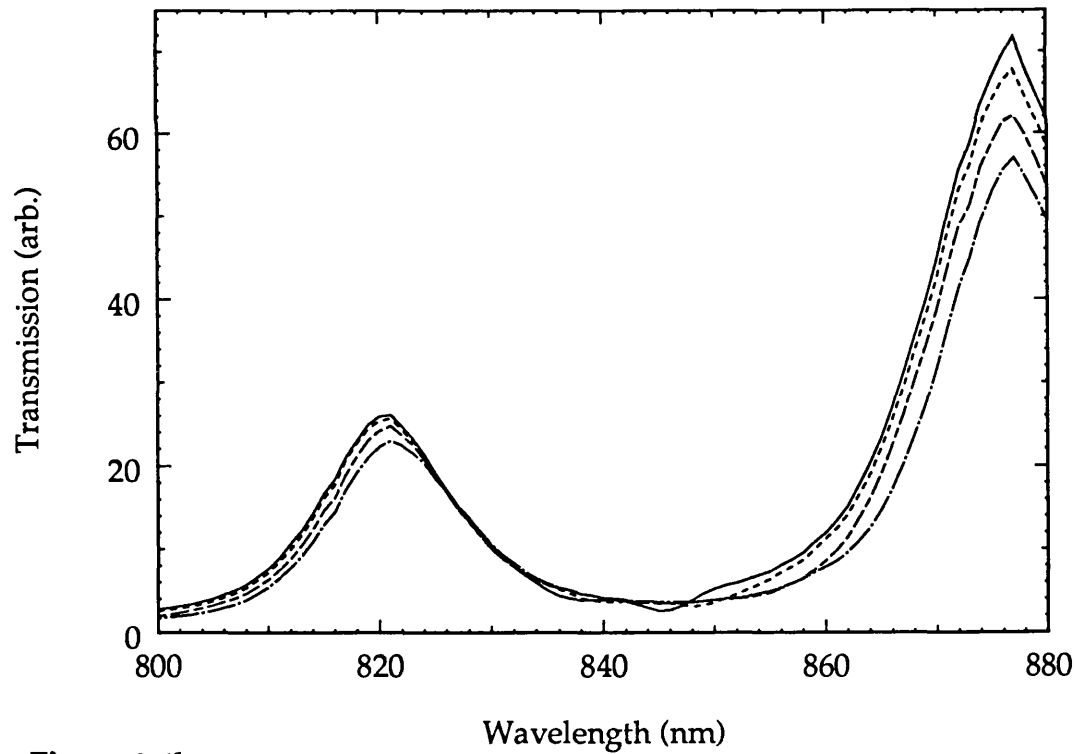


Figure 3.4b

Figure 3.4a Photocurrent and b transmission spectra for a modulator with 80Å wells, at applied voltages of 2V, 6V, 10V and 14V.

to pick out the excitonic features.

There have been proposals to use the interference effects to enhance modulation. References 24 and 25 use the refractive index change associated with the QCSE to switch a high finesse FP cavity on and off. Recently Whitehead noted that absorption change could be used to switch an asymmetric cavity, with high back and lower front reflectivities [26]. It is theoretically possible to obtain zero reflected light, and contrasts in excess of 20dB were obtained with 9V applied bias [14]. The device was grown on gallium arsenide had a multilayer stack as the back reflector and used the uncoated top surface as the front reflector.

The usual way to consider an absorptive Fabry-Perot cavity is through the standard equations relating the transmission T and reflection R to the cavity parameters.

$$\begin{aligned} R &= \frac{B + F \sin^2 \frac{1}{2} \phi}{1 + F \sin^2 \frac{1}{2} \phi} \\ T &= \frac{C}{1 + F \sin^2 \frac{1}{2} \phi} \end{aligned} \quad (3.1)$$

where

$$C = \frac{(1-R_b)(1-R_f)e^{-\alpha d}}{(1-R_\alpha)^2} \quad B = \frac{R_f \left(1 - \frac{R_\alpha}{R_f}\right)^2}{(1-R_\alpha)^2}$$

and

$$F = \frac{4R_\alpha}{(1-R_\alpha)^2} \quad R_\alpha = \sqrt{R_f R_b} \exp(-\alpha d)$$

R_f and R_b are the front and back surface reflectivities, ϕ is the phase length of the cavity, and d is the thickness of absorber with absorption coefficient α . Equations (3.1) are derived by considering the summation of many waves circulating in the cavity (figure 3.5a), with the amplitude of each wave decaying due to a fraction of its energy leaving the cavity and being absorbed during a round trip. The obvious complexity of the standard equations makes it difficult to pick out relationships between parameters. This work will present a new simpler model of the FP cavity which, although it is equivalent to the present model, permits a more intuitive understanding of the phenomena involved.

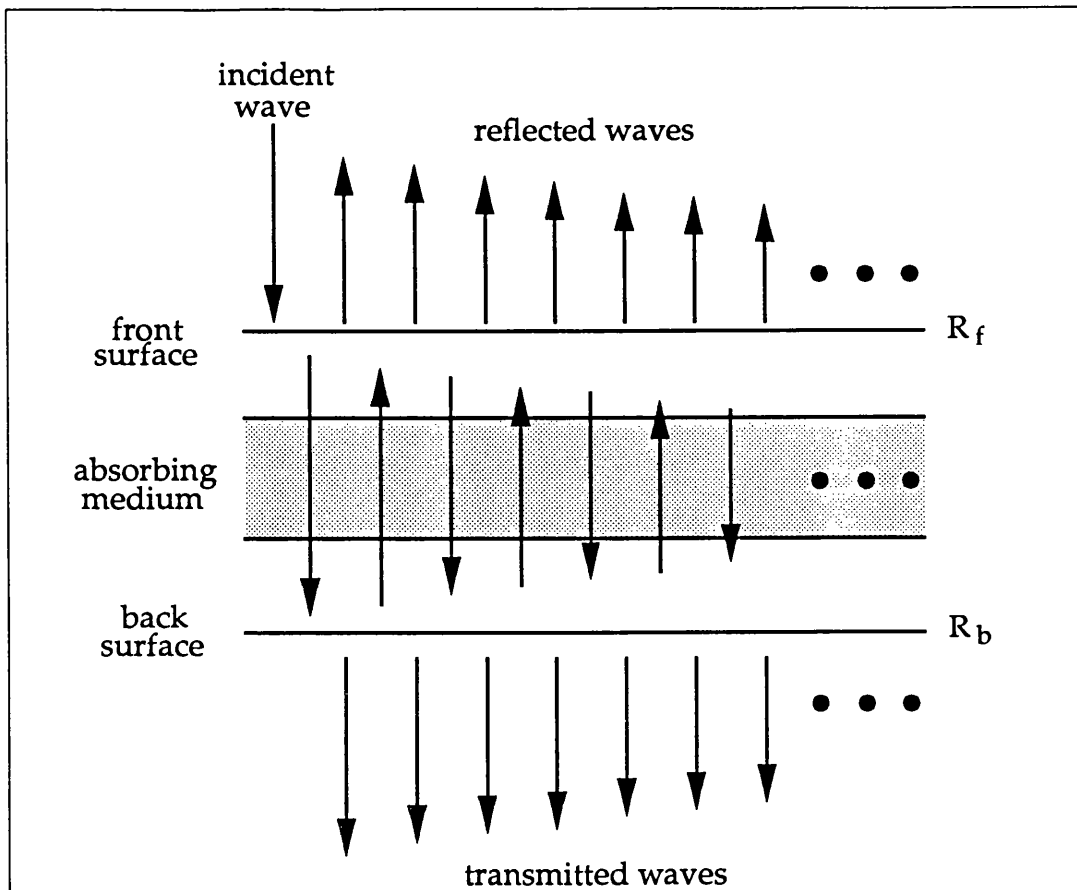


Figure 3.5a

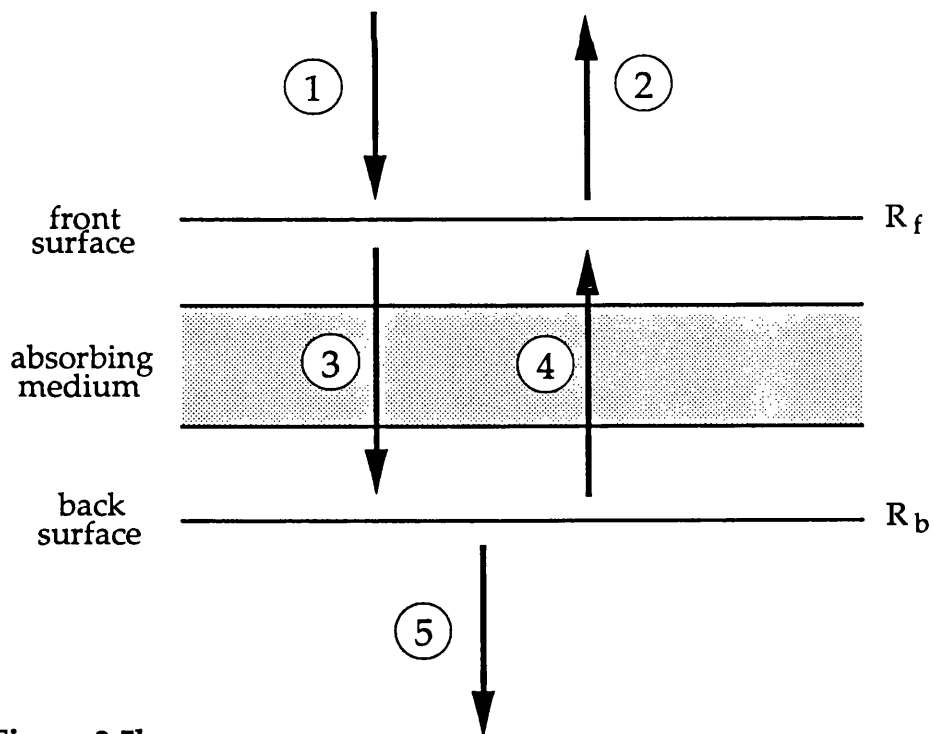


Figure 3.5b

Figure 3.5a Circulating waves in a Fabry-Perot cavity.
 b The five wave model of the Fabry-Perot cavity.

3.3.1 Five wave model of the Fabry-Perot cavity

If all the downward waves in each of the three regions of figure 3.5a, above, inside and below the cavity, are summed together taking phase into account, and so are the upward going waves, we arrive at the situation of figure 3.5b. There are two waves above the cavity, of which the downward going wave (1) is just the original incident wave on the cavity; there are two waves (3) and (4) within the cavity, which undergo attenuation during their passage through the absorbing medium; and a single wave (5) emerges from the cavity.

Treating all the individual waves together as one is valid provided we remember the consequences of interference: when one of our new waves arrives at an interface the proportion of its energy splitting into reflected and transmitted waves depends on the phase of the wave relative to all other waves arriving at the interface. Dealing with the situation in figure 3.5b is greatly simplified when we realise that **interference occurs only at the front face of the cavity**. The proportion of the energy in wave (3) going into waves (4) and (5) is constant and determined only by the back surface reflectivity R_b . How waves (1) and (4) divide into (3) and (2) can be written down given the phase angle ϕ between (1) and (4) and also R_f . To convince any sceptics the standard Fabry-Perot equations are derived using the five wave model in Appendix A.

The fact that there is no interference at the back surface means that there is a simple relationship between cavity absorption A and transmission T which is independent of the phase angle ϕ and so of all Fabry-Perot effects.

3.3.2 Relationship between absorption and transmission

Figure 3.6a is a more detailed version of figure 3.5b where waves (3) and (4) have each been separated into two parts a and b, before and after passing through the absorbing medium. The amplitudes of (3a), (4a) and (5) are X , Y and Z respectively, and those of the waves (3b) and (4b) are X' and Y' . The relationships of the waves before and after passing through the absorber are

$$\begin{aligned} X' &= X \exp\left(-\frac{1}{2}\alpha d\right) \\ Y' &= Y \exp\left(-\frac{1}{2}\alpha d\right) \end{aligned}$$

Wave (4b) is split at the back interface according to

$$\begin{aligned} Y &= r_b X' \\ Z &= \sqrt{1-r_b^2} X' \end{aligned}$$

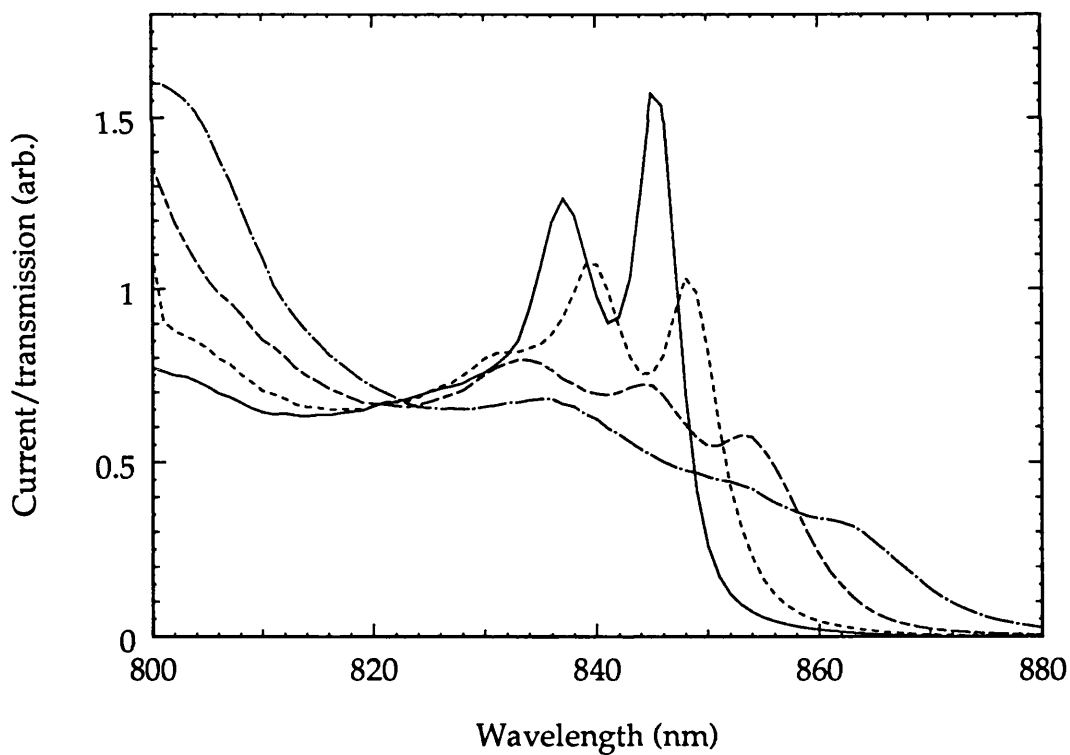
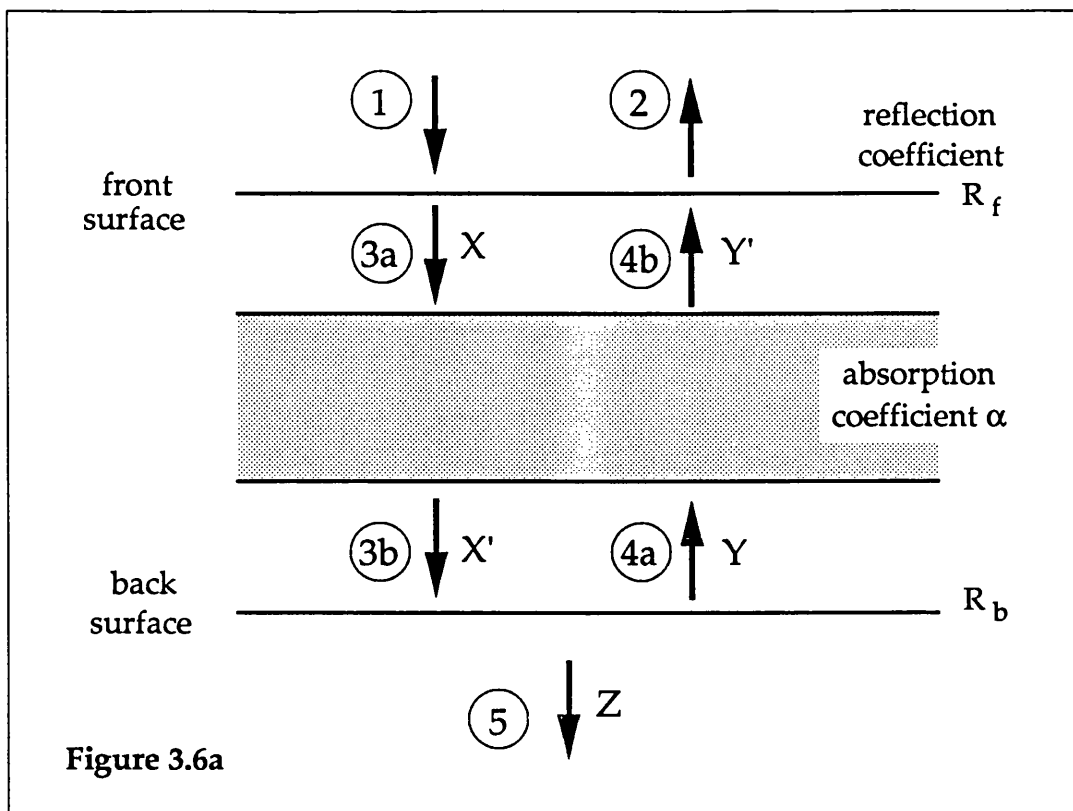


Figure 3.6a Details of the five wave model of the Fabry-Perot cavity.
b Curves obtained by dividing photocurrent (figure 3.4a) by transmission (figure 3.4b). Note that all the Fabry-Perot effects have been eliminated.

where $r_b = \sqrt{R_b}$ is the amplitude reflection coefficient. We have the relationship between X, Y and Z.

$$Y = r_b \exp\left(-\frac{1}{2}\alpha d\right)X$$

$$Z = \sqrt{1-r_b^2} \exp\left(-\frac{1}{2}\alpha d\right)X$$

The absorption in the cavity is

$$A = (X^2 + Y^2)(1 - \exp(-\alpha d))$$

$$A = X^2(R_b \exp(-\alpha d) + 1)(1 - \exp(-\alpha d))$$

while the transmission is

$$T = Z^2$$

$$T = X^2(1 - R_b) \exp(-\alpha d)$$

So

$$\frac{A}{T} = \frac{R_b + \exp(\alpha d)}{1 - R_b} (1 - \exp(-\alpha d)) \quad (3.2)$$

which is independent of ϕ . The absorption in the cavity is proportional to $(1 - \exp(-\alpha d))$, and the function in (3.2) always increases with increasing absorption, being roughly proportional to A at low αd .

3.3.2.1 MQW absorption measurement in transverse modulators

The measurement of photocurrent, which is proportional to the absorption in the pin diode, and transmission can both fail to pick out the features of the MQW absorption spectrum because of the FP effects, as was illustrated in figure 3.4. From the treatment above it is clear that dividing the photocurrent by the transmission will give a spectrum which is independent of any Fabry-Perot effects. The function obtained from the curves of figure 3.4 is plotted out in figure 3.6b. The cavity resonances have been eliminated completely and the excitonic features which remain are crystal clear. The forbidden e1-hh2 transition can be seen at 833nm in the 10V spectrum, although it is invisible in the original spectra of figure 3.4.

The function of equation (3.2) is not directly proportional to absorption but this dividing out method enables all the absorption features to become visible. The photocurrent and transmission need not be evaluated exactly, but only to a multiplicative constant. Clearly we have a simple method of removing the Fabry-Perot resonances from MQW modulator data. The new five wave model of the FP cavity made picking out the key relationships easy; it is considerably more difficult to make the same deductions from equations (3.1).

3.3.2.2 Other applications of the five wave model

Two more possible ways of using the results of the five wave approach are presented here, but these methods have not been tested experimentally.

Some very useful data about quantum well absorption can be taken by propagating light along the layers in a waveguide. The long interaction lengths (typically $\approx 100\mu\text{m}$) make the method appropriate for measuring the weak absorption in the band tails. Both photocurrent and transmission measurements can readily be made, but there are two problems associated with interpreting the data. The waveguide is a Fabry-Perot cavity, and the rapid oscillations in transmission and current with wavelength act as noise on the measurement. The experiments of Wood et al. [27] on MQW waveguides were only accurate to 25% because of this noise contribution. There are other approaches which take the FP effects into account, such as the Walker method [28] where the waveguide is heated by a proximate heat source and watched sweeping through the FP resonances as it cools, but this is only useful for low loss waveguides and is not applicable close to the band edge. The usual method of launching light in these experiments is end fire coupling, which involves large and variable insertion losses. The loss has to be estimated in order to find $\exp(-\alpha d)$ exactly, and not to within a multiplicative constant. However, because equation (3.2) involves only the back surface reflectivity we should be able to use it directly in the waveguide case to determine αd precisely, if the photocurrent and transmitted power are measured exactly (assuming unity quantum efficiency in the pin detector). The limit to the accuracy of this approach should come from waveguide scattering loss and parasitic absorption in the cladding.

Equation (3.2) gives A/T as an expression independent of ϕ . T is easily measured in a transmission modulator and A comes from the photocurrent. In a reflection modulator such as the AFPM [26] we know A and R but not T . The purpose of growing these structures is to make working modulators, but it would still be useful to be able to pick out the MQW absorption spectrum from measured spectra. The excitons are less visible in the data from the high finesse reflection modulator cavities than in the ordinary transmission modulator data. By using the identity

$$A + T + R = 1$$

equation (3.2) becomes

$$\frac{A}{1 - A - R} = \frac{R_b + \exp(\alpha d)}{1 - R_b} (1 - \exp(-\alpha d))$$

and the MQW spectra can be deduced from the observable reflection and photocurrent spectra. Note that in this case it is necessary to evaluate A and

R exactly, and not to within a multiplicative constant.

3.4 Graded well width MQW modulators

While the contrast available from a single pass transverse MQW modulator increases with the number of wells, there is a limit to how many one can pack in. Thicker intrinsic regions means higher operating voltages, but as well as this the non-zero background doping mean that beyond a certain number of wells there is no improvement in modulator performance. A possible solution to this problem is to grade the widths of the quantum wells. Some preliminary modelling work has been done on this technique, but the conclusion as to how useful it might be is fairly open.

There are three major broadening mechanisms for the exciton peak of a multiple quantum well absorption spectrum [11]; phonon broadening - the exciton has a finite lifetime due to ionisation by phonons and its resonance is broadened by the uncertainty principle, random well width fluctuation broadening, and background doping broadening. All of these reduce the height of the exciton peak and degrade the performance of the quantum well as a modulator. In some structures the latter mechanism is the most important, and it is this broadening which can be eliminated by grading the well widths.

The pin diode structure in which MQW modulators are grown should enable a constant electric field to be applied across the quantum wells, but since the intrinsic region really has a finite background doping the electric field falls off linearly across the region. The well at one end of the structure experiences a different electric field from the one at the other end, and its exciton peak has a different location - the peak in the overall spectrum is broadened. In principle by varying the widths of the wells the exciton peaks can be made to have the same location at some convenient applied voltage. Perfect lining up would require exact prediction of the background doping and precise growth, but the background doping broadening in such a device would be zero.

Grading the well widths can only be useful in a structure where background doping is the dominant broadening mechanism without grading. It is important to establish which modulators fall into this category. The exciton energy as a function of applied voltage is shown for the two end wells of a possible modulator in figure 3.7a. At low voltages only one exciton is shifted because the structure is only partially depleted, as shown in figure 3.7b. All wells are depleted only above a certain voltage, and the difference between the two curves at this voltage is W_{\min} . Since all wells

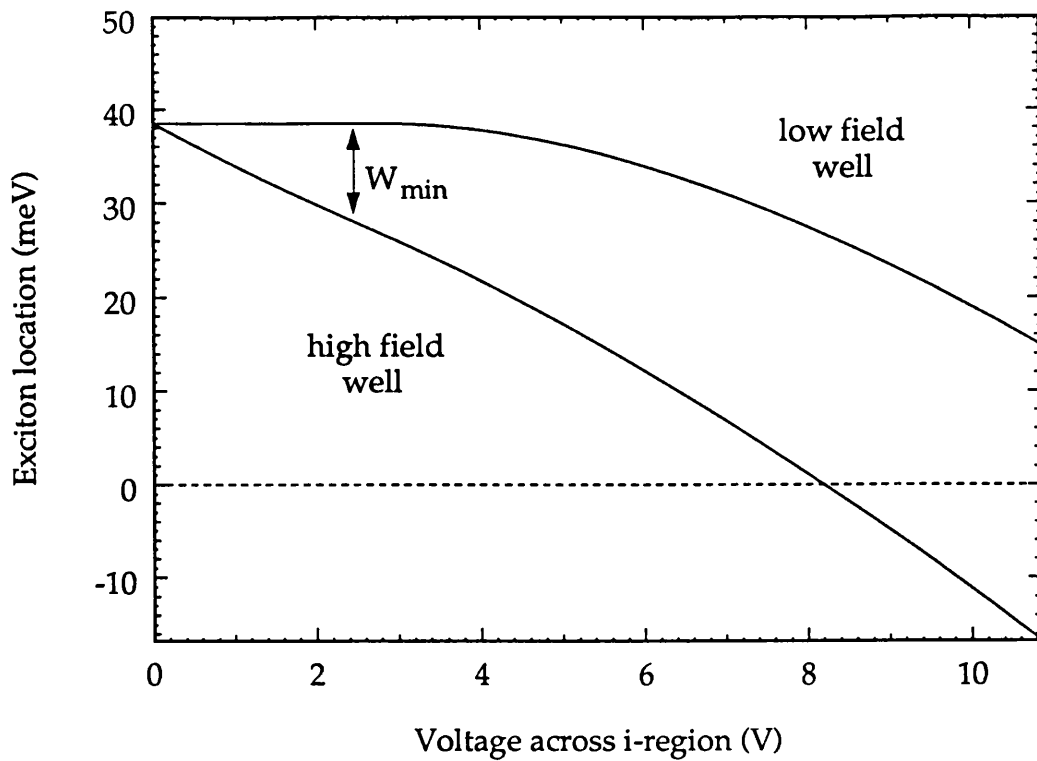


Figure 3.7a

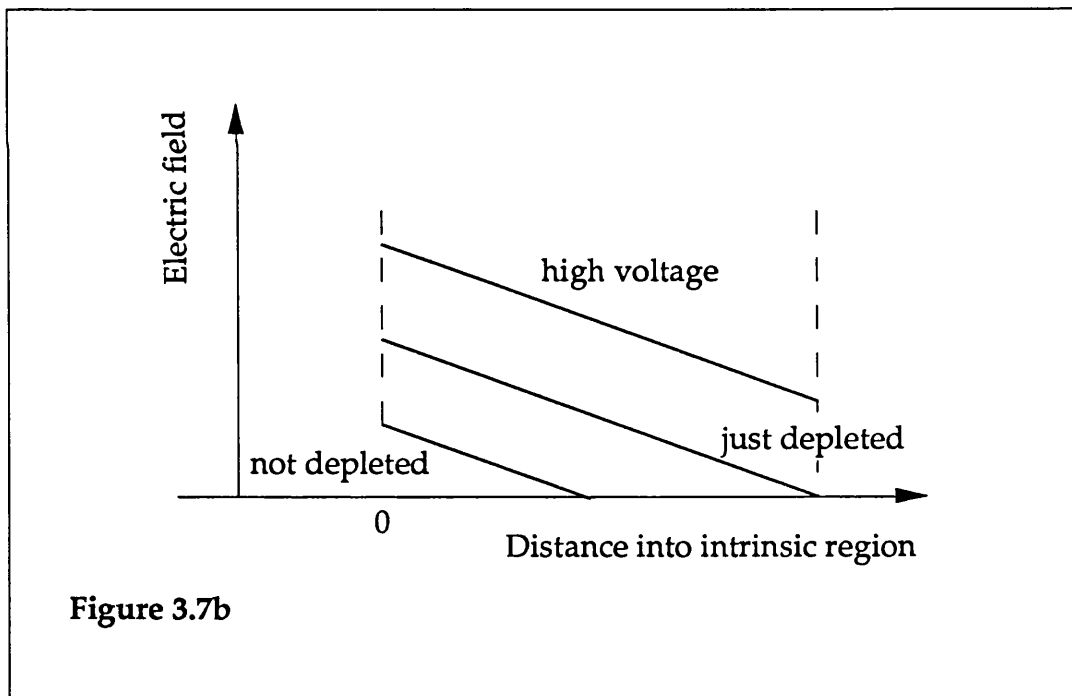


Figure 3.7b

Figure 3.7a Exciton energies (relative to GaAs band gap) of the two end wells in a $0.8\mu\text{m}$ wide intrinsic region with a background doping of $6 \times 10^{15} \text{ cm}^{-3}$. The structure becomes fully depleted at 2.6V.
b Electric field profiles within the (nominally) intrinsic region above, below and at the depletion voltage.

must be depleted for good modulation, if W_{\min} is larger than the other contributions to the exciton's breadth (notably the phonon broadening) then background doping is the dominant broadening mechanism. This is the case with, for example, 1 micron total of 100Å GaAs:AlGaAs MQW region and a background doping of $5 \times 10^{15} \text{ cm}^{-3}$, at 293K and with one monolayer random well width fluctuations [11], which is about the best specification possible for MOCVD grown material [29,30]. The background doping is less important with narrow (e.g. 60Å) wells, but these modulators need a large applied field to produce a useful shift, and if a high contrast is required (i.e. a large number of wells) the operating voltages become prohibitively high. The various long wavelength systems seem to have lower absorption strengths than their GaAs relatives and require more wells to achieve the same contrast ratio, as well as suffering from higher background dopings in general anyway. The regime in which graded well width modulators might be useful then is large intrinsic regions and high background dopings, which is the case for long wavelength modulators and MOCVD grown GaAs:AlGaAs devices.

The variations in well width starting from 100Å for a modulator with 0.8 microns of quantum wells and a background doping of $6 \times 10^{15} \text{ cm}^{-3}$ have been calculated using data obtained by P. Stevens using a tunnelling resonance computation [11], and are plotted in figure 3.8a. The excitons line up when the structure is just depleted. Curves of exciton energy against applied voltage for the two end wells are shown in figure 3.8b; the curves follow one another closely up to high voltages, and so the field inhomogeneity broadening is small over a wide range of voltage. The energy difference between the two curves is evaluated explicitly in figure 3.9, as is the equivalent parameter for 100Å ungraded wells. The contribution from phonon broadening is constant at about 6meV, and the new background doping broadening is less than this up to high voltages (the diagram only goes up to 11V because this was all the data available), and is much less than the corresponding breadth without grading. The area under the exciton peak is fixed, so by reducing the overall breadth by a factor of two, as is apparently possible, gives an advantage of two in the absorption coefficient (a very rough calculation).

The curve of figure 3.8a shows a smooth well width function across the structure. Obviously it is impossible to grow sub-atomic variations in width between adjacent layers but it should be possible to follow the general trend. The broadening of the exciton peak due to monolayer variations in the well widths has already been investigated by Dr. Stevens and was found to be

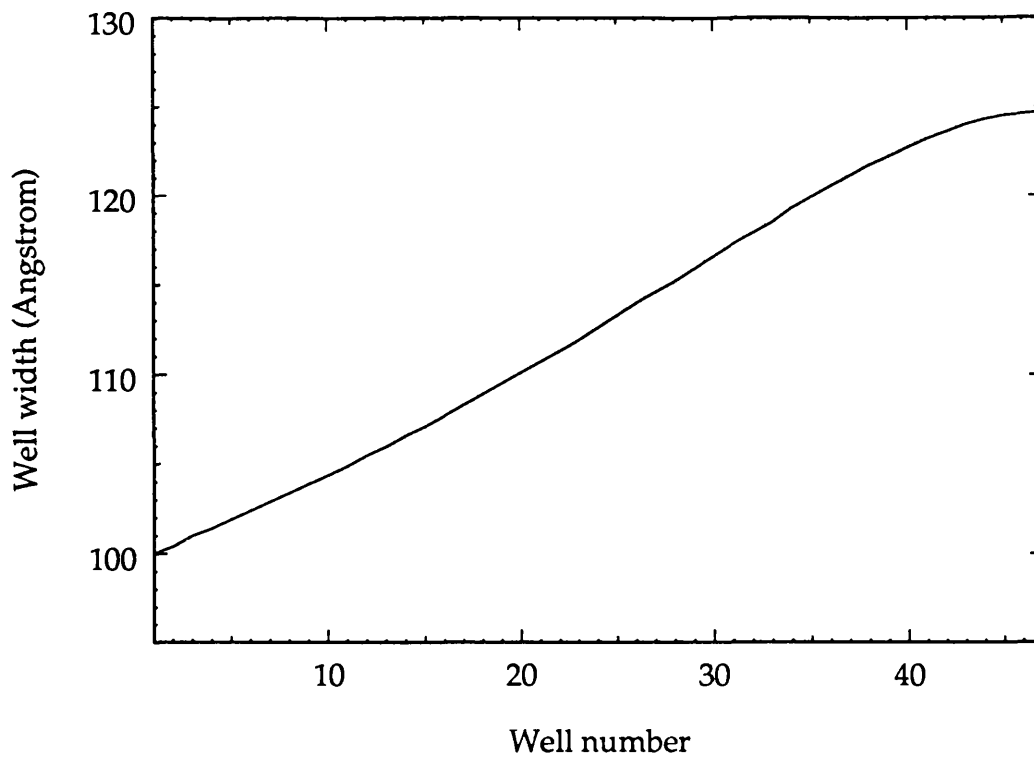


Figure 3.8a

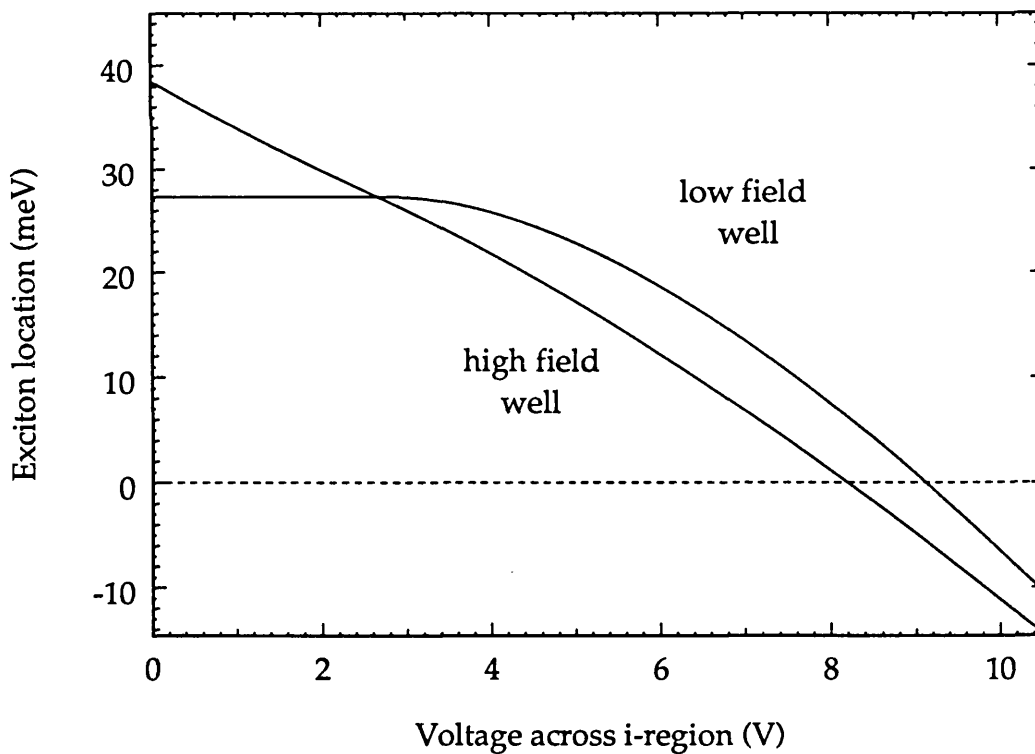


Figure 3.8b

Figure 3.8a Variation in well width required for grading the structure described in the text.

b Heavy hole exciton energies of the two end wells in a graded structure as a function of applied voltage.

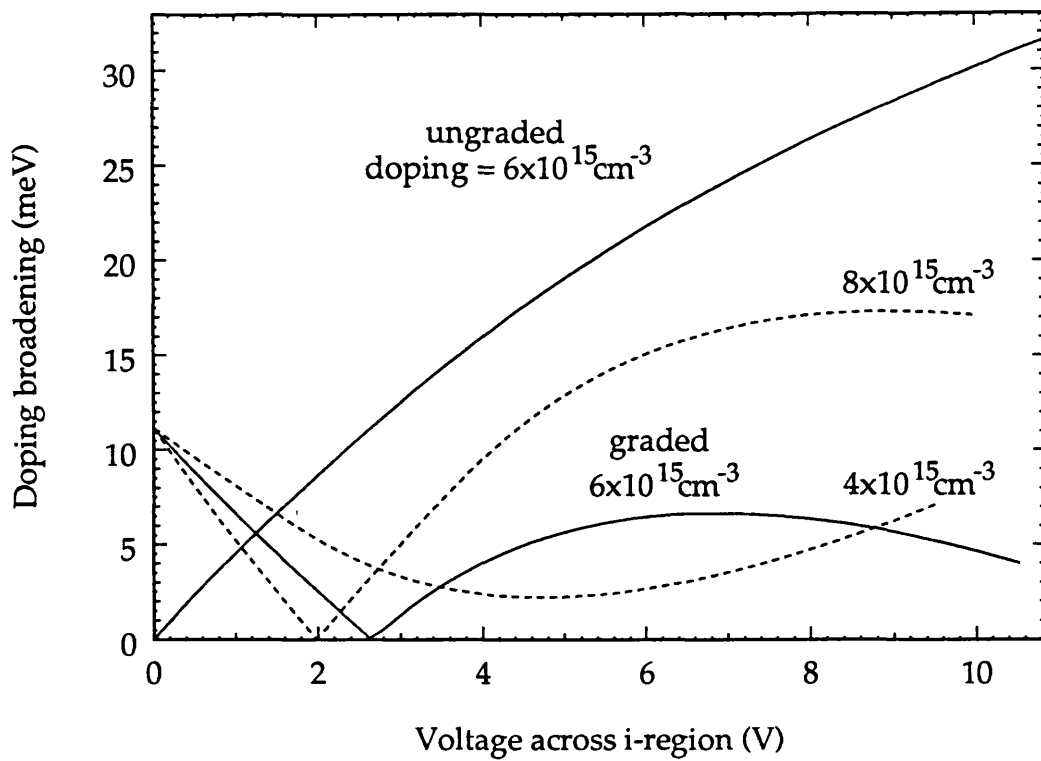


Figure 3.9a

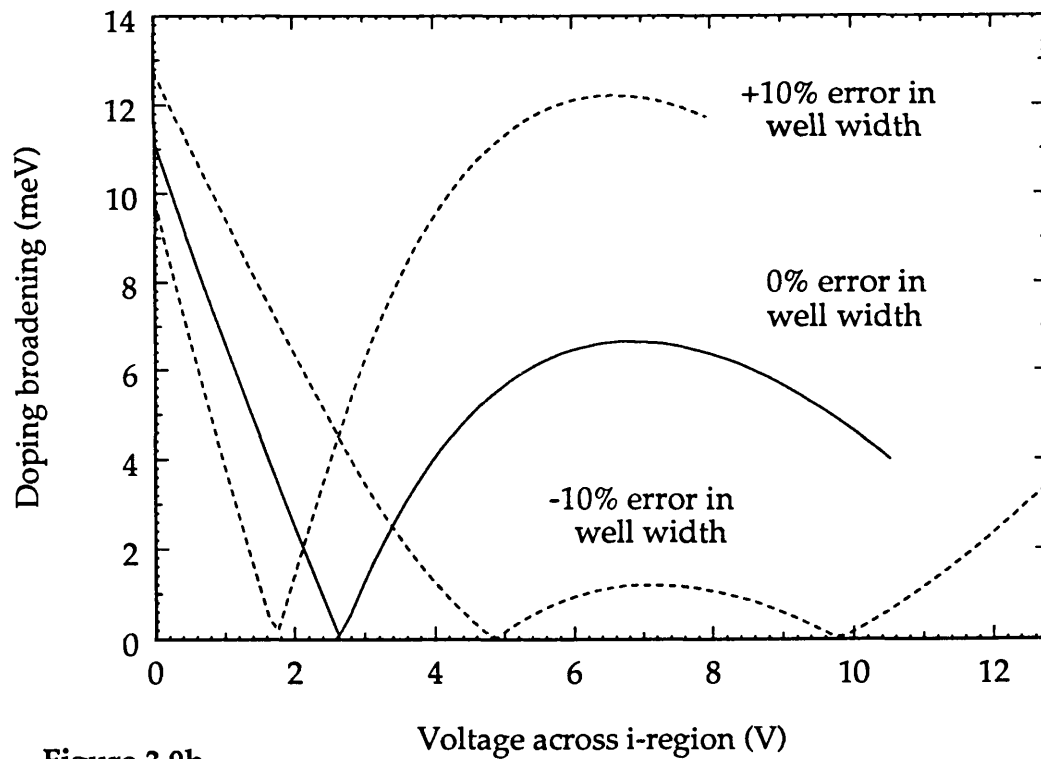


Figure 3.9b

Figure 3.9a The contribution to the width of the exciton peak from background doping broadening, for uniform and graded wells (solid lines). Also curves for graded wells when the background doping is not as predicted (dashed lines).

b Results of an overall error in well width.

less important than the other two broadening mechanisms, so one would expect the step changes in well width of a real graded structure to be relatively unimportant. An accurate prediction of the doping is very difficult, as is growing the wells to their nominal widths. The field inhomogeneity breadths which result if the background dopings are actually 4×10^{15} and $8 \times 10^{15} \text{ cm}^{-3}$ are also shown in figure 3.9a, and the results of an overall 10% error in the well thicknesses are shown in figure 3.9b. The broadening contributions in all these cases are still significantly less than the exciton breadth for the ungraded structure, and hence there is an advantage to be gained in grading the well widths even accounting for the tolerances in growth. To obtain the greatest advantage out of grading the wells accurate growth and in particular accurate prediction of the background doping are necessary.

Modelling the breadth of the exciton peak associated with the finite background doping has shown that there may be some advantage in grading the well widths. The next steps should be to model absorption spectra, and then to try to grow a graded structure. There have been some developments since the work presented here was done. Firstly, it is apparent that background doping levels below 10^{15} cm^{-3} can be achieved in GaAs:AlGaAs using MBE, and so graded wells are not likely to be needed in this system. Also it has been shown that very high contrasts can be obtained in reflection from a Fabry-Perot enhanced modulator [14] as was mentioned above. This development does not entirely displace the graded well width option. For example, long wavelength modulators on transparent substrates are good as transmission modulators, and in an application such as a readily inserted modulator for optical fibers working in reflection is not feasible. On the negative side for graded wells it should be pointed out that there are other parameters which can be varied during growth to offset the effect of background doping, such as the exact composition and so the energy gap of the well material [30]. It is not clear yet which is the best alternative.

3.5 Conclusions

The principle of operation of the MQW Stark effect modulator has been described. The effects involved are well understood and the devices are readily made.

In a real modulator Fabry-Perot effects can make it difficult to extract information about the quantum well absorption spectrum. A new model of the FP cavity has been introduced which is equivalent to the standard model but which makes it easier to understand the interference effects. In particular

some simple relationships between measurable parameters have been picked out using the new five wave model, which make it possible to extract absorption data which are untainted by FP noise in a variety of experimental configurations. The method has been clearly demonstrated for a transverse transmission modulator.

It has been noted that the finite doping in the nominally intrinsic region of a pin MQW modulator can limit the contrast ratio possible. A way to avoid the effects of the background doping by grading the well widths in a multiple well sample has been introduced. Some preliminary modelling suggests that the approach may be valuable, though there are other ways to obtain a high contrast. The area where graded well width modulators might prove useful is in transmission modulators in one of the long wavelength material systems.

3.6 References

- [1] R. Dingle, W. Wiegmann & C. H. Henry, "Quantum states of confined carriers in very thin $\text{Al}_x\text{Ga}_{1-x}\text{As-GaAs-Al}_x\text{Ga}_{1-x}\text{As}$ heterostructures," *Phys. Rev. Lett.*, vol. 33, pp. 827-830, 1974.
- [2] R. D. Dupuis, P. D. Dapkus, N. Holonyak Jr., E. A. Rezek & R. Chin, "Room-temperature laser operation of quantum-well $\text{Ga}_{(1-x)}\text{Al}_x\text{As-GaAs}$ laser diodes grown by metalorganic chemical vapour deposition," *Appl. Phys. Lett.*, vol. 32, pp. 665-667, 1978.
- [3] T. Ishibashi, S. Tarucha & H. Okamoto, "Excitation associated optical absorption spectra of AlAs/GaAs superlattices at 300K," *Int. Symp. GaAs and Related Compounds*, Inst. Phys. Conf. Series No. 63, pp. 587-588, 1981.
- [4] E. E. Mendez, G. Bastard, L. L. Chang & L. Esaki, "Effect of an electric field on the luminescence of GaAs quantum wells," *Phys. Rev. B*, vol. 26, pp. 7101-7104, 1982.
- [5] D. S. Chemla, T. C. Damen, D. A. B. Miller, A. C. Gossard & W. Wiegmann, "Electroabsorption by Stark effect on room-temperature excitons in GaAs/GaAlAs multiple quantum well structures," *Appl. Phys. Lett.*, vol. 42, pp. 864-866, 1983.
- [6] T. H. Wood, C. A. Burrus, D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard & W. Wiegmann, "High-speed optical modulation with GaAs/AlGaAs quantum wells in a *p-i-n* diode structure," *Appl. Phys. Lett.*, vol. 44, pp. 16-18, 1984.
- [7] T. H. Wood, C. A. Burrus, R. S. Tucker, J. S. Weiner, D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard & W. Wiegmann, "100ps waveguide multiple quantum well (MQW) optical modulator with 10:1 on/off ratio," *Electron. Lett.*, vol. 21, pp. 693-694, 1985.
- [8] T. H. Wood, C. A. Burrus, D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard & W. Wiegmann, "131ps optical modulation in

- semiconductor multiple quantum wells (MQW's)," IEEE J. Quantum Electron., vol. QE-21, pp. 117-118, 1985.
- [9] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood & C. A. Burrus, "Electric field dependence of optical absorption near the band gap of quantum-well structures," Phys. Rev. B, vol. 32, pp. 1043-1060, 1985.
- [10] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood & C. A. Burrus, "Band-edge electroabsorption in quantum well structures: the quantum confined Stark effect," Phys. Rev. Lett., vol. 53, pp. 2173-2176, 1984.
- [11] P. J. Stevens, M. Whitehead, G. Parry & K. Woodbridge, "Computer modelling of the electric-field dependent absorption spectrum of multiple quantum well material," IEEE J. Quantum Electron., vol. QE-24, pp. 2007-2016, 1988.
- [12] P. J. Stevens & G. Parry, "Limits to electroabsorption modulation in GaAs/(GaAl)As multiple quantum well diodes," J. Lightwave Technol., vol. 7, pp. 1101-1108, 1989.
- [13] G. D. Boyd, D. A. B. Miller, D. S. Chemla, S. L. McCall, A. C. Gossard & J. H. English, "Multiple quantum well reflection modulator," Appl. Phys. Lett., vol. 50, pp. 1119-1121, 1987.
- [14] M. Whitehead, A. Rivers, G. Parry, J. S. Roberts & C. Button, "Low-voltage multiple quantum well reflection modulator with on:off ratio > 100:1," Electron. Lett., vol. 25, pp. 984-985, 1989.
- [15] K. Wakita, Y. Kawamura, Y. Yoshikuni, H. Asahi & S. Uehara, "High-speed long-wavelength optical modulation in InGaAs/InAlAs multiple quantum wells," Electron. Lett., vol. 21, pp. 951-952, 1985.
- [16] K. Wakita, Y. Kawamura, Y. Yoshikuni & H. Asahi, "Long-wavelength waveguide multiple-quantum-well (MQW) optical modulator with 30:1 on/off ratio," Electron. Lett., vol. 22, pp. 907-908, 1986.
- [17] U. Koren, B. I. Miller, R. S. Tucker, G. Eisenstein, I. Bar-Joseph, D. A. B. Miller & D. S. Chemla, "High-frequency InGaAs/InP multiple-quantum-well buried-mesa electroabsorption optical modulator," Electron. Lett., vol. 23, pp. 621-622, 1987.
- [18] U. Koren, B. I. Miller, T. L. Koch, G. Eisenstein, R. S. Tucker, I. Bar-Joseph & D. S. Chemla, "Low-loss InGaAs/InP multiple quantum well optical electroabsorption waveguide modulator," Appl. Phys. Lett., vol. 51, pp. 1132-1134, 1987.
- [19] A. K. Tipping, G. Parry & P. Claxton, "Comparison of the limits in performance of multiple quantum well and Franz-Keldysh InGaAs/InP electroabsorption modulators," IEE Proc. J, vol. 136, pp. 205-208, 1989.
- [20] W. Dobbelaere, D. Huang, M. S. Unlu & H. Morkoc, "AlGaAs/GaAs multiple quantum well reflection modulator grown on Si Substrates," Appl. Phys. Lett., vol. 53, pp. 94-96, 1988.
- [21] P. Barnes, P. Zouganeli, A. Rivers, M. Whitehead, G. Parry, K. Woodbridge & C. Roberts, "GaAs/AlGaAs multiple quantum well

optical modulator using multilayer reflector stack grown on Si substrate," *Electron. Lett.*, vol. 25, pp. 995-996, 1989.

- [22] S. Schmitt-Rink, D. S. Chemla & D. A. B. Miller, "Linear and nonlinear optical properties of semiconductor quantum wells," *Advances in Physics*, vol. 38, pp. 89-188, 1989.
- [23] M. Whitehead, P. Stevens, A. Rivers, G. Parry, J. S. Roberts, P. Mistry, M. A. Pate & G. Hill, "The effects of well width on the characteristics of GaAs/AlGaAs multiple quantum well modulators," *Appl. Phys. Lett.*, vol. 53, pp. 956-958, 1988.
- [24] R. J. Simes, R. H. Yan, R. S. Geels, L. A. Coldren, J. H. English, A. C. Gossard & D. G. Lishan, "Electrically tunable Fabry-Perot mirror using multiple quantum well index modulation," *Appl. Phys. Lett.*, vol. 53, pp. 637-639, 1988.
- [25] Y. H. Lee, J. L. Jewell, S. J. Walker, C. W. Tu, J. P. Harbison & L. T. Florez, "Electrodispersive multiple quantum well modulator," *Appl. Phys. Lett.*, vol. 53, pp. 1684-1686, 1988.
- [26] M. Whitehead & G. Parry, "High contrast reflection modulator at normal incidence in an asymmetric multiple quantum well Fabry-Perot structure," *Electron. Lett.*, vol. 25, pp. 566-568, 1989.
- [27] T. H. Wood, "Direct measurement of the electric-field-dependent absorption coefficient in GaAs/AlGaAs multiple quantum wells," *Appl. Phys. Lett.*, vol. 48, pp. 1413-1415, 1986.
- [28] R. G. Walker, "Simple and accurate loss measurement technique for semiconductor optical waveguides," *Electron. Lett.*, vol. 21, pp. 581-583, 1985.
- [29] A. C. Jones, P. R. Jacobs, S. Rushworth, J. S. Roberts, C. Button, P. J. Wright, P. E. Oliver & B. Cockayne, "Growth of $\text{Al}_x\text{Ga}_{1-x}\text{As}$ by MOVPE using a new alkylaluminium precursor," *J. Cryst. Growth*, vol. 96, pp. 769-773, 1989.
- [30] J. S. Roberts, private communication, 1990.

CHAPTER FOUR

The Integrated t-SEED

After introducing the t-SEED this chapter presents the results of program to develop a monolithically integrated t-SEED device, highlighting the decisions which were made at each stage. There is also some discussion of further possibilities for the device, and a projection of its dynamic performance.

4.1 SEED development

Shortly after the first demonstrations of the multiple quantum well modulator came the SEED (self electrooptic effect device), a term which refers to some component combined with an MQW modulator to give a device with optical input and output [1]. Like the modulator, most of the work is in the GaAs:AlGaAs system. The first SEED was the resistor biased SEED (r-SEED) [2] which was a resistor electrically in series with a modulator. The combination showed optical bistability, with a hysteresis loop suitable for optical logic operations. It was noted that different resistances gave different operating powers and switching speeds, so that a wide range of device characteristics were possible simply by changing the resistor. There soon followed proposals for the diode biased SEED (d-SEED) [3], which can be used as self-linearized modulator or an incoherent-to-coherent light converter, and the SEED oscillator [4-5], where an LC circuit is linked to the modulator to give an oscillating optical output which can lock into a weak optical input signal [6].

The first integrated SEED device was a d-SEED [7] where the pin photodiode was grown on top of the pin MQW modulator. It was demonstrated with visible light illuminating the biasing diode so that it acted as a constant current source, making a bistable logic device. There are other examples of components being integrated with MQW modulators. Nichols et al. [8] have integrated a charge coupled device array with MQW material to make a spatial light modulator with CCD addressing. Tokuda et al. [9] have grown two pin MQW modulators on top of one another to act as a selective wavelength demultiplexer. All these schemes, and the t-SEED itself, involve vertical integration, that is where elements are arranged on top of one another.

The alternative approach is horizontal integration, where the MQW modulator is connected to neighbouring components via metal tracks. The simplest example is the symmetric SEED (s-SEED) [10-11] where two modulators are electrically in series, and which is used as a bistable logic element. Related to this is the multistable m-SEED [12], which has m modulators in series and can be applied to some more complex logic tasks. The f-SEED has the MQW modulator driven by a MESFET, which is formed on the top n layer of the pin diode. It has been demonstrated as a logic element with a gain of 25, with an input photodiode linked to the FET amplifier which drives the modulator [13]. The unit had external biasing resistors and was not fully integrated. In comparing horizontal to vertical integration we can say that the vertical approach is useful only for some simple device configurations. While the technology needs to be properly developed, it seems that horizontal integration can yield SEEDs with tailored and more complex functionality, and ultimately being able to link modulators with electronic circuitry would be ideal. The window of operation for vertically integrated devices like the t-SEED, then, has to be where large densely packed arrays are required.

The favoured logic element at the moment is the s-SEED. It has a pair of beams as input and output, and utilises a kind of dual rail logic. The states ON and OFF are encoded on whether one of the beams is larger or smaller than the other, and so provides an immunity to variations in power distribution when the device is implemented in an array processing system. S-SEEDs have been demonstrated in several simple optical computing systems [14-17], and have been fabricated in 32×64 arrays [18].

There has been work on the t-SEED idea. Andrews and Rees [19] grew an npn transistor on top of a pin modulator, but with the emitter of the transistor underneath the base. They observed the correct properties but with a transistor gain of about unity. This is quite usual in GaAs:AlGaAs heterojunction bipolar transistors in the emitter-down configuration. They also demonstrated integration of an MQW modulator between the base and collector of a transistor in an npi(MQW)n structure. This device was bistable and did not have independent access to the transistor and modulator parts; it had a low switching energy because of the power gain, but it had no optical gain between input and output.

4.2 Principle of operation of the t-SEED

The t-SEED was proposed by Wheatley [20], though it is covered by a patent by Miller [21]. It comprises a phototransistor electrically in series with an

MQW modulator (figure 4.1a). The modulator is illuminated with a constant power 'pump' beam at a wavelength just off the zero field band edge, and the amount of light falling on the transistor - the input - determines the power emerging from the modulator - the output.

The operation of the device follows Kirchoff's current law. At low input powers all the applied voltage falls across the transistor; the quantum efficiency of the pin diode as a detector is low, and a small current flows through the device. The output from the modulator is high and does not vary greatly with input power. With a high optical power falling on the phototransistor all the voltage falls across the modulator. The output power is low and constant. In between there is a region where some of the total voltage falls across each component, and the output from the modulator is intermediate between the two saturation values. The input-output characteristic of the device therefore has the form of figure 4.1b. The significance of the gain in the phototransistor is that only a small input beam is required to switch a larger pump beam on and off - the device has a real gain.

The current through the two components must be the same, and so the operating point can be determined by the intersection of current-voltage curves as in figure 4.2a. The thicker line is the iv curve for a modulator illuminated at a wavelength just off the exciton peaks. The initial increase in photocurrent is because of the rise in quantum efficiency of the detector, and then the slow rise is due to the increase in optical absorption associated with the Stark effect. The decrease in optical transmission follows the increase in absorption. The fine lines are standard transistor curves showing the current at different optical powers. At low input power there is saturation limit associated with high transmission, and at high powers there is a limit with low transmission, as explained in the previous paragraph. For an antireflection coated device the slope of the intermediate part of the transfer characteristic is equal to the gain of the transistor.

The inverting characteristic with gain is what is required to do logic operations. In proposing the t-SEED Wheatley examined other optical logic gates, including other SEEDs, where the input beam is modulated to become the output and which use bistability to perform logic. He concluded that bistability is not necessary for optical logic, but a transfer characteristic with two hard limits to encode the logic levels unambiguously is what is required. He also noted that some gain is needed to overcome any propagation losses between gates, and by employing a 'three terminal' device, where the output is the modulation of a constant power pump beam, a real gain is obtained,

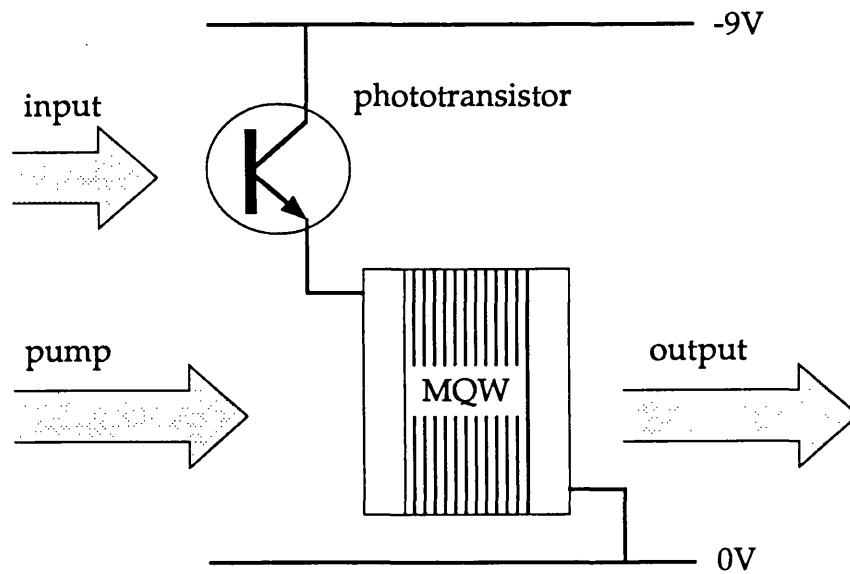


Figure 4.1a

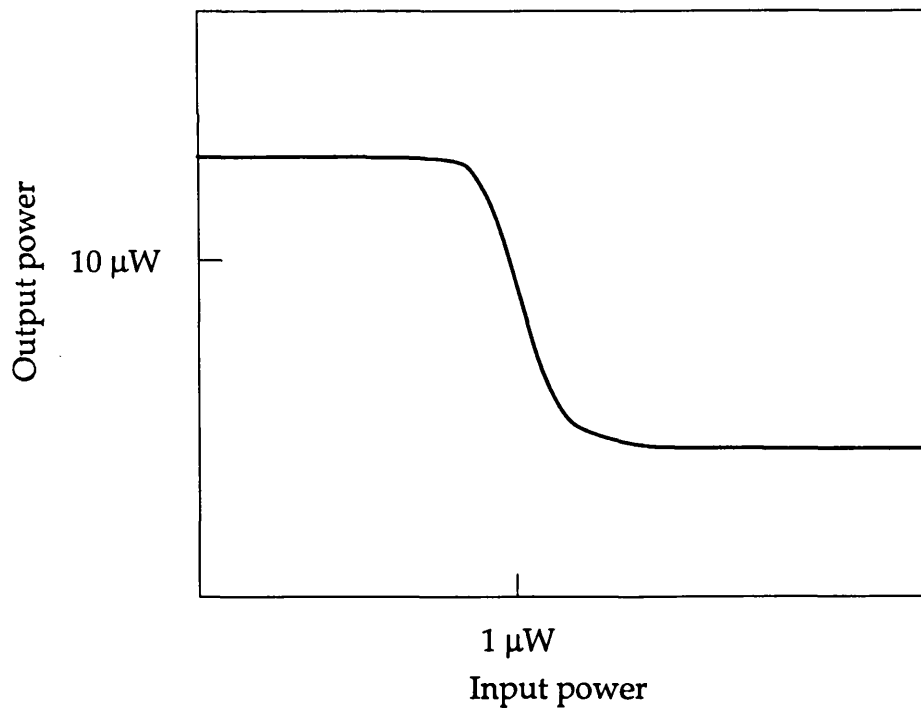


Figure 4.1b

Figure 4.1a The t-SEED.
b T-SEED transfer characteristic.

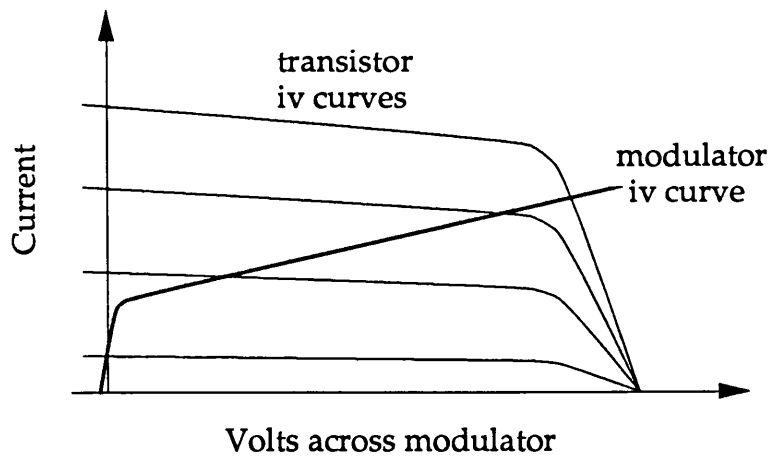


Figure 4.2a

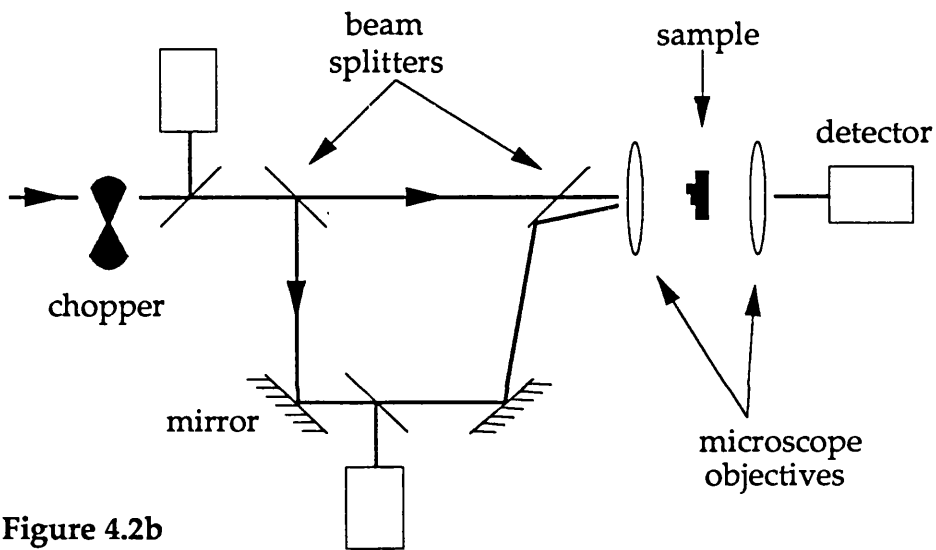


Figure 4.2b

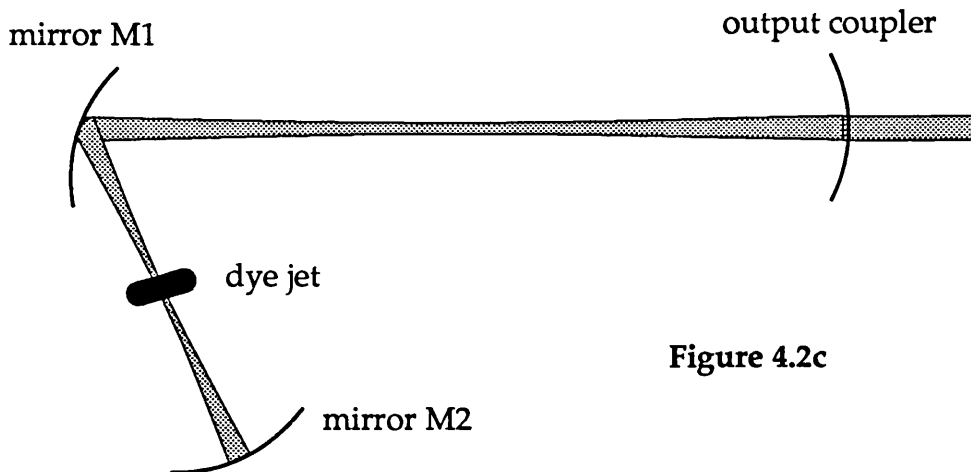


Figure 4.2c

Figure 4.2a Intersecting iv curves for transistor and modulator.
 b Experimental arrangement for examining t-SEEDs.
 c Diagram of dye laser cavity. The jet propagates into the page.

which makes building a system easier than using a differential gain as was done with the bistable devices. The arguments are laid down in full in reference 22. The t-SEED, then, has the essential features from a systems point of view of an optical logic device. With just two components it is quite simple, and so potentially suitable for use in the large densely packed arrays mentioned earlier.

The t-SEED switching characteristic has been demonstrated with a discrete device [20], that is a separate phototransistor and modulator, but has not been shown using an integrated device. The development of the integrated device is the subject of this chapter.

4.3 Material growth and processing

All material discussed here was grown by J. Roberts and C. Button by atmospheric pressure MOVPE using trimethyl gallium, trimethyl aluminium and arsene as main sources, together with silane (n-type) and dimethyl zinc (p-type) as dopant sources. Details are given in references 23-24. The processing of wafers into devices was done by G. Hill, M. Pate, A. Rivers and S. K. Lim, using standard techniques [25]. The mesas were wet etched using an ammonium hydroxide, hydrogen peroxide, water (1:1:20) mixture, and the substrate window was removed with a freon 12 plasma which selectively etched gallium arsenide.

4.4 Practical considerations in investigating devices

Before going into any device details the experimental arrangement will be discussed. A tunable dye laser (Coherent I499) with Styrl 9 dye pumped by an argon ion laser (Coherent I90) was the primary apparatus for investigating the devices. To illuminate the phototransistor and modulator portions separately, the laser beam was divided into two arms (figure 4.2b) and then combined with a small angle between them at a beamsplitter, so that the microscope objective immediately in front of the sample focussed the beams into two spots about 250 μ m apart. The optical powers in the two arms could then be controlled independently. The light was chopped at source and lock-in amplifiers were used to monitor the current in the device and the powers falling on the detectors. The wavelength control, the lock-in amplifiers and the power supply biasing the device were controlled by a computer. This made measurements easy, and also, since long term drift in the laser power was a problem for some measurements, it gave better results because the data points were gathered quickly.

The main hindrance to obtaining clean curves from the data acquisition

system was noise, which principally came from noise on the laser beam. The dye laser was a standard folded cavity laser, depicted in figure 4.2c. The dye solution was introduced into the cavity as a flat-shaped jet flowing square to the page in figure 4.2c, and was pumped by an Ar⁺ laser beam (not shown). It seems that most of the laser noise was due to instabilities, such as bubbles, in the dye jet. It was found that the noise was significantly reduced by taking care to focus the mirrors M1 and M2 properly onto the jet, since the cavity mode then sampled a smaller area of jet and so fewer noise generating events.

An early cause of erroneous results was trusting the image seen by the camera too much. The focal plane of the imaging system and the plane in which the laser spots were focussed were not the same, and so it was possible for two small spots to appear on the monitor when in reality they were not focussed tightly on the device. The proper procedure was to use the camera to position the beams in the right place, and then do the final focussing by translating the device and watching how rapidly the photocurrent or transmitted power fell to zero when the relevant beam was moved off the device.

The other instrument used was a curve tracer. This was very informative about breakdown voltages and responsivities when used on devices illuminated by the laser. Often what we were interested in about a device under test was how much voltage was falling across the transistor and how much across the modulator. A useful technique here was to take spectra at several applied biases, and to infer the potential across the modulator from the Stark shift. Attempts to put a contact on an intermediate layer failed, probably because we were trying to contact AlGaAs which is always difficult. However, the experimental techniques available were always found to be adequate to determine the behaviour of the t-SEED devices.

4.5.1 Device MV247

The first attempt at integration was simply to grow a standard heterojunction bipolar phototransistor on top of an MQW modulator. The device is illustrated in figure 4.3 and precise details of the epitaxial layers are given in Appendix B, as they are for all the layer structures described here.

After a series of experiments on MV247 devices it was decided that they were not working properly. In particular the transistor seemed to have no influence on the behaviour of the device, and all effects associated with the input beam were really due to light leaking through the transistor onto the modulator. The current through the device did not fall to zero when the

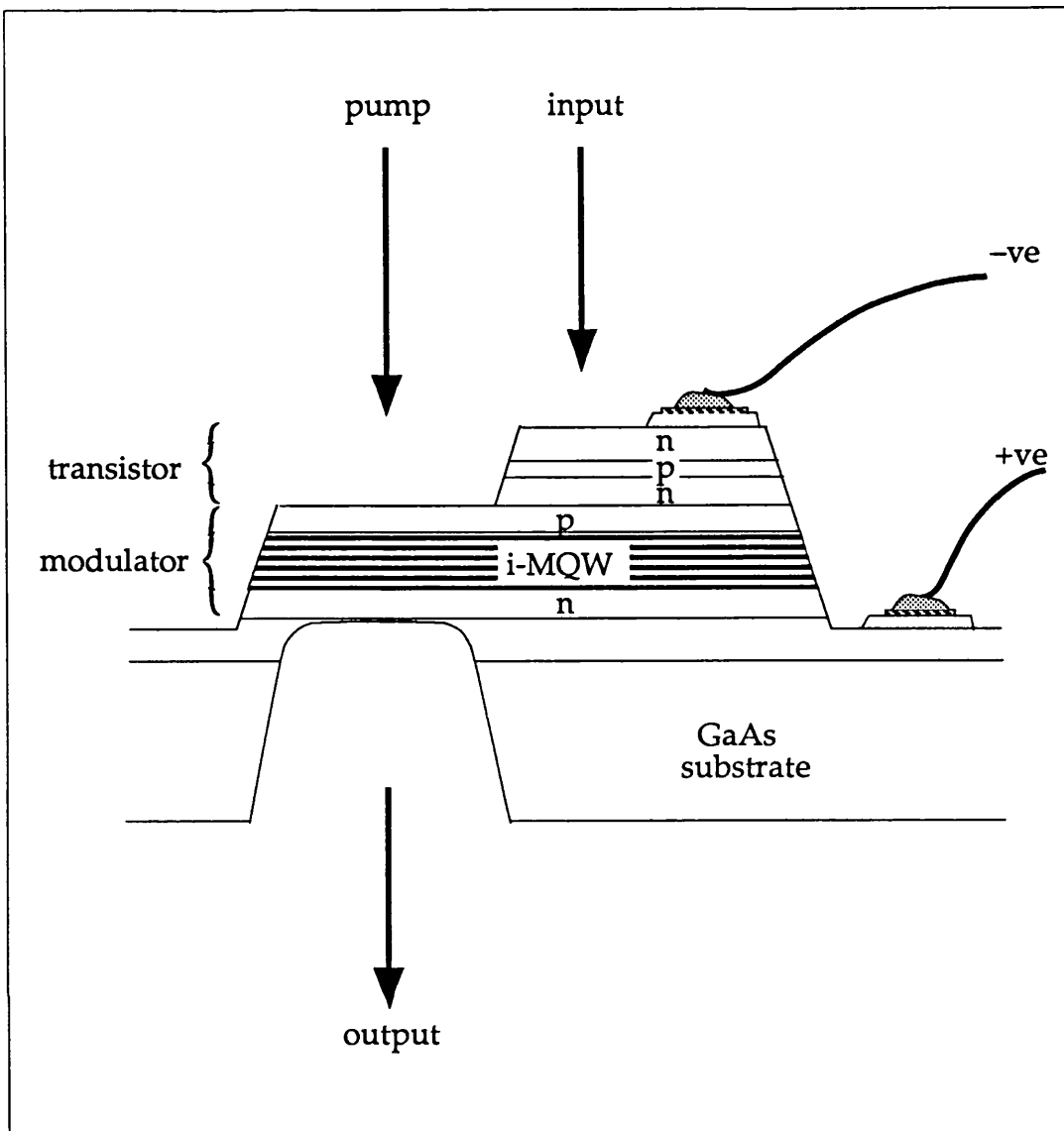


Figure 4.3 T-SEED configuration used for early devices, like MV247.

transistor was not illuminated, as it should be for two photodetectors in series. A crucial experiment was to measure a series of current spectra at different biases with only the modulator illuminated. In this situation there should be zero volts across the modulator, but as can be seen in figure 4.4a there is a considerable Stark shift, which indicates all the applied bias falling across the modulator. Several conventional transistors with base contacts were made from MV247, and all of these had a gain of ≈ 50 , as is illustrated in figure 4.4b.

An explanation was found for this behaviour in the effect of the pn junction between the two components of the device. Even though this junction is forward biased it still prevents the transistor from working properly. The band diagram one might expect for the top four layers of the structure is shown in figure 4.5a. The emitter-base junction is forward biased and as it is a heterojunction the current it passes comprises almost

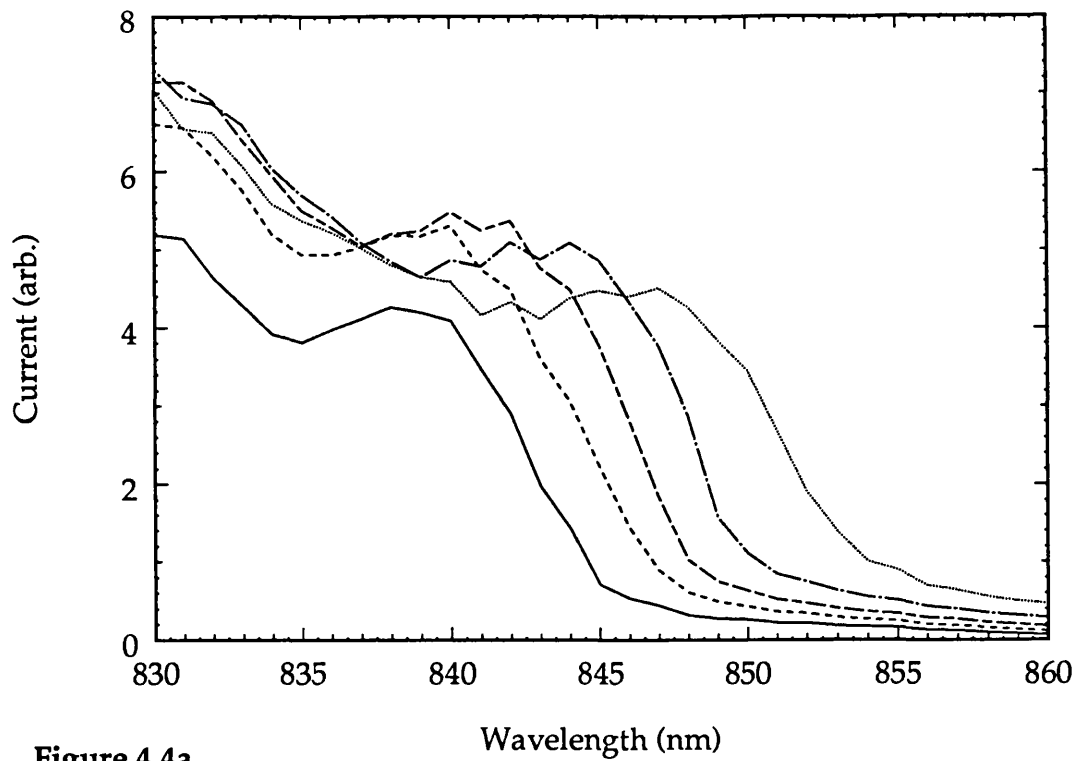


Figure 4.4a

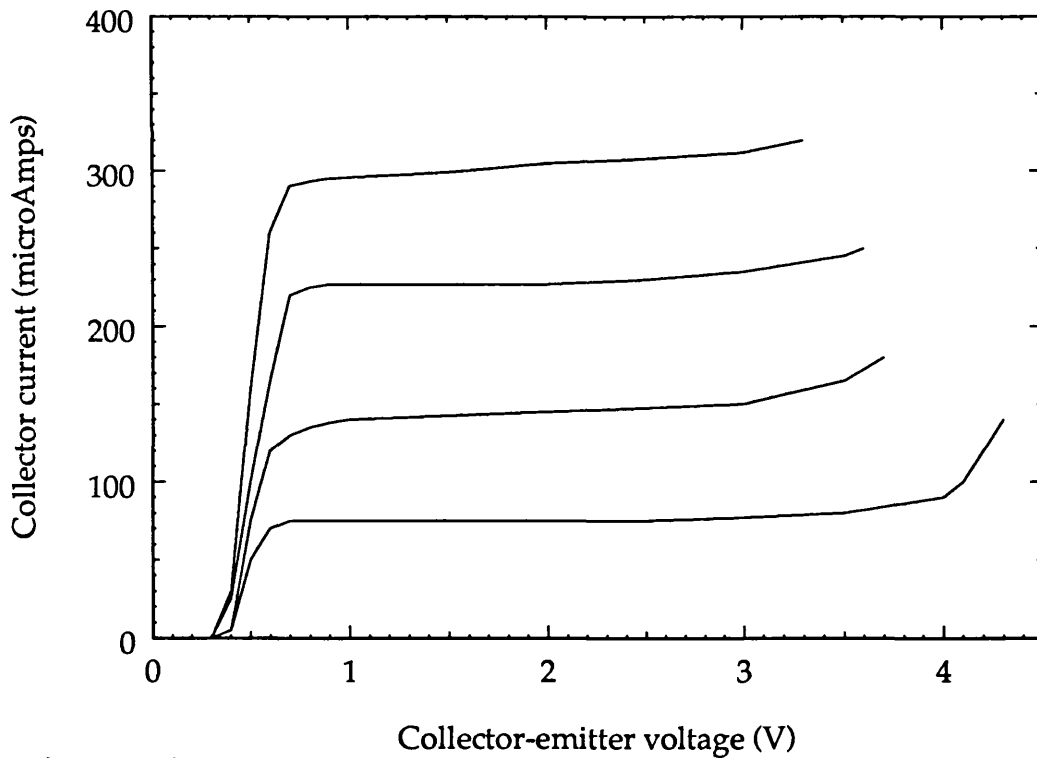


Figure 4.4b

Figure 4.4a Current spectra for MV247 at 2V (solid), 4V, 6V, 8V and 10V (dotted), taken with only the modulator window illuminated.
b Common emitter curves for the transistor part of MV247 measured with an electrical contact to the base. Base currents are (with increasing collector current) 1 μ A, 2 μ A, 3 μ A and 4 μ A, so the gain is about 75.

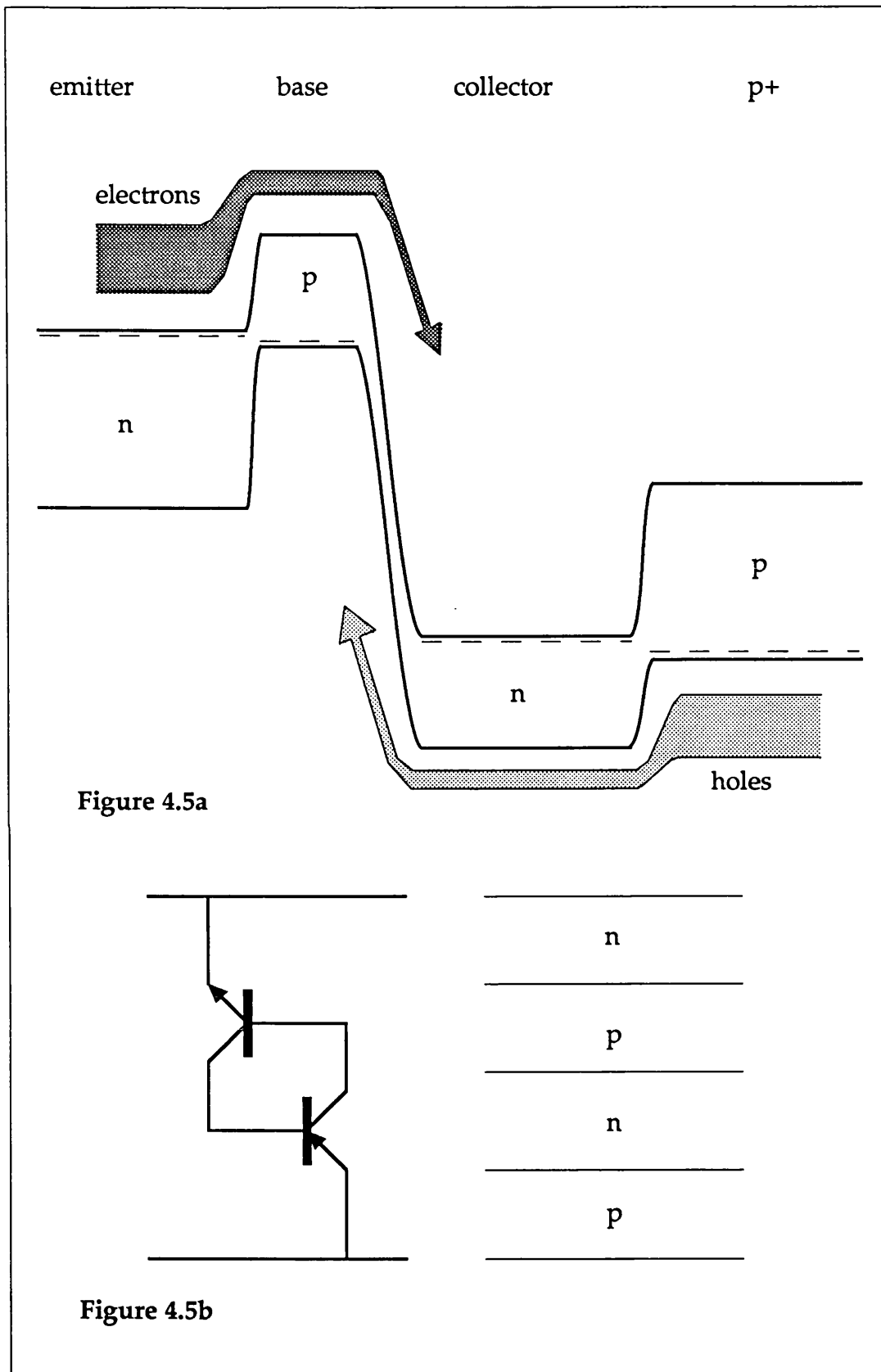


Figure 4.5a Band diagram expected for an npnp sequence of layers, as at the top of the MV247 structure.

b Two back-to-back transistors, which models the behaviour of the npnp structure.

entirely electrons injected from the emitter into the base. The base width L is much less than one diffusion length L_n , and the proportion of the injected electrons which reach the collector before recombining with majority holes in the base is $\text{sech}(L/L_n)$ [26 p.52]. In a phototransistor the holes generated by photon absorption in the base and the collector-base depletion region are stored in the base, and a large enough current must be injected across the base (that is collector current) to maintain a recombination rate in the base equal to the photogeneration rate. The ratio of collector current to effective base current is thus

$$\beta_{\text{nnp}} = \frac{\text{sech}(L/L_n)}{1 - \text{sech}(L/L_n)} \approx \frac{2L_n}{L^2} \quad (4.1)$$

which is the common-emitter current gain.

If a lead were connected to the collector via an ohmic contact the electrons would be able to leave the collector region, and the simple picture of transistor action above would be valid. However, the transistor in the integrated t-SEED is grown on a p+ layer; the depletion region at the n-GaAs p-AlGaAs junction constitutes a high potential barrier which the electrons must surmount in order for their current to flow through the modulator and into the bottom contact of the device. Electrons therefore accumulate in the collector region, reducing the depletion region widths at both ends of the collector and moving both the collector junctions into forward bias. The collector-p+ interface is a heterojunction, and when in forward bias the current it passes consists almost entirely of holes injected from the p+ into the n region. Some of these holes recombine with the majority electrons in the collector, but the remainder are swept into the base region of the HBT in a manner which is entirely the converse of npn transistor action. All three junctions of figure 4.5a will continue to become more forward biased until the middle junction (the collector-base junction) is actually in forward bias. When this happens both electrons and holes can cross the transistor layers unimpeded, which ties in directly with our experimental observations.

The behaviour of an npnp structure is well documented, since it constitutes a thyristor [26]. The four layers can be considered as two back-to-back transistors, as in figure 4.5b. The collector of the npn transistor is the base of the pnp transistor, and vice versa. There is regenerative feedback, and the arrangement sits in the low impedance state when the round trip gain is greater than unity, that is when

$$\beta_{\text{nnp}} \cdot \beta_{\text{pnp}} > 1$$

We can estimate this quantity for our device. β_{nnp} is about 100, and the other gain can be calculated from the analogue of equation (4.1) for pnp

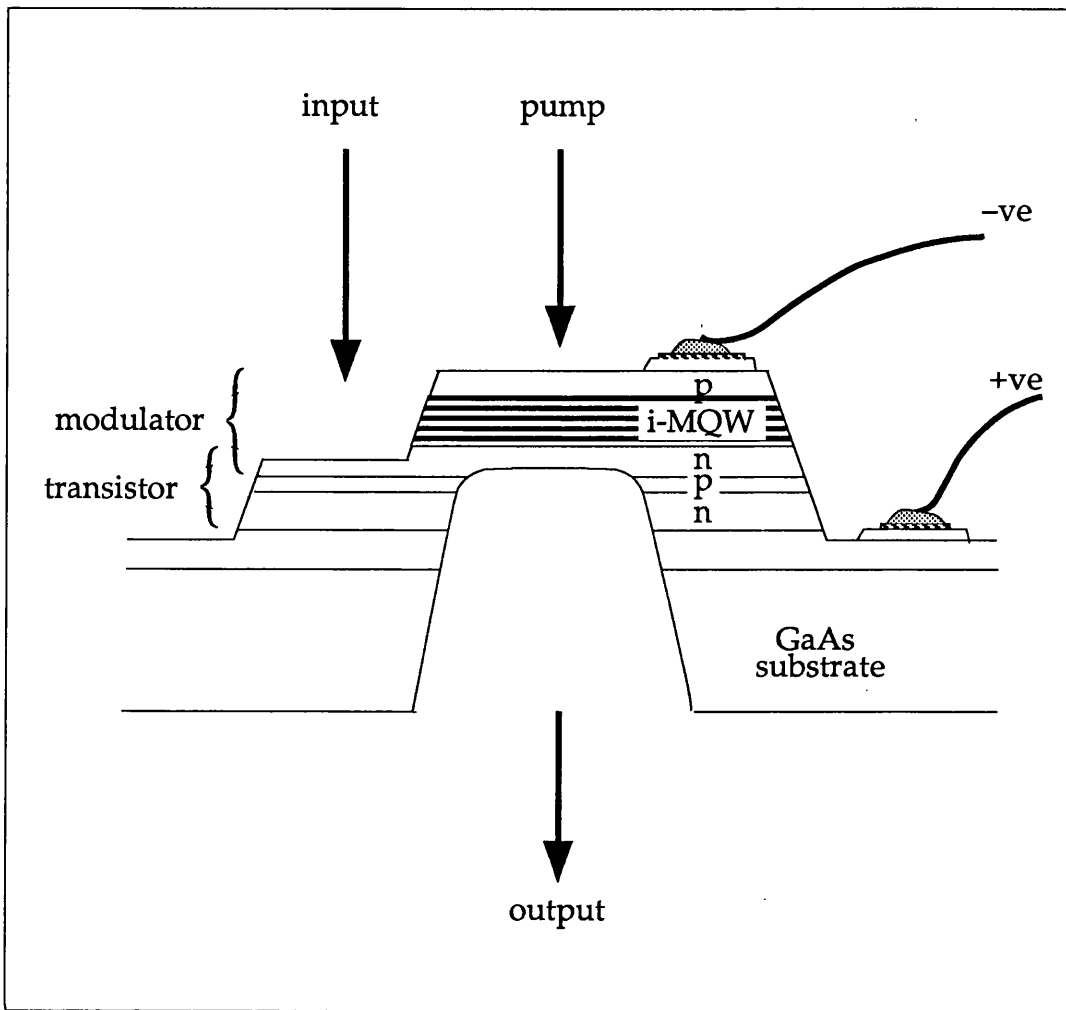


Figure 4.6 New t-SEED configuration with modulator underneath transistor, as applies to CB27 and CB245.

transistors. Using [27]

$$L_p \approx 2 \mu\text{m}$$

gives

$$\beta_{\text{pnp}} \approx 30$$

and so the product of the two transistor gains is

$$\beta_{\text{npn}} \cdot \beta_{\text{pnp}} \approx 1000$$

which is obviously much greater than unity. We therefore have a sound explanation for the poor results from MV247.

4.5.2 Device CB27

The next structure had the layers rearranged to avoid any thyristor effects. The transistor was located underneath the modulator in material CB27, and the two components shared an n-layer between them. In the GaAs:AlGaAs system it is always necessary to grow the emitter over the base. Each device had a double mesa again (figure 4.6), the main difference in processing being

that the exit window was under the middle of the structure. This could be achieved using the same mask set as before, so there were no delays in making any new masks. Note that the window has two functions in the new device - to let transmitted light leave, and also to prevent any transmitted light entering the phototransistor which would permanently switch it on.

Considerably better results were obtained from CB27. Blocking off the light falling on the transistor reduced the current to zero, so we were observing both hard limits of operation this time. There were two problems with the phototransistor, however; it had a low gain and a low breakdown voltage. The common-emitter characteristic curves are shown in figure 4.7a. The current in parentheses is the equivalent base current assuming full conversion of photons into useful electron-hole pairs. The gain is about 1.5, and the breakdown voltage is only 6V for a moderate input power. The input-output characteristic of figure 4.7b was obtained for the device. The slope of the gain region is 1.

The explanation for the low gain was thought to be that the light collection efficiency of the phototransistor was low. The observed phototransistor gain β_{photo} is the product of the electrical gain β_{npn} and the light collection efficiency η .

$$\beta_{\text{photo}} = \eta \cdot \beta_{\text{npn}}$$

η is the fraction of incident photons which are absorbed to produce holes which find their way into the base region. In CB27 the collector-base depletion region was very narrow, so that the light collecting region was the 0.2 μm wide base layer. At the operating wavelength of 860nm the absorption coefficient of gallium arsenide is 8500 cm^{-1} [28], and so with 30% reflection at the top surface it is possible that the light collection efficiency was only 10%, and the electrical gain of the transistor was as high as 15. The main solution to the problem was to widen the collector-base depletion region by reducing the doping level in the collector.

There are two usual mechanisms for transistor breakdown. Punch through occurs when the collector-base depletion region meets the emitter-base depletion region, but in our device with a highly doped base this should happen at about 28V. The other mechanism is avalanche breakdown of the collector-base junction assisted by transistor gain. From the data in reference 26 p.101 the breakdown voltage should be somewhere below 12V. The highest observed breakdown voltage was 13V, and so it seems that the avalanche mechanism is the important one. The breakdown voltage can be reduced by reducing the doping level in the collector, so we have a cure for both difficulties.

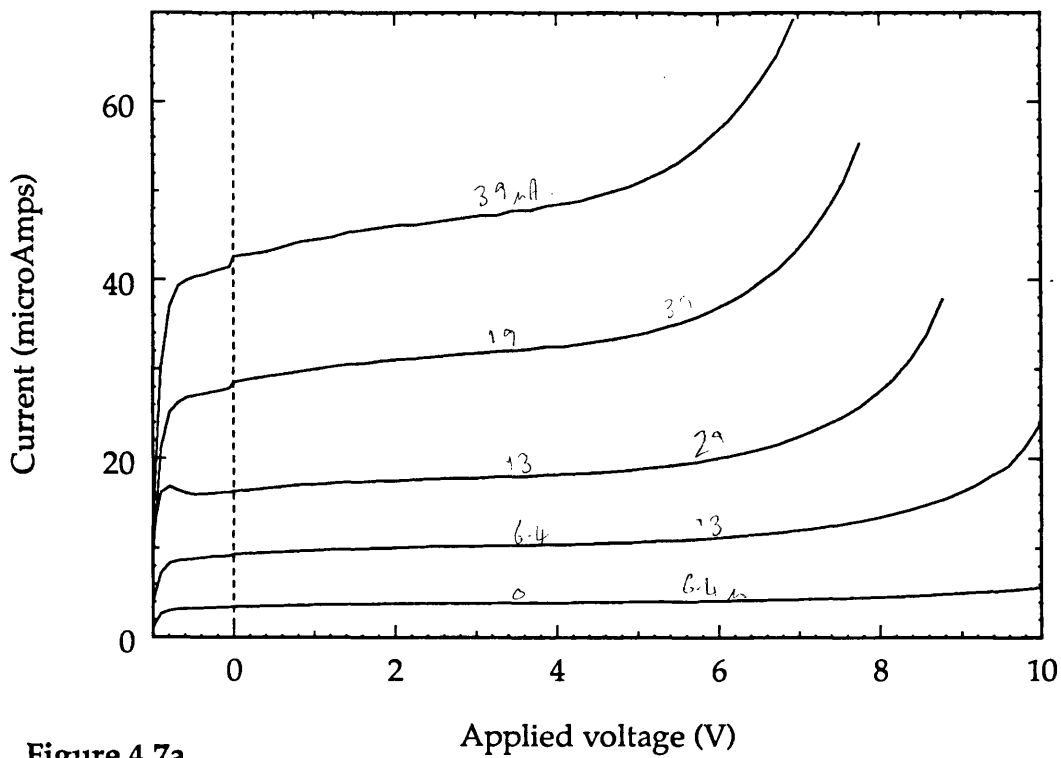


Figure 4.7a

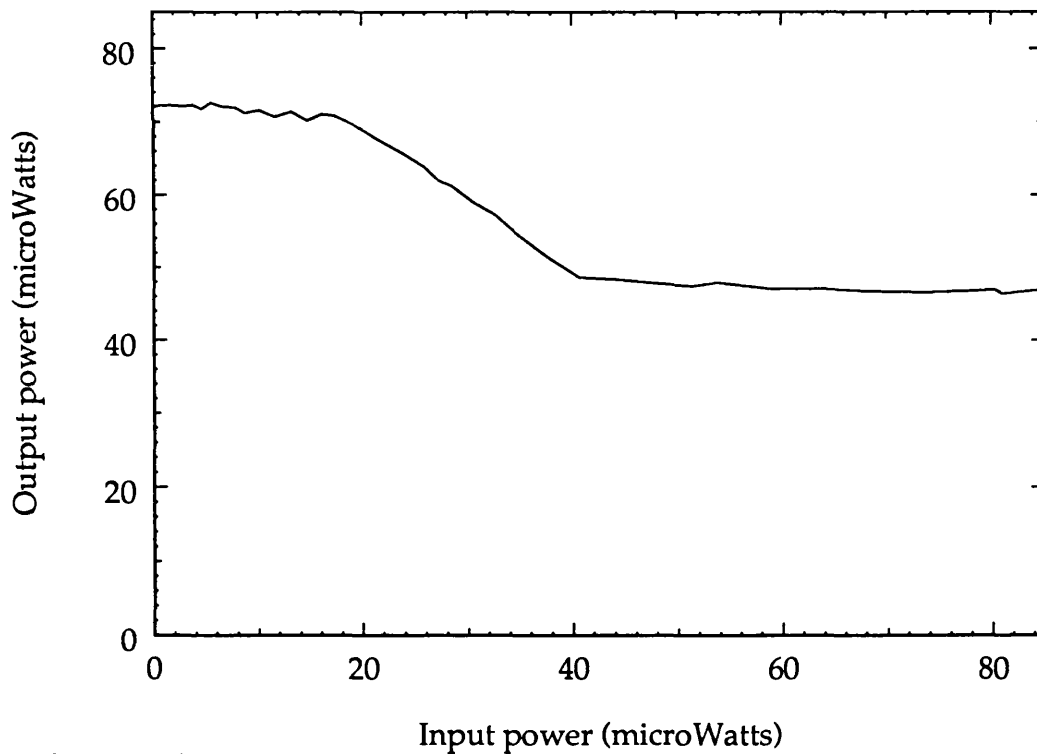


Figure 4.7b

Figure 4.7a Transistor curves for CB27, measured with the modulator strongly illuminated and with different optical powers on the phototransistor. Going upwards powers are (equivalent base current in brackets) $0\mu\text{W}$, $9.8\mu\text{W}$ ($6.4\mu\text{A}$), $20\mu\text{W}$ ($13\mu\text{A}$), $29\mu\text{W}$ ($19\mu\text{A}$) and $39\mu\text{W}$ ($26\mu\text{A}$).
 b Input-output characteristic measured for CB27.

4.5.3 Device CB245

The next growth specification differed from CB27 in two ways: the quantum wells were 85Å instead of 100Å so that the device would operate at a shorter wavelength where the GaAs absorption is stronger, and a 1µm wide collector with the lowest possible n-type doping was requested. The collector was grown on an n+ buffer layer, and a moderate voltage should be sufficient to deplete the 1µm layer fully, giving a 1.2µm wide light collecting layer at higher biases.

Devices from the first processing run on CB245 had the wrong electrical properties - the transistor behaved like a forward biased diode, as though the collector layer was not there. It was quickly realised that the outer mesa had been etched only as far as the low doped part of the collector. We were trying to make an Ohmic contact to an n- layer, which is very difficult. Whatever the mechanism for the diode-like behaviour, it was clearly due to the contact being in the wrong place.

In the next processing run the outer mesa was etched all the way to the n+ part of the collector, and there was no recurrence of the problems with the previous batch of devices. The breakdown voltage of the transistor was 20V, but unfortunately the gain was still not very large. A typical transfer characteristic is given in figure 4.8a, and it shows a fanout of 1.6. Figure 4.8b is a family of common-emitter curves which indicates a gain of about 3. The dc gain is plotted against current in figure 4.9a; it is possible to some extent to diagnose the cause of the low gain by looking at the gain as a function of current. The standard mechanism for reduction in gain at low current levels is the defect current associated with traps and recombination centres in the emitter-base junction. An ideality factor for the emitter-base diode is obtained by plotting out $\log(\text{gain})$ against $\log(\text{current})$ [26 p.142]. The slope is equal to $(1 - m^{-1})$, where m is an ideality factor. The CB245 data gave a good straight line with slope 0.18 ± 0.03 , giving $m = 1.2$, which is close to the ideal value of 1. This indicates clean injection of electrons across the emitter-base junction, and so a high defect current does not explain our low gain. At high injection levels the Kirk effect reduces the gain [26 p.145]. This effect is associated with low doped collectors, but it occurs when the number of injected carriers approaches the normal carrier density which should not be the case in our device, and in any case the Kirk effect should cause a rapid drop in gain with increasing current. The contrast ratio of the modulator changed at the highest currents, and so it is possible thermal effects were involved, though of course this does not explain the low gain.

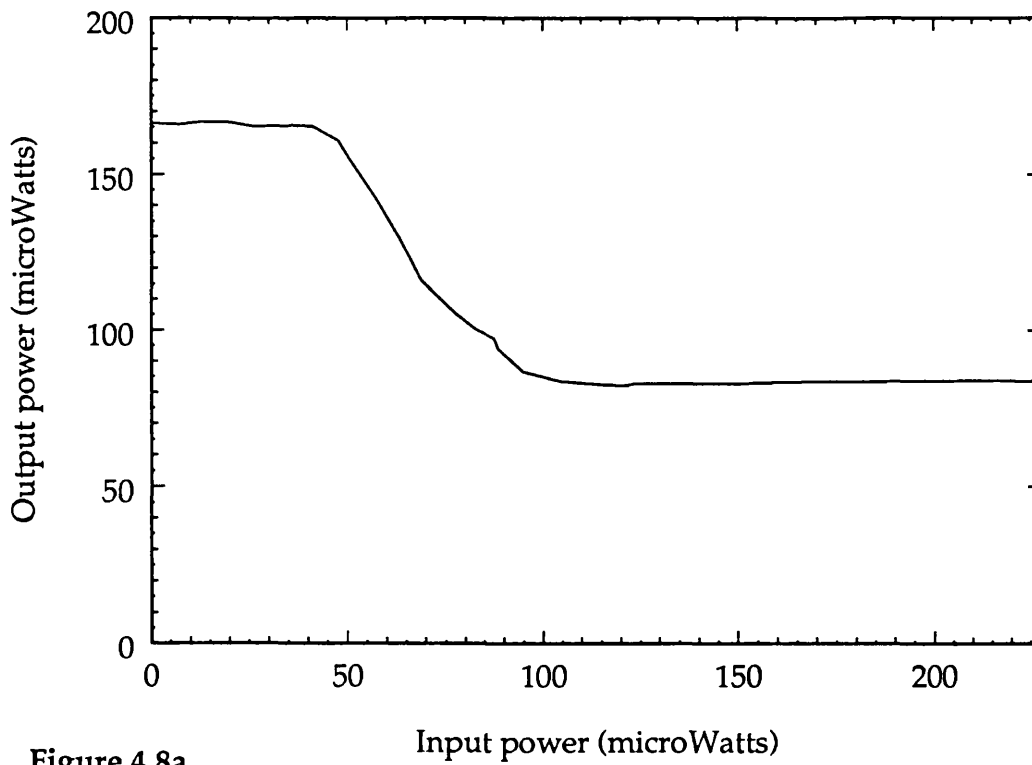


Figure 4.8a

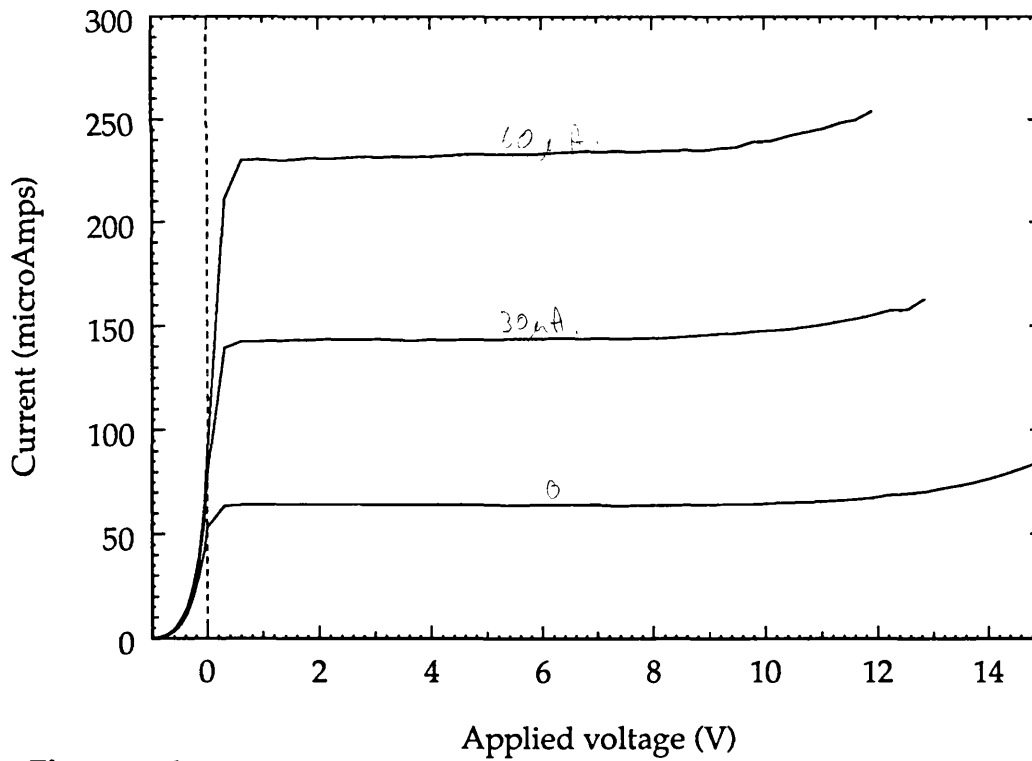


Figure 4.8b

Figure 4.8a Input-output characteristic for CB245, second processing run.
 b Transistor curves for CB245 (processing run 2) at optical powers (going upwards) of $0\mu\text{W}$, $46\mu\text{W}$ ($30\mu\text{A}$) and $92\mu\text{W}$ ($60\mu\text{A}$). The gain is about 3.

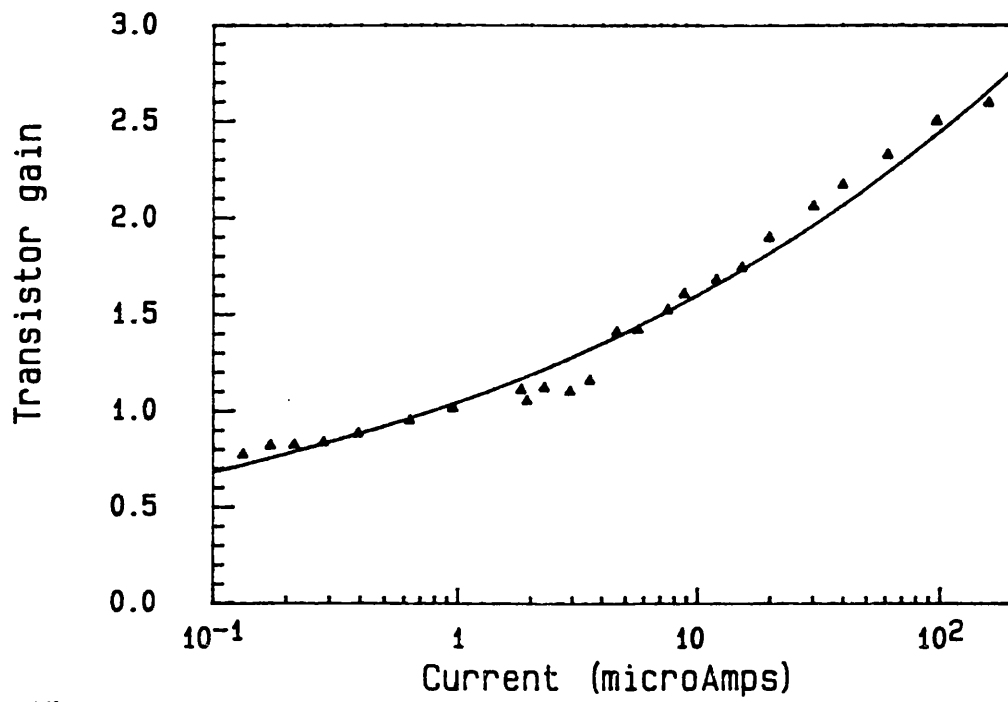


Figure 4.9a

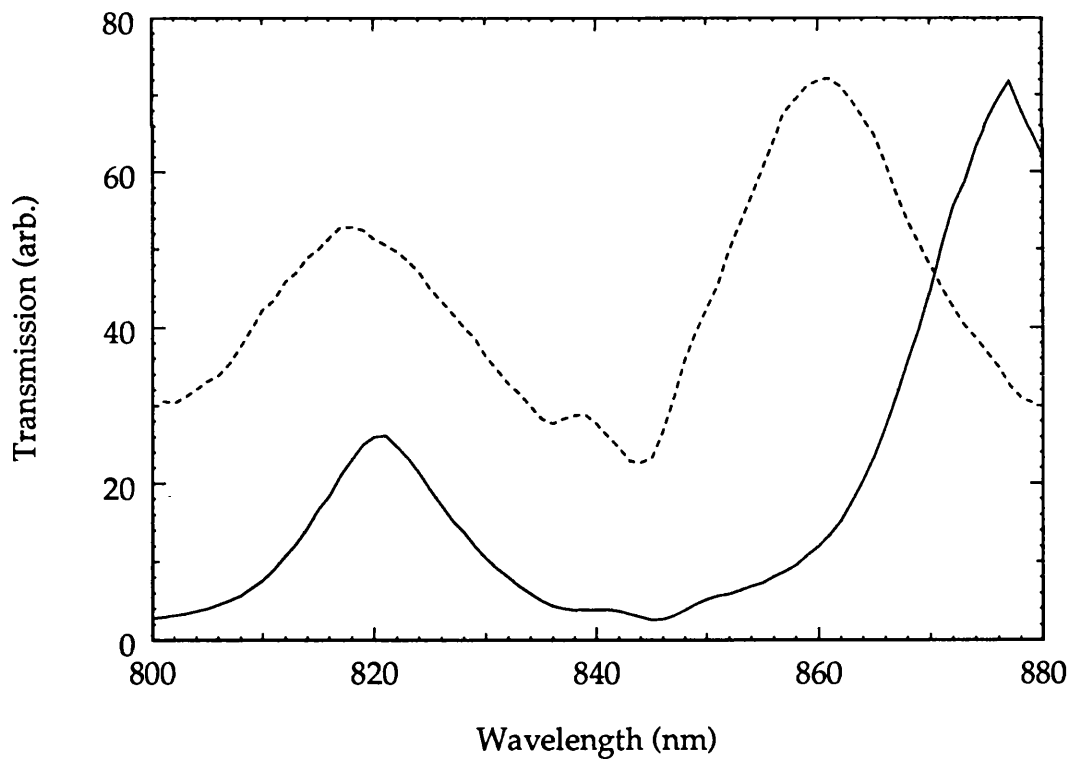


Figure 4.9b

Figure 4.9a Variation in transistor gain as a function of current for device CB245. The gain is evaluated as the difference in current with light on and off, divided by the equivalent base current. The curve is a fitted exponential.
b Transmission spectra for CB245 devices from the first (solid line) and second (dashed line) processing runs.

At this point something of an impasse was reached in that no explanation for the lack of gain could be found. The hurdle was passed when it was realised that there was surplus gallium arsenide on the back of the substrate window, and that this could radically affect the properties of the transistor. The Fabry-Perot peaks in the transmission spectrum of the modulator provide information on the thickness of the film left by the window etching process. Figure 4.9b compares the transmission of a device from the first processing run (solid line), which failed because the bottom contact was in the wrong place, with the spectrum of a device from the second processing run (dashed line). The resonances are closer together for the second device, indicating a wider cavity. The positions of the resonances are consistent with the window etching process stopping $0.7\mu\text{m}$ short of the AlGaAs emitter layer. There is usually a colour change when the plasma etch has reached an AlGaAs layer which is used as a criterion to decide when the etch has stopped, but a $0.7\mu\text{m}$ layer of GaAs is so thin that the the colour change might have been observed anyway, and the device was removed early.

It would seem that the base and $0.5\mu\text{m}$ of the low doped collector layer remained underneath the emitter. There would not be enough collector for proper transistor operation, and so the layers under the modulator would behave as a forward biased diode connected across the emitter-base junction. This diode effectively shorts out the emitter-base junction; the voltage drop across the junction is lower than for normal transistor operation, and almost all the base current, that is the primary photocurrent, can flow through the diode rather than through the transistor. The transistor gain becomes unity. When the top surface reflectivity and the less-than-unity light collection efficiency are taken into account one might expect an effective gain of about 0.5, which is indeed what was observed at low currents (figure 4.9a). At high currents the lateral voltage drop across the base layer should cause the transistor to become more turned on, and so the gain should be higher. This is what was observed, and so once again we have reasonable agreement between theory and experiment when the theory is correct.

The solution to the problem was simply to repeat the window etching, making sure that it did not stop short of the emitter. A third processing run was dogged by mechanical problems with the plasma etching machines, and only one batch of three good devices was made before the material was exhausted. These devices were left in the plasma etcher for a long period of time to ensure complete window removal. They all had a normal breakdown

voltage in excess of 20V, as is expected, but they had a reverse breakdown of only 5V, while it is usually higher than 10V. In the reverse mode the pin modulator diode and the collector-base junction are forward biased, and all the voltage falls across the reverse biased emitter-base junction. It is disappointing that this junction had a low reverse breakdown voltage because this indicates some degree of damage during the window etching process. When examined optically, all the devices had low gain. It is well known that for a high gain transistor, especially a heterojunction transistor, the emitter-base interface must be defect-free to allow clean injection of electrons from the emitter across the base. It appears that the plasma etching process had damaged the walls of the window, which is an inside-out mesa, and in particular the emitter-base junction was sufficiently damaged to kill the transistor gain.

What needs to be done next is to try different etches to see whether this affects the transistor gain, confirming the theory above, and to find an etch which does not damage the walls of the window. Perhaps most of the window could be etched with the freon plasma, and it could be finished with a selective wet etch. The flip-chip structure proposed in section 4.7 offers a solution because it requires no substrate window. Unfortunately, there is no more material to work with, and these experiments are beyond the time scale of this project.

The best transfer characteristic that has been obtained is that of figure 4.8a, with a gain of 1.6. While this is obviously not good enough for logic operations, it is still the highest real gain that has ever been obtained from any SEED configuration. The results we have indicate that some process development is what is needed to obtain useful high gain devices. It is certainly true that no serious fault in the t-SEED has been found which demands that the development program be abandoned.

4.6 Dynamic response of the integrated t-SEED

No dynamic measurements were made on any of the t-SEED devices since they did not have good static characteristics, and in any case to demonstrate really low switching times would require much smaller devices than the 250x750 μ m ones used in the investigations to date. Because the switching speed is quite important some simple theoretical predictions will be presented here to give an idea of what might be expected.

The t-SEED circuit diagram is given in figure 4.10a, together with a standard equivalent circuit [29] for evaluating the dynamic response. It is usual to consider a phototransistor as a photodiode across the collector-base

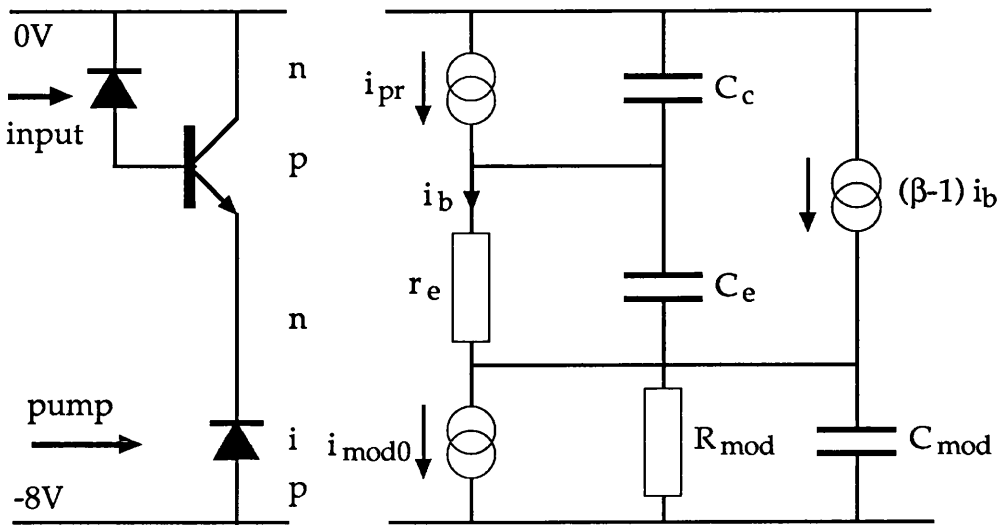


Figure 4.10a

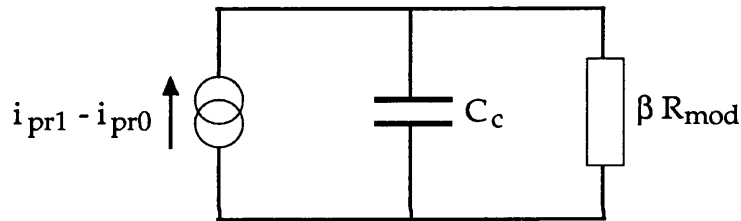


Figure 4.10b

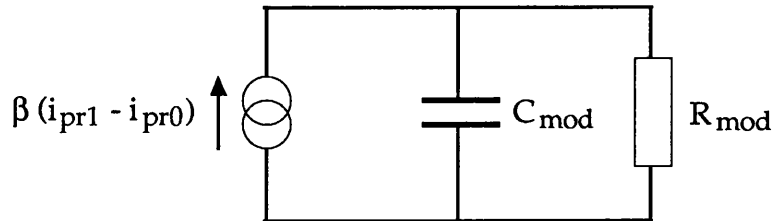


Figure 4.10c

Figure 4.10a Equivalent circuit for the t-SEED. Simplified equivalent circuits for b the input stage and c the output stage of the t-SEED.

junction of a transistor, and the equivalent circuit has the a current source generating the primary photocurrent i_{pr} . This current source has to charge the collector and emitter capacitances C_c and C_e . The fact that there is a small voltage change across the emitter-base junction during switching is represented by an emitter resistance r_e . The current flowing through this resistor is i_b , and the transistor amplifies this to a current βi_b (shown as a current source) which drives the modulator load. The iv characteristic of the modulator can be approximated to a straight line of slope R_{mod} which crosses $V=0$ at a non-zero current i_{mod} . The modulator is represented as a current source i_{mod} across a resistor R_{mod} and a capacitor C_{mod} , which is the capacitance of the pin junction.

The time constant associated with the emitter $r_e C_e$ is given by τ_r/β , where τ_r is the recombination time in the base. This time constant is very short and so the emitter junction can be neglected from the analysis. Suppose the primary photocurrent is abruptly changed from i_{pr0} to i_{pr1} . We can set

$$i_{pr0} = i_{mod0}/\beta$$

The base resistor develops less voltage across it than the collector capacitance, but it draws the same amount of current as a resistance βR_{mod} , so that the input equivalent circuit is a current source $(i_{pr1} - i_{pr0})$ driving a capacitance C_c and a resistor βR_{mod} , as shown in figure 4.10b. The amplified current then drives the modulator, as is represented by the circuit of figure 4.10c. If the current drawn by the modulator is about i_0 and the voltage across the t-SEED is V_0 , then $R_{mod} = \partial V/\partial i \approx V_0/i_0$, and the time constant of the output stage is $V_0 C_{mod}/i_0$. The expression $(i_{pr1} - i_{pr0})$ is of order i_0/β , and the time taken to charge the collector capacitance to V_0 is $V_0 C_c/\beta i_0$. The presence of β in the denominator means that the limiting time here is with the transistor, that is the time taken for the input beam to charge up the collector junction.

Using the expression for depletion capacitance on p.79 of reference 26, with a doping of 10^{16} cm^{-3} , a voltage of 5V and a device $20\mu\text{m} \times 10\mu\text{m}$, gives a mean value for C_c of 25fF. The size of the current i_0 depends on the pump power. In a real system we might want to use a single laser to illuminate an array of devices, and a 500mW laser feeding a 1000 element array (32×32) gives each device 500 μW . The responsivity of a modulator is about 20%, so i_0 would be about 100 μA . The voltage V_0 has to be large enough to give a large contrast from the modulator, say 8V. The gain required from the t-SEED might be 10. This combination of parameters yields a switching time of 20ns. This could be fast enough given that the devices are operated in a parallel architecture. A multistage amplifier driving the modulator would be faster,

but this would require the new horizontal integration technologies for it to be realisable.

4.7 Further improvements to the t-SEED

Even anticipating successful results from devices where the substrate window is wet etched, there are still improvements necessary for the devices to be useful for array processing in a real system. The thin films left behind by the window etching are very fragile. The standard way to avoid this problem is to grow a multilayer reflector stack and operate in reflection, with no substrate removal. A second reason for working in reflection is that very high contrast ratios can be obtained using Fabry-Perot enhancement in the AFPM configuration, as was discussed in chapter 3.

It is not immediately clear where to put the reflector stack in the present device (figure 4.11a), given that the rear window was needed to prevent pump light falling on the transistor. The arrangement of figure 4.11b should do the trick. On the substrate are grown first the pin diode modulator, then the transistor and then the reflector stack. The structure is inverted and mounted on some convenient material, say sapphire, and then the gallium arsenide substrate is fully removed and a double mesa is etched. We would then have independent access to the modulator and transistor portions, with the modulator being operated in reflection. This flip-chip approach with total substrate removal has not been tried at UCL, but a similar process is used at other laboratories [4,7,13]. The layer order of the modulator and transistor parts is exactly as before, and the transistor is grown in the emitter-up configuration, so that we can expect the proper electrical behaviour from the device. The lower contact can be put down on the top of the reflector stack, so that the electric current does not pass through the stack, which should make it easier to grow the reflector and may bring benefits for high speed operation. This flip-chip device proposal then retains the present phototransistor-modulator layer combination and involves development only of new processing steps.

4.8 Conclusions

The history of SEED development has been outlined, and the principle of operation of the t-SEED has been discussed.

The first attempt at growing an integrated t-SEED had the transistor on top of the modulator. This device failed because of thyristor action involving the top npnp layer combination. The next structure had the transistor underneath the modulator, and while this device did not suffer from

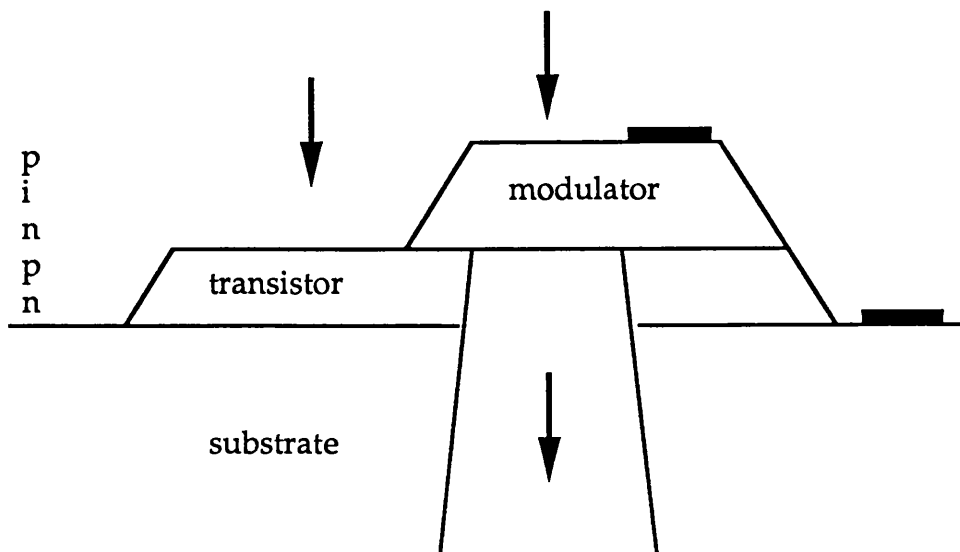


Figure 4.11a

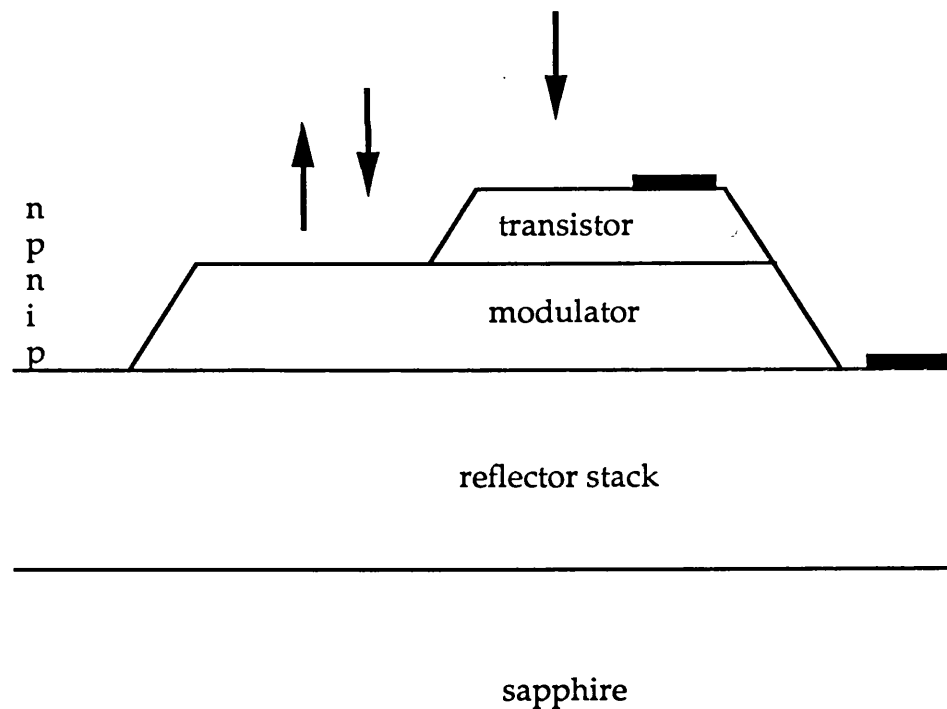


Figure 4.11b

Figure 4.11a Diagram of the present t-SEED configuration.
 b A new flip-chip configuration which allows reflection operation.

thyristor effects it had a low gain which was thought to be due to inadequate conversion of incident light on the phototransistor into primary photocurrent. The third structure dealt with this problem by having a very low doped collector, so that the collector-base depletion region would be wide enough to collect most of the light. The first devices in this material had low gain because the substrate window was not etched up to the emitter layer - it is critical to remove the collector and base layers completely. When the window was etched away fully the devices had low gain, which is thought to be due to damage of the emitter-base junction during etching. Further work on the etching process may yield good devices.

The switching time of the device has been studied theoretically; it is limited by the time taken to for the light falling on the transistor to discharge the collector capacitance through a voltage equal to the applied bias. By putting in reasonable values for the parameters involved a switching time of 20ns has been projected. A new device configuration has been proposed which involves removing all of the substrate and working in reflection. Arrays of these devices should be much more robust than the windowed devices, and high contrasts should be possible with Fabry-Perot enhancement of the MQW modulator.

4.9 References

- [1] D. A. B. Miller, "Quantum-well self-electro-optic effect devices," *Opt. Quantum Electron.*, vol. 22, S61-S98, 1990.
- [2] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood & C. A. Burrus, "Novel hybrid optically bistable switch: The quantum well self-electrooptic effect device," *Appl. Phys. Lett.*, vol. 45, pp. 13-15, 1984.
- [3] D. A. B. Miller, D. S. Chemla, T. C. Damen, T. H. Wood, C. A. Burrus, A. C. Gossard & W. Wiegmann, "Optical-level shifter and self-linearized optical modulator using a quantum-well self-electro-optic effect device," *Opt. Lett.*, vol. 9, pp. 567-569, 1984.
- [4] D. A. B. Miller, D. S. Chemla, T. C. Damen, T. H. Wood, C. A. Burrus, A. C. Gossard & W. Wiegmann, "The quantum well self-electro-optic effect device: optoelectronic bistability and oscillation, and self-linearized modulation," *IEEE J. Quantum Electron.*, vol. QE-21, pp. 1462-1476, 1985.
- [5] C. R. Giles, T. H. Wood & C. A. Burrus, "Quantum-well SEED optical oscillators," *IEEE J. Quantum Electron.*, vol. 26, pp. 512-518, 1990.
- [6] C. R. Giles, T. Li, T. H. Wood, C. A. Burrus & D. A. B. Miller, "All-optical regenerator," *Electron. Lett.*, vol. 24, pp. 848-850, 1988.
- [7] D. A. B. Miller, J. E. Henry, A. C. Gossard & J. H. English, "Integrated

- quantum well self-electro-optic effect device: 2x2 array of optically bistable switches," Appl. Phys. Lett., vol. 49, pp. 821-823, 1986.
- [8] K. B. Nichols, B. E. Burke, B. F. Aull, W. D. Goodhue, B. F. Gramstorff, C. D. Hoyt & A. Vera, "Spatial light modulators using charge-coupled-device addressing and electroabsorption effects in GaAs/AlGaAs multiple quantum wells," Appl. Phys. Lett., vol. 52, pp. 1116-1118, 1988.
- [9] Y. Tokuda, K. Kanamoto, Y. Abe & N. Tsukada, "Dual-wavelength multiple quantum well *n-i-p-i-n* photodetector using an optically bistable abrupt absorption edge," Appl. Phys. Lett., vol. 56, pp. 227-229, 1990.
- [10] A. L. Lentine, H. S. Hinton, D. A. B. Miller, J. E. Henry, J. E. Cunningham & L. M. F. Chirovsky, "Symmetric self-electro-optic effect device: Optical set-reset latch," Appl. Phys. Lett., vol. 52, pp. 1419-1421, 1988.
- [11] A. L. Lentine, H. S. Hinton, D. A. B. Miller, J. E. Henry, J. E. Cunningham & L. M. F. Chirovsky, "Symmetric self-electrooptic effect device: optical set-reset latch, differential logic gate, and differential modulator/detector" IEEE J. Quantum Electron., vol. 25, pp. 1928-1936, 1989.
- [12] A. L. Lentine, D. A. B. Miller, J. E. Henry, J. E. Cunningham & L. M. F. Chirovsky, "Multistate self-electrooptic effect devices," IEEE J. Quantum Electron., vol. 25, pp. 1921-1927, 1989
- [13] D. A. B. Miller, M. D. Feuer, T. Y. Chang, S. C. Shunk, J. E. Henry, D. J. Burrows & D. S. Chemla, "Field-effect transistor self-electrooptic effect device: integrated photodiode, quantum well modulator and transistor," IEEE Photonics Tech. Lett., vol. 1, pp. 62-64, 1989.
- [14] A. L. Lentine, D. A. B. Miller, J. E. Henry, J. E. Cunningham, L. M. F. Chirovsky & L. A. D'Asaro, "Optical logic using electrically connected quantum well PIN diode modulators and detectors," Appl. Opt., vol. 29, pp. 2153-2163, 1990.
- [15] M. E. Prise, N. C. Craft, R. E. LaMarche, M. M. Downs, S. J. Walker, L. A. D'Asaro & L. M. F. Chirovsky, "Module for optical logic circuits using symmetric self-electrooptic effect devices," Appl. Opt., vol. 29, pp. 2164-2170, 1990.
- [16] T. J. Cloonan, M. J. Herron, F. A. P. Tooley, G. W. Richards, F. B. McCormick, E. Kerbis, J. L. Brubaker & A. L. Lentine, "An all-optical implementation of a 3-D crossover switching network," IEEE Photonics Tech. Lett., vol. 2, pp. 438-440, 1990.
- [17] E. Kerbis, T. J. Cloonan & F. B. McCormick, "An all-optical realization of a 2x1 free-space switching node," IEEE Photonics Tech. Lett., vol. 2, pp. 600-602, 1990.
- [18] A. L. Lentine, F. B. McCormick, R. A. Novotny, L. M. F. Chirovsky, L. A. D'Asaro, R. F. Kopf, J. M. Kuo & G. D. Boyd, "A 2 kbit array of symmetric self-electrooptic effect devices," Photonics Tech. Lett., vol. 2, pp. 51-53, 1990.
- [19] S. R. Andrews & P. K. Rees, "Multiple quantum well bistable optical

devices with gain," IEE Colloquium on Optics in Computing, digest no. 1988/122, 17th November 1988.

- [20] P. Wheatley, P. J. Bradley, M. Whitehead, G. Parry, J. E. Midwinter, P. Mistry, M. A. Pate & J. S. Roberts, "Novel nonresonant optoelectronic logic device," *Electron. Lett.*, vol. 23, pp. 92-93, 1987.
- [21] D. A. B. Miller, U.S. Patent 4 546 244, 1985.
- [22] P. Wheatley, "Devices for optical logic," Ph.D. Thesis, University of London, 1987.
- [23] J. S. Roberts, N. J. Mason & M. Robinson, "Factors influencing doping control and abrupt metallurgical transitions during atmospheric pressure MOVPE growth of AlGaAs and GaAs," *J. Cryst. Growth*, vol. 68, pp. 422-430, 1984.
- [24] J. S. Roberts, M. A. Pate, P. Mistry, J. P. R. David, R. B. Franks, M. Whitehead & G. Parry, "MOVPE grown MQW pin diodes for electro-optic modulators and photodiodes with enhanced electron ionisation coefficient," *J. Cryst. Growth*, vol. 93, pp. 877-844, 1988.
- [25] S. M. Sze, "Semiconductor devices: physics and technology," John Wiley & Sons, 1985.
- [26] S. M. Sze, "Physics of semiconductor devices," 2nd Ed., John Wiley & Sons, 1981.
- [27] H. C. Casey, B. I. Miller & E. Pinkas, "Variation of minority-carrier diffusion length with carrier concentration in GaAs liquid-phase epitaxial layers," *J. Appl. Phys.*, vol. 44, pp. 1281-1287, 1973.
- [28] M. D. Sturge, "Optical absorption of gallium arsenide between 0.6 and 2.75 eV," *Phys. Rev.*, vol. 127, pp. 768-773, 1962.
- [29] G. J. Ritchie, "Transistor Circuit Techniques: discrete and integrated," Van Nostrand Reinhold, 1983.

CHAPTER FIVE

Systems employing t-SEEDs

Having talked about the t-SEED device and its development, the purpose of this chapter is to look at how the device fits into systems. As has already been established in chapters 1 and 2, what optics has to offer is an interconnect ability, but it is not entirely clear how to make use of the interconnect advantage. Here we will put forward two examples of processors using t-SEEDs, a neural network and an image processing machine, which can be considered as examples of what might be possible.

5.1 Requirements on t-SEED characteristic

The operation of a t-SEED has been described in chapter 4. The feature which should make it useful is the inverting input-output characteristic. We can develop some simple conditions which this characteristic must satisfy for several of the devices to operate together.

Figure 5.1a gives the essential features of a transfer characteristic. The output is on with transmission T_1 at input powers below S_0 , and off with transmission T_0 at inputs above S_1 . Suppose we wish to have the output of one t-SEED drive the input of another. An attenuated version of output PT_1 must be larger than S_1 , and that of PT_0 must be less than S_0 . This leads to the condition

$$\frac{T_1}{T_0} > \frac{S_1}{S_0} \quad (5.1)$$

A second condition is that the device output is large enough to drive an input, that is that the transistor has gain.

A bistable unit is formed from two t-SEEDs by feeding the output of one device into the input of the other, and vice versa. The result is an RS flip-flop (figure 5.1b). One of the devices is on while the other is off, and the situation is reversed by applying a triggering pulse to one of the inputs. The two transfer characteristics should intersect as in figure 5.1c, but it can readily be verified that the stable intersection points will involve the hard limits of the two curves only if equation (5.1) is satisfied, and if the devices have gain.

In practice by changing the operating wavelength of the t-SEED the switching points S_0 and S_1 can be made to be almost equal, so condition (5.1) becomes

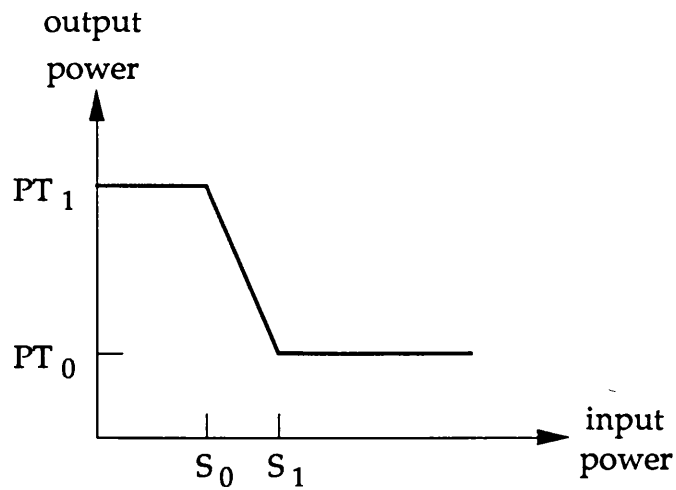


Figure 5.1a

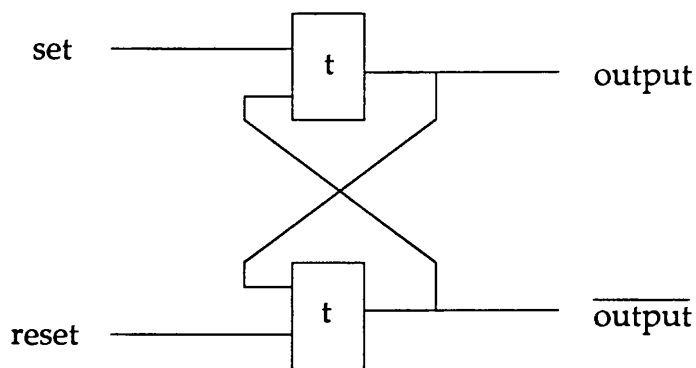


Figure 5.1b

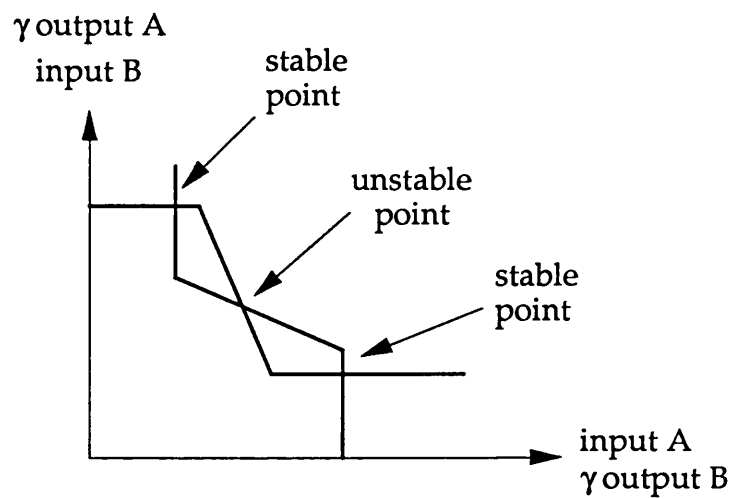


Figure 5.1c

Figure 5.1a Essential features of the t-SEED transfer characteristic.

b A set-reset flip-flop formed from two t-SEEDs. The t-SEEDs behave as NOR gates. A fraction γ of each output is fed into the other's input.

c How the characteristics of the two SEEDs intersect to make the flip-flop.

$$\frac{T_1}{T_0} > 1$$

If we want to have N outputs driving the input of one device, in a NOR or a NAND gate mode, the condition becomes

$$\frac{T_1}{T_0} > N$$

So the contrast determines the fan-in. We have two conditions for the devices to have a wide window of operation: high contrast from the modulator and high gain from the transistor. It is hoped that the work described in chapter 4 indicates that a t-SEED with these properties should be possible in the future.

5.2 t-SEED in a neural processor

An area in which it is thought optics could make a great impact is in neural processors. These machines have a structure like the brain, and are intended to solve problems which humans are better suited to than conventional computers, such as associative memory (or content addressable memory) rather than location addressed memory.

A neural network has a large number of neurons, each of which is connected to many neighbours by synapses. The neurons are bistable, either on or off, and the synapses have different weights of connection associated with them. A neuron changes state according to the total strength of activity it perceives from the other neurons through the synapse connections. When a binary pattern is introduced to the network, the states of the neurons evolve to a new pattern which is one of those memorised by the network. The memorised patterns and the way they are associated with the inputs are stored in the synapses in a distributed format - if one synapse fails the memory remains intact. The network is also tolerant to noise and to variations in the components, so it is not like an ordinary computer memory.

The Hopfield model [1] expresses the synapse weights as a matrix, and the remembering process is a sequence of matrix-vector multiplications (figure 5.2a). The states of the neurons at any time are expressed as a vector whose elements are +1 or -1. The neuron vector one cycle later is found by multiplying by the weight matrix to give a vector of continuous values, and then thresholding this vector to a set of bistable elements. In the thresholding process a value which is greater than zero becomes +1, less than zero -1. The synapse matrix is the sum of the outer products of the memorised vectors. It has been shown that after several iterations the network can converge onto the memorised vector which is closest to the original input vector.

Implementing a system like the Hopfield network in electronics becomes difficult for a large number of neurons, because the movement of data and

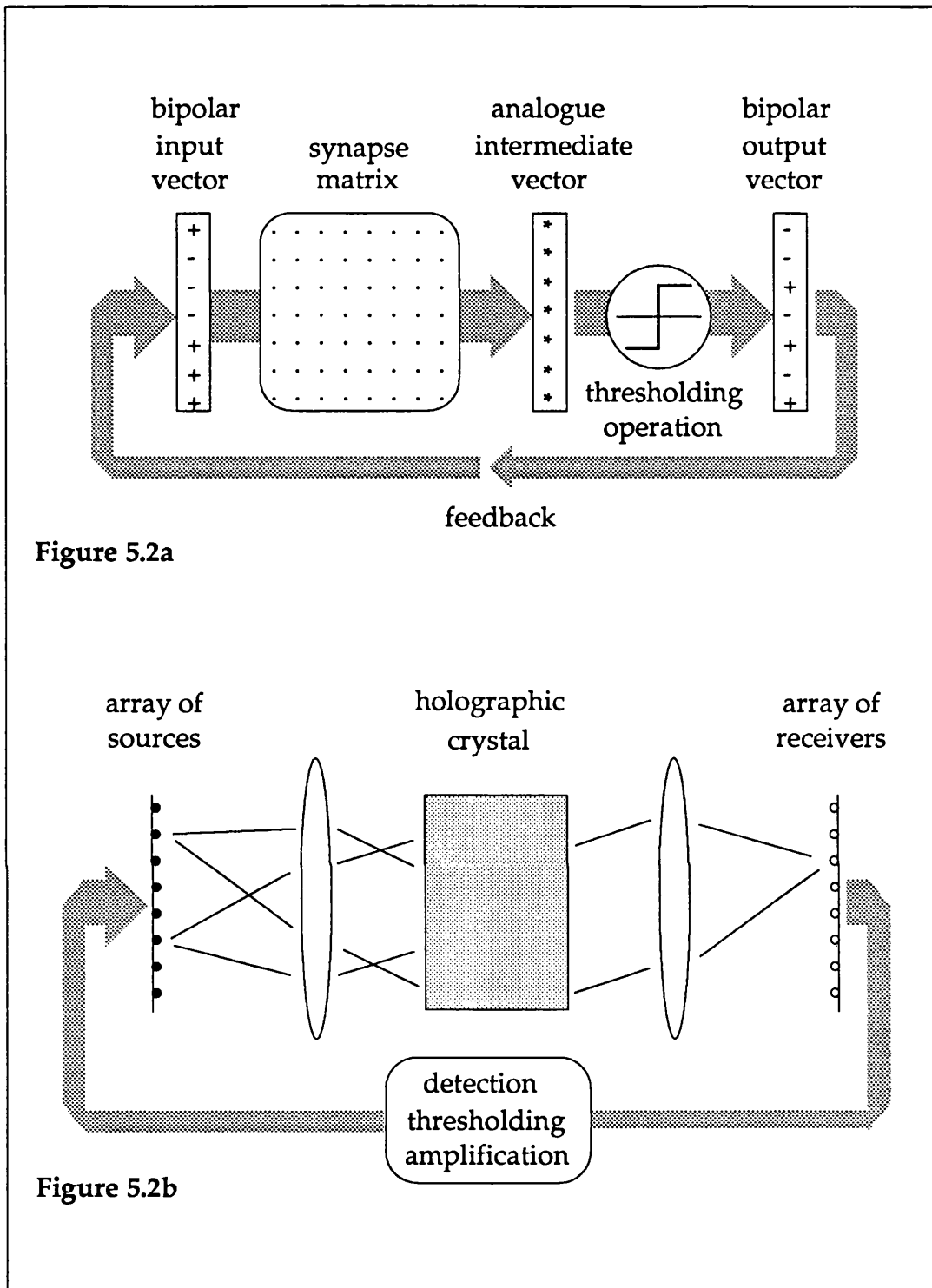


Figure 5.2a The Hopfield network.
 b An optical arrangement for the matrix-vector multiplication part of the Hopfield network.

the matrix manipulations are not suited to that technology. The situation is reversed with optics [2]. A volume hologram in the Fourier domain can interconnect two planes of neurons with variable weights of interconnection, so that it performs the matrix-vector multiplication, and such holograms can cope with a huge number of interconnections. The optical implementation of the Hopfield network then looks like figure 5.2b. There have been experimental demonstrations using a column of LEDs on the left and a row of photodetectors on the right, with a two dimensional mask as a matrix and a microprocessor to control the feedback [3], while the most advanced experiments use thick holograms and a two-sided liquid crystal light valve as a neuron array [4]. The light valve is a reflective spatial light modulator whose elements are controlled by the amount of light falling on the far side; it performs the detection, amplification and thresholding operations, and so completes the feedback loop of figure 5.2b.

The fastest light valves have ferroelectric liquid crystals which can switch in several microseconds. Multiple quantum well modulators might be able to operate 100 times faster than this, and the t-SEED is the best device to do the job. What is required of a neuron element is a device with two output levels, which has a high output for inputs above a certain level. It must have gain, the value of which depends upon the way the holographic crystal is used. The signal can be encoded on the amplitude of the light wave E , in which case there must be coherence between all the light sources so that the input to the next layer of neurons is the true sum of the fields; or the signal can be encoded on the optical power E^2 , and the sources can be mutually incoherent. In a system with N neurons the crystal has N^2 gratings and the mean power efficiency of a grating is up to N^{-1} . This means that in the case of incoherent illumination the neuron must have a gain G of

$$G_{\text{incoherent}} \approx N^{-1}$$

but because the waves add constructively with coherent illumination the equivalent parameter is [5]

$$G_{\text{coherent}} \approx 1$$

In fact the gain must be large enough to overcome any propagation losses, and the transistor gain of the t-SEED should be appropriate. The coherent method cannot be used if each neuron emits light. To have coherence the neuron element must modulate an external beam, as is the case with the t-SEED.

A combination of two t-SEEDs one after the other (figure 5.3a) has the right shape of forward characteristic. The way in which the output of one device drives the input of the next is shown in figures 5.3b&c. It has been

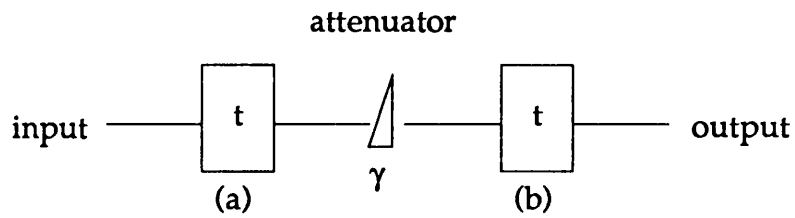


Figure 5.3a

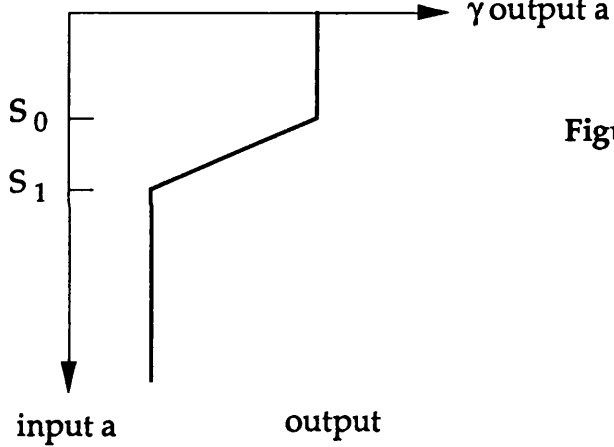
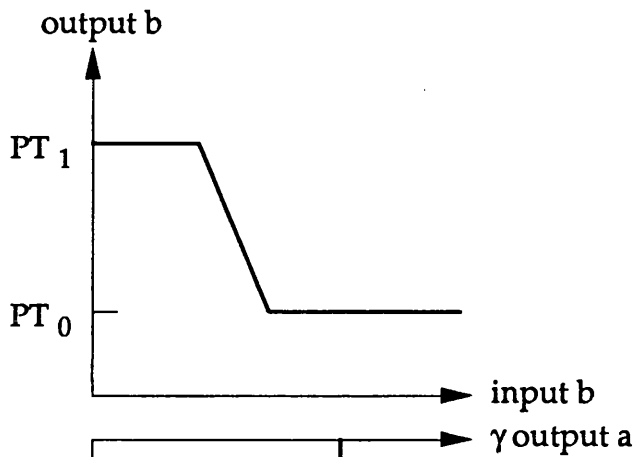


Figure 5.3b

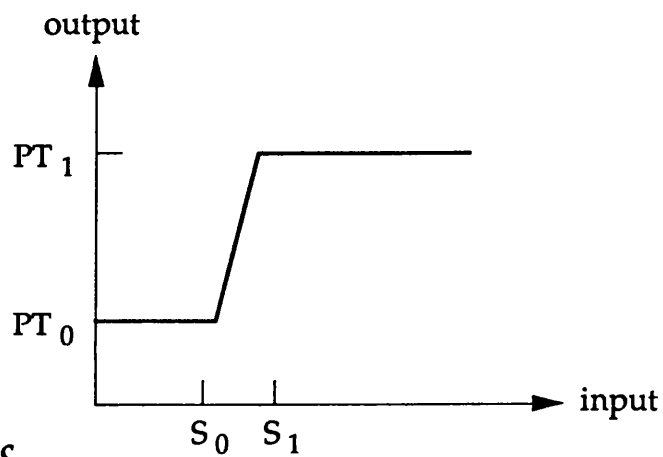


Figure 5.3c

Figure 5.3a Two t-SEEDs in sequence.
b How the characteristics of SEEDs (a) and (b) meet.
c The input-output curve of the combined unit.

noted that the Hopfield network can converge to an incorrect state [6], but this can be avoided by starting off with a soft neuron threshold and making it sharper as the remembering process progresses. Such a time varying transfer curve shape can be achieved with t-SEEDs. When the output of one t-SEED is fed back into its own input the effective transfer characteristic is softened (figure 5.4a). A feedback fraction γ causes the output o as a function of input i to become subject to a load line

$$i = \gamma o + i_{\text{ext}}$$

where i_{ext} is the externally applied input (figure 5.4b). The feedback leads to the characteristic of figure 5.4c. A t-SEED with feedback followed by another t-SEED, as shown in figure 5.5a, can have the range of transfer characteristics of figure 5.5b. This is done by varying two of the attenuators. In practice one would use the same attenuators for a whole array of neurons, and so the shape of the thresholds of all the neurons can easily be changed together.

The Hopfield network strictly requires neuron outputs of +1 and -1, and positive and negative weights in the tensor matrix. This is difficult to realise using optics, and the problem has been considered by other workers. Psaltis [5] notes that with coherent illumination where the signal E is recovered by mixing with a reference, it might be possible to encode positive and negative numbers. White [7] has shown that the positive and negative values can be separated into two distinct networks which evolve separately, and the result required is the difference of the two convergent solutions. However, the reason the Hopfield network is so well established is that it provides a theoretical link between neural processes and the way they recover stored information. It is almost certainly the case that an effective associative memory can be made from a hologram-threshold feedback loop without sticking to the principles of the Hopfield network.

It might become possible to assemble other detector-amplifier-quantum well modulator combinations which have better performance, say speed, than the t-SEED. It is clear though that what is required for neural processors is not very fast devices - the speed comes from the parallelism - or exact device characteristics. A useful neural network must have a large number of neurons, and so the simple vertically integrated t-SEED which can be made in densely packed arrays may be ideal for these applications.

5.3 Applications of t-SEED arrays as logic gates

The t-SEED was devised as a logic gate, and the neural application above does not really utilise it as a gate. Architectures are being developed

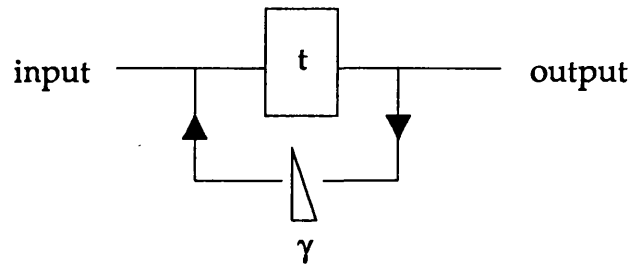


Figure 5.4a

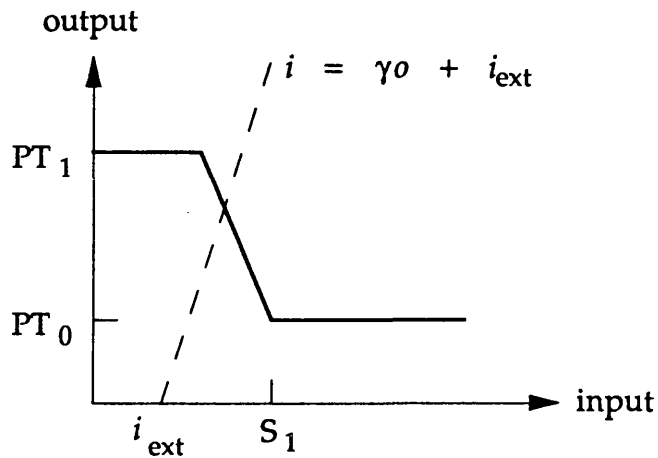


Figure 5.4b

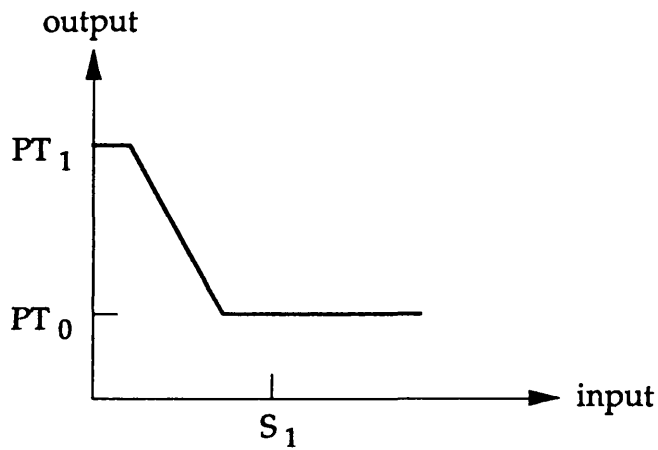


Figure 5.4c

Figure 5.4a A t-SEED with a fraction γ of the output fed back into the input.

b How the new characteristic is determined by a load line.

c The input-output characteristic of the unit.

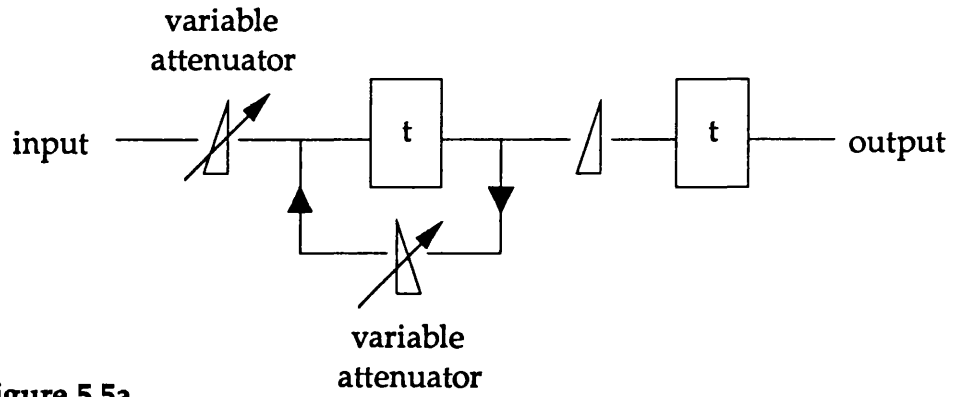


Figure 5.5a

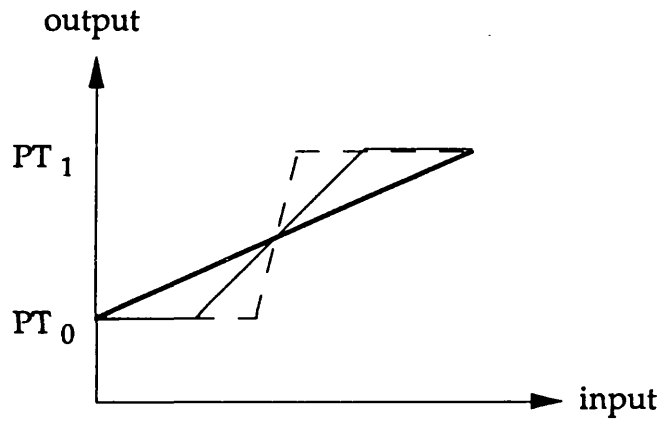


Figure 5.5b

Figure 5.5a Arrangement of two t-SEEDs in sequence and two variable attenuators to give a variable forward characteristic.

b Three examples (fine line, dashed line, broad line) of the transfer characteristics possible using the arrangement in a.

specifically for processing planes of binary data with arrays of optical gates, of which the t-SEED is just one example. The foremost of these architectures is symbolic substitution [8], which has close relatives in OPALS [9], binary image algebra [10] and proposals for optical cellular processors [11-12].

Symbolic substitution deals with planes of bits, encoded in intensity or in polarisation [13], by picking out a search pattern and replacing it with a new pattern. The searching is done by performing AND operations of the image with copies of itself that have been displaced through different vectors. A 1 results wherever the search pattern is found, and this is turned into the replacement pattern by the converse of the search process - making several replicas, displacing them and ORing them with one another. The hardware required is a way of splitting an image into copies, a mechanism for displacing an image through a certain vector, and an array of optical gates for the logic and image regeneration. The t-SEED is suitable for the latter task. The ability to replace one pattern of bits with another translates into a way of doing Boolean logic operations in parallel, and so symbolic substitution can be applied quite generally. It has been demonstrated in simple tasks like edge detection [14] and in a full adder [15], and a methodology for expressing any algorithm in symbolic substitution form has been laid down [16], though this involved considerable redundancy in the optical gates. However, it has been noted that symbolic substitution does not appear to be practical for general purpose computing, and specialist applications must be justified by the need for the connectivity of optics [17].

We know from chapter 2 that the regime for optical interconnects is large distance communication, so our optical array processing architecture should either use global interconnects (as opposed to local interconnects as in a cellular processor) or have a pipelined multistage form making use of long range one-to-one mappings. It seems to this writer that the problem of translating a real computational task into an algorithm for optical gate arrays which makes effective use of optical interconnects has not been solved yet.

5.3.1 An image processing network

To investigate the applicability of the t-SEED device for logic operations a simple image processing system was devised, based around the multistage machines of chapter 2.

The network of figure 5.6a appears in Stone [18] as a demonstration of how the perfect shuffle interconnection can be used for parallel computing. The task performed by the machine is to evaluate the sum of eight numbers in parallel, together with the seven running sums, that is the seven

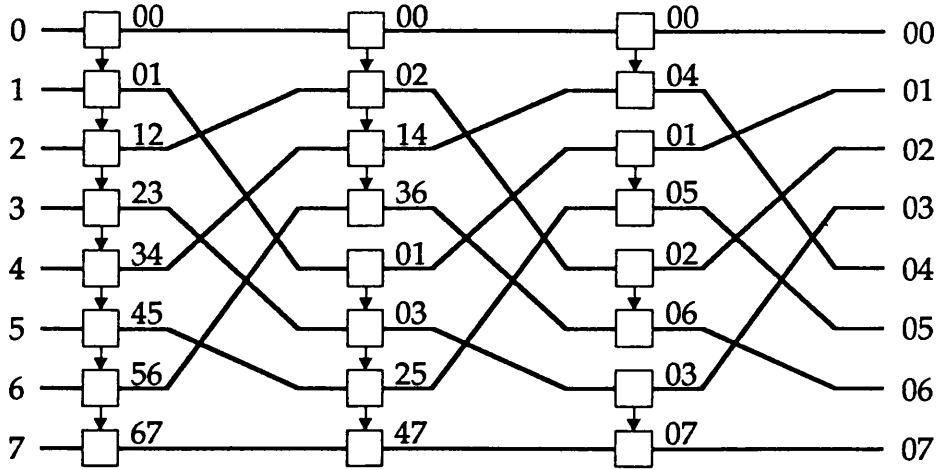


Figure 5.6a

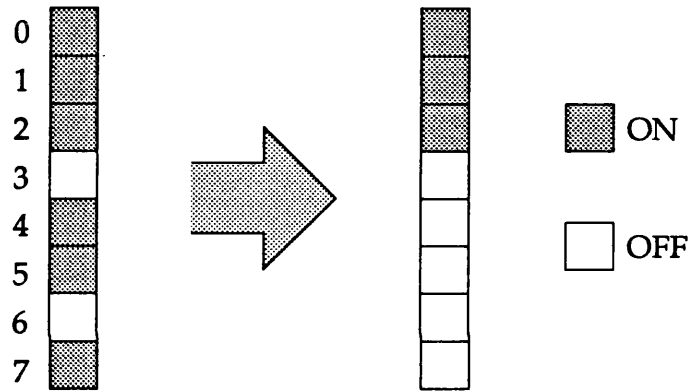


Figure 5.6b

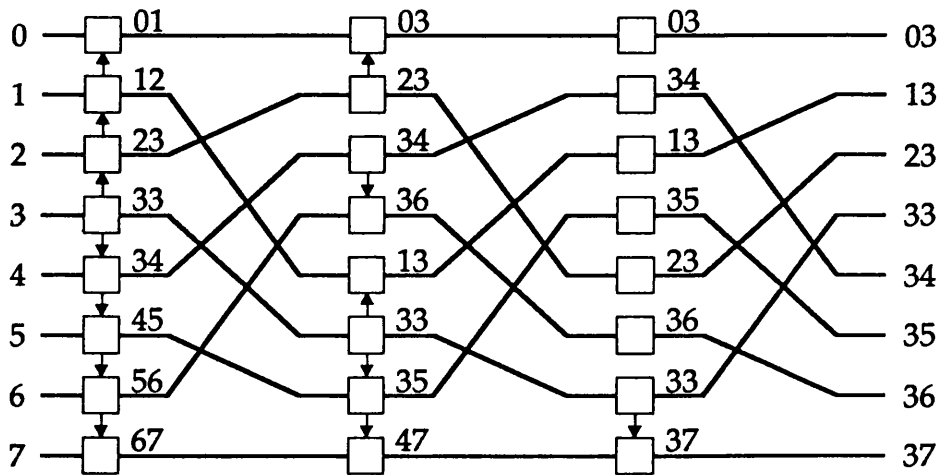


Figure 5.6c

Figure 5.6a A multistage network for evaluating the sum of eight numbers together with the intermediate sums.

b Example input to the image processing network, and the resulting output.

c A variation of a where the test point is at pixel 3 instead of pixel 0.

intermediate results which would be obtained by adding the numbers in sequence. Each box is an adder. The data is input on the left and flows to the right. Where a box has an arrow leaving it, it passes its input to the neighbouring box. A box having two numbers input to it adds them together and outputs the sum; a box with one input returns that number. The pair of figures written above each output in figure 5.6a are the start and finish of the range of numbers over which that line carries the sum. The line with "3 6" above it has the sum of number 3, number 4, number 5 and number 6. It can easily be checked by following the reverse perfect shuffles that the results on the right hand side are the running sums "0 0" to "0 7".

The image processor has the same structure as this parallel adder. The task it performs is this: given a one dimensional pixellated binary image (a line of ONs and OFFs) which contains several objects, the processor returns an image including only the object, if there is one, which makes contact with the top edge. The example in figure 5.6b contains three objects. Only the object comprising pixels 0, 1 and 2 touches the top edge, and so the output has only these three pixels ON, the remainder are OFF.

This process can be performed by a network like the one shown in figure 5.6a where the boxes are AND gates rather than adders. Each output at the right hand side now says whether there is a continuous line of ONs from pixel 0 to that pixel number. The output marked "0 3" is pixel 0 AND pixel 1 AND pixel 2 AND pixel 3. There is no need to redraw figure 5.6a in going from the summing to the image processing network because the flow of data is exactly the same.

It is possible for the test point to be somewhere other than pixel 0. Figure 5.6c shows a network whose result is the object contacting pixel 3. The only difference between this and the original network is the arrows between boxes, which is the local movement of data between the processing elements.

Clearly a two dimensional version of this machine would be considerably more interesting than the 1D processor. The 2D machine would perform an equivalent process: given an image of several objects as in figure 5.7a, it would pick out just the object touching the test point. The 1D network can indeed be expanded into two dimensions, and in the spirit of the work of chapter 2 this network comprises stages of two dimensional reverse perfect shuffles and layers of logic. The logic required is AND operations again. Instead of one way of passing data to the next cell, there are now three kinds of local data movement: horizontal, vertical and diagonal, and a mixture of these is used at each plane. To process an $N \times N$ image, $\log_2 N$ stages are needed.

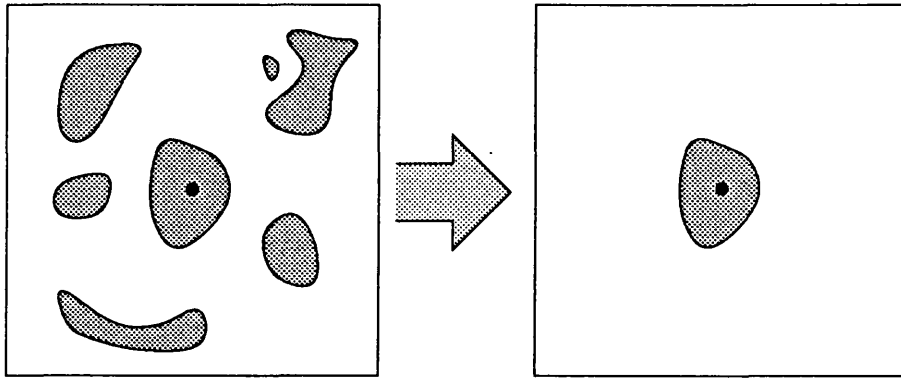


Figure 5.7a

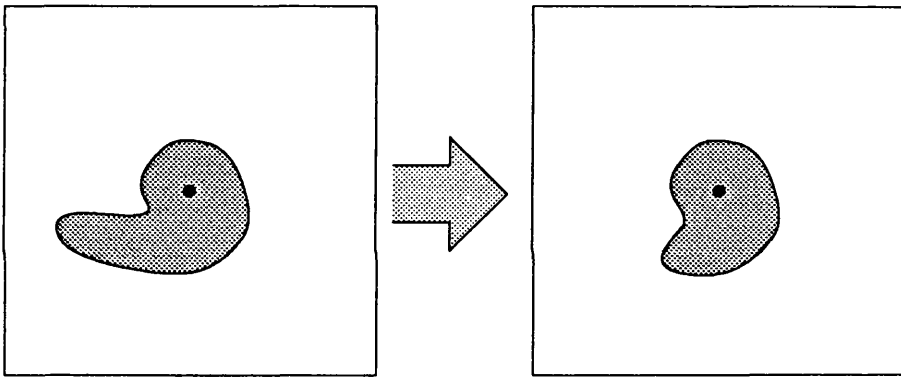


Figure 5.7b

Figure 5.7a Selection of the object making contact with a test point out of an image containing several objects.

b How the image processor can erroneously deal with an object which does not have sufficiently regular shape.

The processor works by looking at a specific path from a pixel (x, y) to the test point $(0, 0)$ to see if it is filled with 1s. This means that the machine can make mistakes, and reject pixels where there is a path to the test point that is not the one being examined. For example the object on the left of figure 5.7b is processed into the one on the right, with a portion erroneously chopped off. Despite these distortions it seems probable that an application might exist for this clutter removing machine.

The hardware required, then, is a means of performing a reverse perfect shuffle, making copies of an image, doing translations on a subset of pixels, and an array of AND gates. The reverse perfect shuffle can be done by the optical arrangement for a forward perfect shuffle, such as that in figure 2.4, working backwards. Extra spots of light will be generated at the output because the system is not quite reversible, but these can be masked out. The displacement of a subset of pixels can be done by masking out pixels which are not to be moved, and then using the translating optics of Brenner et al. [8]. Our candidate for the AND gate array is obviously the t-SEED; two devices in sequence perform an AND operation. All the elements exist then for an optical implementation of the image processor.

The machine performs a global process, in that the final value of a pixel at (x, y) depends on all the pixels between it and the test point. An algorithm which stepped outwards from the test point would take N stages to process an $N \times N$ array; our processor needs only $\log N$ stages because it works in parallel. It is pipelined because it works in a sequence of stages. The conclusion being reached here is that the image processor should fall into the category of machines which do make effective use of optical interconnects. In fact these ideas on data flow were kept in mind when designing the processor more than the need to find a solution for the clutter removal application.

The optical elements are passive, and so the t-SEEDs determine the throughput of the system. If one were to construct a system around 128×128 arrays of t-SEEDs (64×64 arrays of MQW modulators are commercially available at the moment), and achieve the 20ns switching time calculated in section 4.6, this implies a data rate of 800Gbit/s. This is an impressive value. Making a realistic comparison with an electronic computer is very difficult, but reference 19 gives details of some prototype parallel computers, and a data rate of 50Gbit/s is typical for them. Note that any speedup of the optical machine is associated with the architecture more than with the t-SEED processing elements.

5.4 Conclusions

The conditions for a t-SEED to be able to drive another device have been calculated; they are that the transistor should have gain, and the modulator should have a high contrast.

Two processors have been described for which the t-SEED is suitable. The optical implementation of a Hopfield-like neural network uses the t-SEED as a thresholding neuron element, and the potential for the t-SEED to be fabricated in large arrays makes it a good candidate for this application. A new image processing network has been introduced for selecting one object out of a binary image containing several objects. The network is based on 2D perfect shuffle interconnections, and employs arrays of AND gates which could be t-SEEDs. The network is very simple, and was developed chiefly to provide an example of a system which makes good use of the t-SEED by making efficient use of optical interconnects. A simple analysis gave a projected data rate for the machine of 800Gbit/s, which indicates that the t-SEED can be employed in an optically interconnected system so as to give superior performance to electronics.

5.5 References

- [1] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Natl. Acad. Sci. USA*, vol. 79, pp. 2554-2558, 1982.
- [2] D. Psaltis & N. Farhat, "Optical information processing based on associative-memory model of neural nets with thresholding and feedback," *Opt. Lett.*, vol. 10, pp. 98-100, 1985.
- [3] N. H. Farhat, D. Psaltis, A. Prata & E. Paek, "Optical implementation of the Hopfield model," *Appl. Opt.*, vol. 24, pp. 1469-1475, 1985.
- [4] Y. S. Abu-Mostafa & D. Psaltis, "Optical neural computers," *Scientific American*, pp. 66-73, March 1987.
- [5] D. Psaltis, D. Brady & K. Wagner, "Adaptive optical networks using photorefractive crystals," *Appl. Opt.*, vol. 27, pp. 1752-1759, 1988.
- [6] D. R. Selviah, J. E. Midwinter, A. W. Rivers & K. W. Lung, "Correlating matched-filter model for analysis and optimisation of neural networks," *IEE Proc.*, vol. 136, Pt. F, pp. 143-148, 1989.
- [7] H. J. White, "Experimental results from an optical implementation of a simple neural network," *Optical Computing, SPIE*, vol. 963, pp. 570-575, 1988.
- [8] K.-H. Brenner, A. Huang & N. Streibl, "Digital optical computing with symbolic substitution," *Appl. Opt.*, vol. 25, pp. 3054-3060, 1986.
- [9] J. Tanida, M. Fukui & Y. Ichioka, "Programming of optical array logic.

- 2: Numerical data processing based on pattern logic," *Appl. Opt.*, vol. 27, pp. 2931-2939, 1988.
- [10] K.-S. Huang, B. K. Jenkins & A. A. Sawchuk, "Image algebra representation of parallel optical binary arithmetic," *Appl. Opt.*, vol. 28, pp. 1263-1278, 1989.
- [11] T. Yatagai, "Cellular logic architectures for optical computers," *Appl. Opt.*, vol. 25, pp. 1571-1577, 1986.
- [12] M. J. Murdocca, "Digital optical computing with one-rule cellular automata," *Appl. Opt.*, vol. 26, pp. 682-688, 1987.
- [13] K.-H. Brenner, "New implementation of symbolic substitution logic," *Appl. Opt.*, vol. 25, pp. 3061-3064, 1986.
- [14] A. K. Cherri & M. A. Karim, "Edge detection using optical symbolic substitution," *Opt. Comm.*, vol. 74, pp. 10-14, 1989.
- [15] T. M. Merklein, W. Stork & H. Yajima, "Optical full adder," *Appl. Opt.*, vol. 28, pp. 4313-4315.
- [16] A. Huang "Computational origami - the folding of circuits and systems," *Optical Computing, 1989 Technical Digest Series*, vol. 9, (OSA), pp. 198-201, 1989.
- [17] T. J. Cloonan, "Performance analysis of optical symbolic substitution," *Appl. Opt.*, vol. 27, pp. 1701-1707, 1988.
- [18] H. S. Stone, "High-performance computer architecture," Addison-Wesley, 1987.
- [19] R. J. Offen "Whither image processing?" in *VLSI Image Processing*, R. J. Offen, Ed., McGraw-Hill, 1985.

CHAPTER SIX

Conclusions

The main part of this Ph.D. project was the development of a monolithic t-SEED device, that is the vertical integration of a heterojunction bipolar phototransistor with a multiple quantum well modulator. The first attempt at integration had the transistor grown on top of the MQW modulator. The device failed because of thyristor action involving the top npnp combination of layers. The next structure to be grown had the transistor underneath the modulator with the two components sharing a layer. This device did not suffer from thyristor action, but the transistor had a low gain, which was diagnosed as being due to inadequate conversion of the incident light into primary photocurrent, and also a low breakdown voltage. The solution was to grow a wide low doped collector layer, which would be fully depleted at a low voltage and so collect a large proportion of the light. A third wafer was grown to this specification. It was discovered that it is critical to remove all of the gallium arsenide material when etching the window in the substrate, because any remnants of the base or collector regions reduce the transistor gain to unity. Ultimately no good devices were obtained from the third growth run; the final batch had a low gain which was thought to be because of damage to the emitter-base junction while etching the rear window.

The next step should be to try different ways of etching the window to see if the transistor gain can be recovered. A new flip-chip configuration for the t-SEED has been proposed which would not require windowing the substrate, and so should give much more robust devices. It involves working in reflection, and so high contrasts might be possible using Fabry-Perot enhancement of modulation.

As well as the development of the t-SEED there has been some work on the MQW modulator itself. In a real modulator it can be difficult to extract information about the absorption spectrum because of Fabry-Perot effects. A new model of the Fabry-Perot cavity has been introduced which makes it easier to understand the interference effects and to pick out relationships between measurable parameters. It has been shown that the features of the absorption spectrum can be isolated from the Fabry-Perot resonances by dividing the photocurrent in the modulator by the transmission. Similar methods have been put forward for eliminating cavity effects from the

absorption spectrum in reflection and waveguide configurations.

The non-zero doping in the intrinsic region of the MQW pin modulator can limit the contrast ratio possible. A way to overcome the detrimental effects of the background doping by grading the well widths has been proposed. Some simple modelling suggests that the approach may be useful, but in a material system other than the GaAs:AlGaAs system, and for transmission modulators because high contrasts can be obtained in reflection with other methods.

On the systems side, it was noted that optical interconnects are best used to replace long range electrical interconnects in a parallel processing system, and an optimal system would have islands of electronics linked by optics. While the multistage interconnection network is a good example of a machine which benefits from optical interconnects, it can be improved upon. It was shown by a series of examples how to map any multistage network into two dimensional format, so that it makes full use of the space available to the optics. Mapping rules were also found to vary the size of the switching units, so that they could be fixed to the optimum electronic island size and so make the best use of optics as an interconnect. A brief evaluation of an optical implementation of a multistage switch indicated that it could be superior to an all-electronic one.

The role the t-SEED can play in an optical computing system has been considered. In order for one t-SEED to be able to drive another it must have a high contrast from the modulator and a high gain from the phototransistor. The switching time of the device is limited by the time taken for the input beam to discharge the collector capacitance through the applied voltage. Using reasonable values for the relevant parameters a switching time of 20ns was calculated.

Two systems which employ the t-SEED were discussed. The optical implementation of a neural network uses the device as a neuron element. While many technologies might be capable of doing this job, the t-SEED has the important attributes of high speed and availability in large densely packed arrays. The second system considered was an image processing network which was devised specifically to make the best use of optics as a parallel interconnect, so as to make good use of the t-SEED. The task performed by the image processor is to select one object out of a binary image containing several objects. The throughput of the system depends on the t-SEED switching time, and using the projected value of 20ns a very high data rate was obtained, which indicates that the t-SEED can be applied in such a way as to achieve better performance than electronics by itself.

Contributions have been made to the fields of optical computer design and MQW device development. The program to develop the t-SEED has increased our knowledge of the issues involved in extending the functionality of an MQW modulator by integrating it with other components. The work on the switching networks has shown how optical interconnects can be best employed in a real system. The work in the two areas fits together in demonstrating which systems are appropriate for applying the t-SEED.

APPENDIX A

Proof of the standard Fabry-Perot cavity equations starting from the five wave model

The principles behind the five wave model of the Fabry-Perot cavity are explained in section 3.3.1.

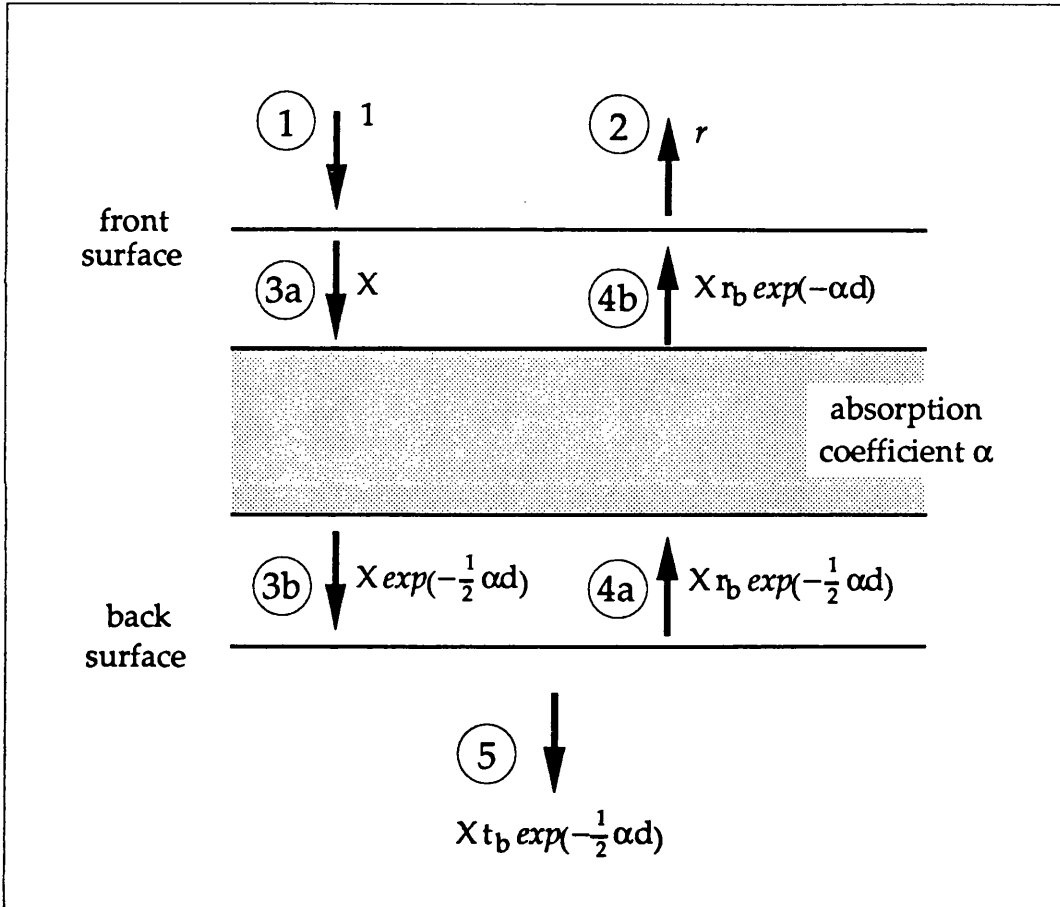


Figure A.1. Amplitudes of waves in a Fabry-Perot cavity.

Figure A.1 shows the amplitudes of the waves in the cavity in the five wave model as far as they can be written down. The amplitude of the incident wave (1) is 1. Wave (3a) has amplitude X , and the amplitudes of (3b), (5), (4a) and (4b) are written down in figure A.1 in terms of X . The amplitude reflection and transmission coefficients of the front and back interfaces are r_f ,

r_b , t_f and t_b (all positive quantities), and the absorbing region attenuates the amplitude of a wave by $\exp\left(-\frac{1}{2}\alpha d\right)$ (note the factor of $\frac{1}{2}$ because we are dealing with amplitudes and not optical powers). Let the amplitude of the reflected wave be r .

Wave (3a) is the sum of (1) and (4b) taking phase into account. If the phase difference between (1) and (4b) is ϕ , then

$$X = t_f + X r_b r_f e^{-\alpha d} e^{j\phi}$$

Solving for X ,

$$X = \frac{t_f}{1 - r_b r_f e^{-\alpha d} e^{j\phi}}$$

$$X = \frac{t_f}{1 - R_\alpha e^{j\phi}}$$

where

$$R_\alpha = r_b r_f e^{-\alpha d}$$

The reflected wave (2) is the sum of the reflected part of (1) and the transmitted part of (4b). Remembering that there is a π phase change involved, the sum is

$$r = -r_f + X r_b t_f e^{-\alpha d} e^{j\phi}$$

$$r = -r_f + \frac{r_b t_f^2 e^{-\alpha d} e^{j\phi}}{1 - R_\alpha e^{j\phi}}$$

$$r = \frac{-r_f + R_\alpha r_f e^{j\phi} + r_b t_f^2 e^{-\alpha d} e^{j\phi}}{1 - R_\alpha e^{j\phi}}$$

which becomes using

$$r_f^2 = R_f$$

and

$$r_f^2 + t_f^2 = 1$$

$$r = \frac{-R_f + (R_\alpha R_f + R_\alpha(1-R_f))e^{j\phi}}{r_f(1 - R_\alpha e^{j\phi})}$$

$$r = \frac{-R_f + R_\alpha e^{j\phi}}{r_f(1 - R_\alpha e^{j\phi})}$$

The power reflection coefficient R is

$$R = r^* r = \frac{R_f^2 - 2R_\alpha R_f \cos\phi + R_\alpha^2}{R_f(1 - 2R_\alpha \cos\phi + R_\alpha^2)}$$

which with

$$\cos\phi = 1 - 2\sin^2\frac{1}{2}\phi$$

becomes

$$R = \frac{(R_f - R_\alpha)^2 + 4R_\alpha R_f \sin^2\frac{1}{2}\phi}{R_f(1 - R_\alpha)^2 + 4R_\alpha R_f \sin^2\frac{1}{2}\phi}$$

$$R = \frac{B + F \sin^2\frac{1}{2}\phi}{1 + F \sin^2\frac{1}{2}\phi} \quad (\text{A.1})$$

where

$$B = \frac{R_f \left(1 - \frac{R_\alpha}{R_f}\right)^2}{(1 - R_\alpha)^2} \quad F = \frac{4R_\alpha}{(1 - R_\alpha)^2}$$

The transmitted amplitude t can be written down without dealing with any interfering waves.

$$t = \chi t_b \exp\left(-\frac{1}{2}\alpha d\right)$$

$$t = \frac{t_b t_f \exp\left(-\frac{1}{2}\alpha d\right)}{1 - R_\alpha e^{j\phi}}$$

The transmitted power is

$$T = t^* t = \frac{t_b^2 t_f^2 e^{-\alpha d}}{1 - 2R_\alpha \cos\phi + R_\alpha^2}$$

$$T = \frac{(1-R_b)(1-R_f)e^{-\alpha d}}{(1-R_\alpha)^2 + 4R_\alpha \sin^2 \frac{1}{2}\phi}$$

$$T = \frac{C}{1 + F \sin^2 \frac{1}{2}\phi} \quad (\text{A.2})$$

where

$$C = \frac{(1-R_b)(1-R_f)e^{-\alpha d}}{(1-R_\alpha)^2}$$

The expressions for R and T in equations (A.1) and (A.2) are the standard Fabry-Perot equations which are to be proved. Therefore, the five wave model of the FP cavity is equivalent to the standard approach.

APPENDIX B

Epitaxial layer structures

B.1 MV247

n+	GaAs	0.1 μm
1.3x10 ¹⁷ cm ⁻³ n	AlGaAs (30%)	0.5 μm
2x10 ¹⁷ cm ⁻³ n	graded	300Å
5x10 ¹⁷ cm ⁻³ p	GaAs	0.2 μm
5x10 ¹⁷ cm ⁻³ n	GaAs	0.5 μm
10 ¹⁸ cm ⁻³ n	GaAs	0.2 μm
10 ¹⁸ cm ⁻³ p	AlGaAs (30%)	0.5 μm
undoped	60 x [82Å GaAs + 60Å AlGaAs (30%)]	0.85 μm
10 ¹⁸ cm ⁻³ n	AlGaAs (30%)	0.56 μm
10 ¹⁸ cm ⁻³ n	AlGaAs (50%)	0.13 μm
n+	GaAs substrate	

B.2 CB27

p+	GaAs	0.1 μm	
$1 \times 10^{18} \text{ cm}^{-3} \text{ p}$	AlGaAs (30%)	0.5 μm	
undoped	AlGaAs (30%)	400 \AA	
undoped	60 x [90 \AA GaAs + 60 \AA AlGaAs (30%)]	0.9 μm	
$5 \times 10^{17} \text{ cm}^{-3} \text{ n}$	AlGaAs (30%)	0.5 μm	n
	parabolic grade of Al	300 \AA	n
undoped	GaAs	100 \AA	n
$5 \times 10^{17} \text{ cm}^{-3} \text{ p}$	GaAs	0.2 μm	p
undoped	GaAs	300 \AA	p n
$5 \times 10^{17} \text{ cm}^{-3} \text{ n}$	GaAs	0.5 μm	n
n+	substrate		

B.3 CB245

p+	GaAs	0.1 μm
$1 \times 10^{18} \text{ cm}^{-3}$ p	AlGaAs (30%)	0.5 μm
undoped	AlGaAs (30%)	400 \AA
undoped	60 x [85 \AA GaAs + 60 \AA AlGaAs (30%)]	0.87 μm
$5 \times 10^{17} \text{ cm}^{-3}$ n	AlGaAs (30%)	0.5 μm
	parabolic grade of Al	300 \AA
undoped	GaAs	100 \AA
$5 \times 10^{17} \text{ cm}^{-3}$ p	GaAs	0.2 μm
undoped	GaAs	300 \AA
$2.5 \times 10^{16} \text{ cm}^{-3}$ n	GaAs	1.0 μm
n+	GaAs buffer layer	
n+	substrate	

Papers written in connection with this thesis

M. G. Taylor & J. E. Midwinter, "A novel two dimensional perfect shuffle network," presented at OSA conference on Photonic Switching, Salt Lake City, March 1989.

M. G. Taylor & J. E. Midwinter, "Two-dimensional perfect shuffle networks," in OSA Proceedings on Photonic Switching, John E. Midwinter & H. Scott Hinton, Eds. (Optical Society of America, Washington D.C., 1989), vol. 3, pp. 180-183.

M. G. Taylor & J. E. Midwinter, "Method of obtaining algorithms for nonblocking two-dimensional optical multistage interconnection networks," *Electron. Lett.*, vol. 25, pp. 1641-1643, 1989.

M. G. Taylor, A. W. Rivers, J. S. Roberts, C. Button, G. Hill & M. Pate, "A multiple quantum well optical logic device," presented at Semiconductor and Integrated Optoelectronics '90, Cardiff, March 1990.

M. G. Taylor & J. E. Midwinter, "Optically interconnected switching networks," submitted to IEEE Journal of Lightwave Technology, March 1990.

M. G. Taylor & J. E. Midwinter, "Multistage switching networks using optical interconnects," IEE Colloquium on Optical Connection and Switching Networks, 14th May 1990.

J. E. Midwinter & M. G. Taylor, "Optoelectronic interconnects in VLSI: The reality of digital optical computing?" IEEE Lightwave Communication Systems Magazine, pp. 40-46, May 1990.

M. G. Taylor, A. W. Rivers, J. S. Roberts, C. Button, G. Hill & M. Pate, "Monolithic integration of a heterojunction transistor with a multiple quantum well modulator," IEE Colloquium on Applications of Quantum Wells in Optoelectronics, 13th June 1990.

Patent applied for

"Optical interconnect networks," UK patent application no. 890421.6, filed 24th February by British Telecom.