INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

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

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Université d'Ottawa · University of Ottawa

THE FABRICATION AND OPTICAL PROPERTIES OF QUANTUM WIRES

by

Paul D. Finnie

A thesis submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Ottawa

1996

Paul D. Finnie, 1996.

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-26116-6

Canada

University of Ottawa

Abstract

THE FABRICATION AND OPTICAL PROPERTIES OF QUANTUM WIRES

by Paul D. Finnie

Chemical beam epitaxy was used to grow GaAs and InP based heterostructures on selectively masked substrates. The natural tendency of crystals to form facets was studied as a function of growth conditions. By exploiting this tendency, patterns were reduced from ~1 micron in width to ~20 nm in width, and inverted-V crescent quantum wire structures were grown entirely *in situ*. Photoluminescence spectra were obtained, showing broad multiple peaks which are related to the structure and migration between facets. Redshifts of over 100 meV were obtained for narrow structures. Mesa roughness and PL line width are identified as difficulties for the fabrication technique. A simple analytic model - the only separable model in two dimensions showing geometrical quantum confinement - is developed in the spirit of the Ben-Daniel Duke approximation for a square well. The model is applied to variational estimates for the exciton binding energy in crescent quantum wire structures and is compared to the experimental data.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

Thank you to my supervisor Alain Roth for introducing me to growth and optics, for giving me free reign on this project, and for support and advice at every stage.

Thanks to the University of Ottawa Department of Physics, especially for allowing me the opportunity to work at the NRC. I believe the University can benefit greatly by taking advantage of the resources that are found in the national capital region. I am grateful that they gave me the chance.

My kindest thanks to the Director General of the National Research Council's Institute of Microstructural Sciences, Dr. Peter Dawson. Without his support for graduate students at IMS I would have missed this excellent opportunity.

Thank you to Charles Lacelle for risking life and lung in epitaxy at the NRC. He kept the CBE running like clockwork, and helped with all manner of problems created when my theoretical background couldn't give me the experimental common sense.

Thank you to Margaret Buchanan for her excellent microfabrication work and friendship throughout my time at NRC.

Thanks to Yan Feng for e-beam lithography, but more importantly showing me the usefulness of the FE-SEM. Thanks to Jeff Fraser for operating the FE-SEM.

Thanks to Sylvain Charbonneau and Philip Poole who made available their excellent photoluminescence skills in a pinch more than a couple of times.

Thanks to Emory Fortin for filling in for Alain when he returned to France, and for financial help when money was tight.

Thanks to Robin Williams and Geof Aers for discussion of diffusion processes, and an extra thanks to Geof for letting me use his office as a library after the library was shut down, not to mention help with Unix accounts.

Thanks also to Abdelhak Bensoula for allowing me to use his CBE machine when time was really tight here. Thanks also to Hakim Bensoula for helping me with it.

I have named a few IMS staff explicitly, but I owe a great debt to many others for assistance and for their friendship. If you see this page and I missed your name, just sign at the bottom - you probably deserve to be on here.

My parents will try to read this thesis but they'll fall asleep not long after this page, so I'll thank them here. Of course I have to thank them for everything, but specifically I should thank them for a telescope they gave me for my fifth birthday which started me out on this path.

Finally, I will never be able thank Erin Yoshida enough, so I'll simply thank her last.

INTRODUCTION

The beginning of the twentieth century saw remarkable progress in physics, a field which was thought to be complete but for a few small problems. The resolution to some of those problems came in the form of the revolutionary theory of quantum mechanics. At the end of the twentieth century, despite insinuations about the end of physics[1], the study of nanostructures is developing. The field encompasses the efforts to engineer structures of such a small scale that their properties must be described quantum mechanically.

Arguably, the foremost problem in nanotechnology is fabrication. It is extremely difficult to fabricate structures on nanometre scales. A particular fabrication technique is investigated here. Chapter 1 develops the background needed to understand it.

After describing the materials and methods, the next two chapters describe the analysis of real structures grown as part of this work. Techniques used routinely here to characterize the structures were scanning electron microscopy* (SEM) and photoluminescence (PL). Other characterization techniques were occasionally used, as well. Chapter 2 and Chapter 3 describe

* Including Field Emission SEM (FE-SEM)

the characterization of real structures made on gallium arsenide and indium phosphide substrates, respectively.

The theory of these structures has rapidly developed in only the last decade. To understand and use these structures, many theoretical tools are still needed. A simple model, the only separable model of geometrical confinement in two dimensions, is described in Chapter 4, where it is placed in the context of other previous theoretical work. As an application of the model, in Chapter 5 a variational calculation of the exciton binding energy is performed for crescent structures. This quantity was calculated since optical properties of good quantum confined semiconductors are usually dominated by excitonic effects, and since it tends to infinity as the potential approaches true one dimensional confinement.

The appendices elaborate on technical issues that could not be given much attention in the body of the thesis.

This field has seen rapid progress in the course of the past few years, but it is by no means mature. New fabrication and characterization techniques appear each year, and the theory develops in tandem. This thesis is an attempt to contribute to a growing field that shows great promise at the end of the first century of quantum mechanics.

Chapter 1

FABRICATION TECHNIQUES

Introduction

The scientific and technical interest in nanometre scaled structures has been clear at least since the well publicized lecture "There's Plenty of Room at the Bottom" by Feynman.[2] Chemists have long been sufficiently adept at manipulating atoms and molecules that they are able to synthesize and manipulate larger, more complex structures.[3] Physicists, coming from the very successful study of bulk crystalline materials, and because of advances in deposition technologies, were able to propose[4,5] and synthesize[5,6] crystalline layers of quantum mechanical thinness (i.e. thinner than the electron mean free path in the solid). Since thin layer deposition has been so successful, it might be possible to extend this technology to produce two and three dimensional patterns on similarly small scales. A major part of this work will be the investigation of one such method. Before describing this work specifically it is necessary to develop some background in the fabrication methods.

Epitaxial Growth Techniques

The word *epitaxy* is derived from the Greek *epi* meaning "on", and *taxis* meaning "arrangement".[7] It refers to the deposition on a crystalline substrate of new atomic layers which are themselves arranged in crystalline form. There are many varieties of epitaxy. Very high quality bulk material can be produced by liquid phase epitaxy (LPE), in which the semiconductor material is precipitated onto a substrate from a melt.[8,9] In hydride vapor phase epitaxy

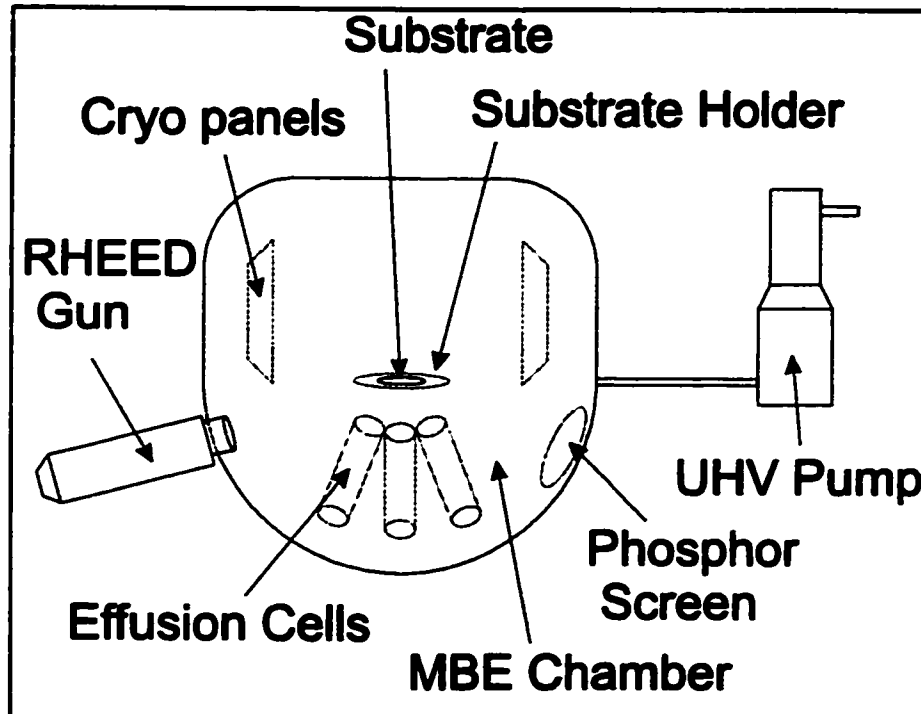


Figure 1. An MBE machine.

The substrate, a thin round wafer, is held on a molybdenum block by the substrate holder. A UHV pump together with cryo panels maintain UHV conditions. Sources are supplied by effusion cells. The surface is monitored by a RHEED gun electron beam which is reflected onto the phosphor screen.

(VPE), in the case of III-V semiconductors, gaseous group V chlorides flow over liquid group III elements to produce gaseous metal chlorides, which are then deposited on a substrate.[8,10] Both of these techniques have been very successful industrially, in particular for the fabrication of bulk materials. For the fabrication of nearly atomically abrupt interfaces between different materials (heterojunctions), however, the most successful techniques are molecular beam epitaxy (MBE) and metalorganic chemical vapor deposition (MOCVD).[9] The growth technique used for this work was chemical beam epitaxy (CBE), a fairly recent combination of these two.[11] Any description of CBE would be incomplete without reference to these more standard growth techniques, so they will be briefly described in the following pages.

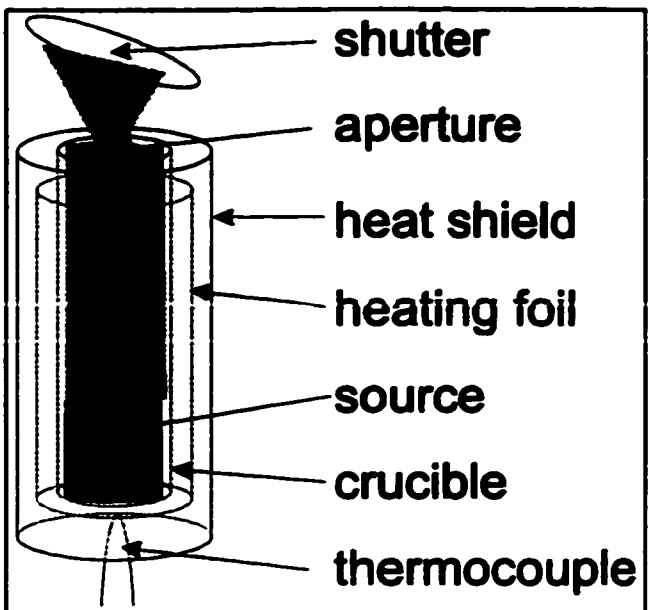


Figure 2. An effusion cell.

The source material is contained in a crucible and heated resistively by a heating foil. A thermocouple measures the temperature. Material sublimates and travels through the aperture towards the sample. A shutter may be used to interrupt the beam.

Molecular Beam Epitaxy

Though not historically the first epitaxial technique, and certainly technically demanding, MBE is probably physically the simplest technique in principle. In MBE[12,13], (Figure 1) a substrate, typically a thin (< 1 mm thick) wafer polished to near perfect flatness across a large area (several cm^2), is placed in ultra-high vacuum (UHV) conditions ($P < 10^{-8}$ atm) and heated to a temperature at which incorporation of other elements is favorable. The elements to form an epilayer are placed in effusion cells (Figure 2), which are essentially cylindrical furnaces each with one open end pointed towards the substrate. (In an idealized picture the open end is a narrow orifice, but with the exception of dopant cells, it is usually necessary to have a wide opening to ensure high enough flux of growth material.) The cell is heated to a temperature such that some of the element sublimates off in a steady fashion. The element escapes the furnace through the opening, usually in a dimer or tetramer form. Since the pressure is so low in the growth chamber, the mean free path exceeds the distance from source to substrate, and the molecules travel directly to the substrate without any scattering. A shutter may be used to block the beam and so abruptly block the flux. The incorporation rate into the substrate crystal is usually much faster than the flux, hence there is no boundary layer on the substrate, except for the incomplete top monolayer on

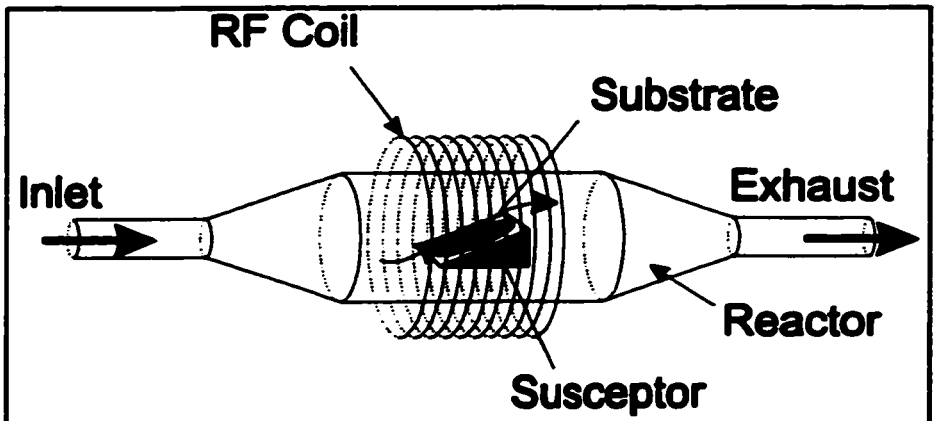


Figure 3. An MOCVD reactor.

Source gases flow through an inlet, across a substrate. The substrate is affixed to a susceptor which is heated by a radio frequency (RF) coil. Exhaust gases are vented.

which atoms and molecules may diffuse. For compound semiconductors, such as III-Vs one cell is typically used for each element, and additional cells are used for dopants.

Metalorganic Chemical Vapor Deposition

In chemical vapor deposition a similar substrate is placed into a reactor chamber. Gaseous sources containing the elements to be incorporated are made to flow past the substrate, which is heated enough to allow the gas to decompose on or near the surface, but not too high so that incorporation of the source elements is thermodynamically favored. Typically, pressures of a fraction of atmospheric are required to ensure a high enough growth rate (1 μm/hour or more).

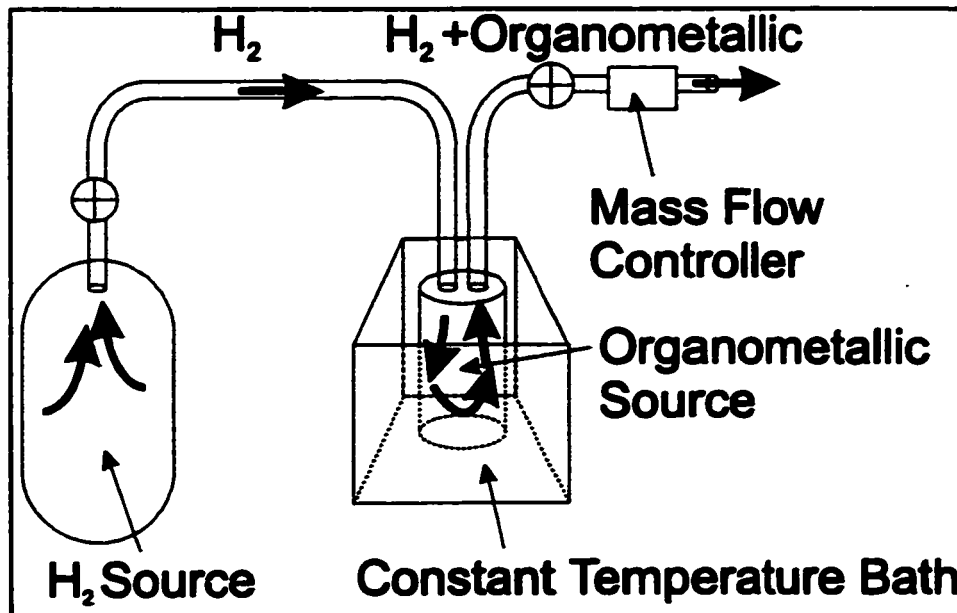


Figure 4. A metalorganic bubbler system.

Hydrogen carrier gas is bubbled through an organometallic source in a flask which is held at a constant temperature by a bath. The outgoing gas is largely hydrogen, together with a smaller amount of organometallic gas. A mass flow controller, here in a downstream configuration, regulates the flow.

For III-V semiconductors, metalorganic sources molecules are the group III sources, hence the name metalorganic chemical vapor deposition (Figure 3). Examples of such sources are trimethylindium (TMI), and triethylgallium (TEG). Because they have low vapor pressures, a hydrogen carrier gas is typically bubbled through a liquid (for TEG) or solid (for TMI) metalorganic source (Figure 4). The group V elements are also introduced in gaseous form, often as hydrides (such as arsine AsH_3 or phosphine PH_3). Since these have a substantial vapor pressure even at room temperature, no carrier gas is

necessary. Flow rates are controlled by constricting the flow with valves known as mass flow controllers.

Chemical Beam Epitaxy

Chemical beam epitaxy (CBE) is a hybrid of the previous two techniques, and inasmuch as this project is concerned, combines advantages of both techniques. Because of the commercial value of extending bandgap engineering (that is, the tailoring of epilayer bandgaps through the choice of material) to new wavelengths, it had been of interest to incorporate phosphorus into epitaxial layers. This had proved to be extremely difficult in MBE, because of the difficulty in evaporating elemental phosphorus sources. In the 1970s the technique of gas source molecular beam epitaxy (GS-MBE) was introduced as an answer to this problem.[14] Rather than supplying the group V elements from effusion cells as evaporated elemental sources, they were supplied as hydrides, which were thermally cracked before introduction to the growth chamber. In CBE this is how the group V elements are introduced. In the growth experiments conducted as part of this work, arsenic and phosphorus were supplied as pure arsine (AsH_3) and phosphine (PH_3). The flow rates, typically in the tens of sccm (standard cubic centimeters per minute) were controlled by mass flow controllers (MFCs). Just as the hydride gas enters the growth chamber, it was cracked by a cracker cell (Figure 5).

In MOCVD, in which materials must be supplied in a molecular vapor form, group III elements are typically supplied as organometallics with a molecular hydrogen carrier gas. In CBE, organometallics are supplied in this way as well, but with a much lower (molecular beam as opposed to gaseous) flow rate. For this reason CBE is sometimes called MOMBE for metalorganic MBE.

Quite apart from the intellectual appeal and technical simplification that result from supplying all sources in gaseous form, CBE has many advantages over earlier epitaxial growth techniques.[11,15] Some of the advantages which relate specifically to this project have to do with selective area growth. (See Selective Area Epitaxy, p18.) In contrast to MBE, the use of organometallic group III sources makes selective area epitaxy possible. In contrast to MOCVD, the beam nature implies that there are no surface flow effects which may detract from uniformity - particularly in the case of selective area epitaxy (see below, p18) in which

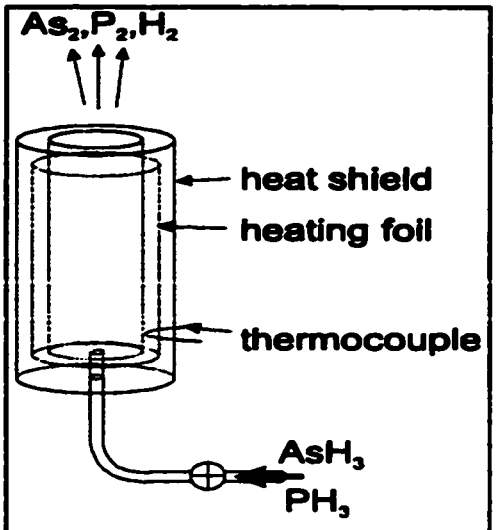


Figure 5. A cracker cell.

Hydrides (here arsine and/or phosphine) flow through a cell which is heated resistively to very high temperatures (~950 °C). This cracks the hydrides, leaving arsenic-only and phosphorus-only compounds (primarily As_2 and P_2), as well as hydrogen gas.

there are variations in the substrate on small (~1 μm) scales. Details of the flow rate control of the gas lines and bubblers are reserved for Appendix A.

Selective Area Epitaxy

All the above growth techniques are used primarily for the deposition of thin *films* of crystalline material. That is to say they deposit basically flat layers across large areas. In fact, historically this is the technological role they were designed to fill. The ability to change composition abruptly has made it possible to produce nanometre scale changes in film composition and bandgap, making the fabrication of quantum wells and heterostructure semiconductor lasers possible. Typically, such film structures are grown, and then, via photolithography, individual devices are fabricated. Finally a metallization step contacts the devices. However, epitaxial interfaces are nearly ideal, whereas lithographically defined interfaces, adequate for many purposes, can be detrimental to the optical properties, in particular causing non-radiative transitions near the etched surfaces. For this reason it is desirable to form structures entirely epitaxially where possible.

Since the bandgap variations are created epitaxially, a possibility for designing structures with bandgap variations in more than one direction would be to exercise some sort of lateral control on the epitaxial growth. Any such lateral control can be termed selective area epitaxy (SAE), but it usually refers to one or more of the three main methods shown in Figure 6. One method involves etching mesas (i.e. raised ridges) into the substrate. Then the growth on top of the mesa is separated from the growth in the valley separating mesas. A second method is to use a shadow mask, in which the beams are interrupted locally by a thin shutter with holes in it. The name shadow mask arises from the fact that the molecular beam is interrupted rather like light would be, and so a "shadow" is cast on the substrate. A third method, used in this work, is to deposit a mask on the substrate, and to etch openings in the mask. If the mask materials are

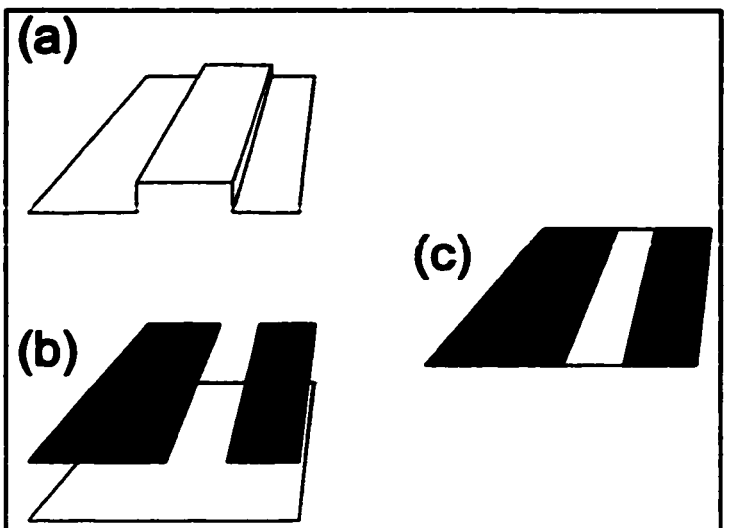


Figure 6. Three methods of selective area epitaxy. (a) ridge overgrowth (b) shadow masking (c) masked overgrowth.

chosen such that the sources cannot stick to them, then growth only occurs in the windows. The advantage of this technique is that growth occurs only where it is wanted. The disadvantage is that growth conditions must be chosen such that source materials do not stick to the mask - sometimes these requirements clash with the ideal growth conditions, leading to compromise, as will be shown.

For a survey of the field of selective area epitaxy see reference [16].

Lithography

The word "lithography" originally referred to a printing technique on paper. The essential feature that the original printing technique and the modern semiconductor processing technique share is that a pattern is transferred by means of a chemical "resist" which is deposited on the sample to be patterned.[17] In the case of this work, a chemical resist is used to protect areas from the action of an etchant. In photolithography the pattern is transferred photographically to a resist on the sample, typically using ultraviolet light which is directed through a mask. The exposed resist can then be developed. This is the standard patterning technique in the semiconductor industry. Less commercialized are the e-beam and ion-beam techniques which will also be discussed below. Though the techniques are common, the specific

application of them here is not. Therefore, in what follows, the lithographic procedures used are outlined. Although both GaAs and InP substrates were used, the processing techniques were very similar. A more detailed description of the techniques described below is found in Williams.[18]

The basic process began with a two inch diameter (industry standard size) wafer of (100) oriented semi-insulating GaAs or InP. After an initial clean in hot solvents, a 100 nm thick layer of silicon dioxide (SiO_2) was deposited by plasma enhanced chemical vapor deposition (PE-CVD). Standard lithographic techniques were then used to pattern the resist on the oxide, and then to etch through the oxide locally to create windows for growth.

Photolithography

The primary technique used to transfer a pattern was photolithography.

First a mask was designed using Autocad with all the features of interest: micron scale lines of various widths, centimetre sized openings (to compare patterned growth to conventional epitaxy), micron scale squares and more complicated shapes (Figure 7). The design was written by e-beam lithography in chrome on a glass mask by a commercial supplier.

The SiO_2 masked substrates described in the previous section were coated with a film of positive resist (Shipley 1400-26 resist) using a high speed spinner (typically 5000 r.p.m.). The resist was exposed to the *i*-line of a

mercury lamp (365 nm) shone through the glass mask pattern onto the substrate using a Karl Suss MJB-3 contact aligner. The pattern was aligned such that the long lines ran parallel to either the major or minor axes. The pattern was immersed in developer (Shipley 351 developer). A two minute, 120°C "postbake" heat treatment was used after development.[19] The pattern was then etched into the SiO_2, down to the semiconductor surface, by a 6:1 buffered oxide etch (abbreviated BOE - chemically NH_4F:HF). The time of the etch, roughly one minute long, was calibrated using a stylus profiler (Dektak) on a similarly treated test piece. Finally, the sample was chemically "stripped" (by Baker PRS1000 stripper), removing the photoresist, leaving the patterned SiO_2 mask and the substrate.

E-beam lithography

Photolithography is fundamentally limited to resolutions approximately given by the wavelength of the light in question. Although our fabrication technique reduced the minimum feature size from one defined lithographically to a much smaller size defined epitaxially, on occasion it was of interest to use patterns which were better defined than is possible optically. In particular, where μm scale square and rectangular patterns were concerned, photolithographic patterns lacked definition. In fact, squares of about 2 μm diagonal turned out nearly circular after photolithography. Since the

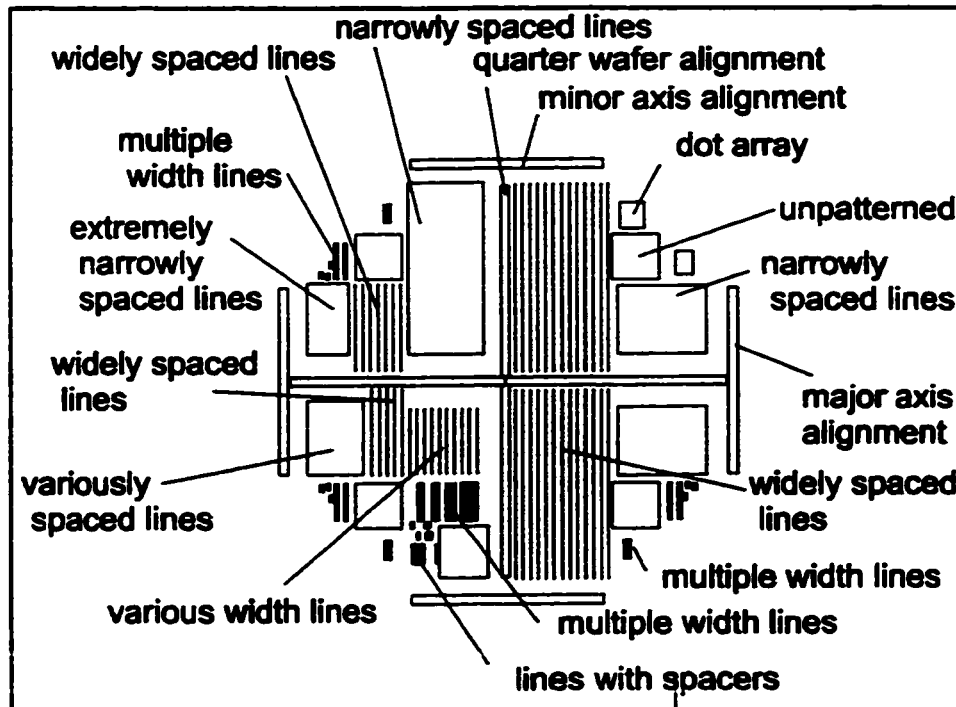


Figure 7. Photomask design.

For a 2" diameter wafer. Narrow lines are nominally 1.6 μm wide. Each quarter has an unpatterned region. Alignment markers are required to orient the mask with respect to crystalline planes.

deBrōglie wavelength of an electron accelerated by a voltage in the kV range in the vacuum is less than one hundredth of a nanometre, much smaller even than the interatomic spacing, e-beam patterns were used to eliminate diffraction effects. Practical considerations currently limit e-beam lithographic resolution to spot sizes of about 10 nm. For a general survey of e-beam techniques the reader is referred to Williams.[18] Specific details of the process outlined below are described in Feng.[20]

As with the photolithographic samples, a wafer was coated with a 100 nm thick SiO_2 layer. Next a thin (<100 nm) film of polymethylmethacrylate

(PMMA) resist was applied using a high speed (typically 5000 r.p.m.) spinner, and then baked for five minutes at 180°C to improve adhesion. PMMA is a polymer which degrades under exposure to an electron beam.

In e-beam lithography, no analog to the photomask exists. Instead, a narrow, focused beam of high energy electrons, like that used in a scanning electron microscope, traces the pattern directly onto the resist. The e-beam, focused to a nanometre scale spot size, was used to write millimetre long lines of submicron width and squares with less than a two micron diagonal. The resist was then developed by immersion in 1:3 methyl isobutyl ketone (MIBK), a solvent, and isopropyl alcohol (IPA) for 20 seconds at 20°C, then rinsed for approximately one minute with IPA. That left the SiO_2 patterns exposed for etching with BOE, to which PMMA is resistant. With the pattern etched into the mask, the remaining PMMA could be removed by rinsing the sample in acetone.

Faceting

Background

A problem with lithography, from the point of view of this project, is resolution. It is desirable to fabricate structures with lateral features in the quantum (<10 nm) range. Standard photolithography is limited to features

close to the wavelength of light (hundreds of nanometres). There is extensive research into higher resolution lithographic techniques, but none has yet matured into commercial production. Even e-beams are presently unable to pattern structures in this size regime.

Recently, however, it has been realized that the inherent anisotropy in crystal growth can be used to create patterns smaller than the etching used to create them. The first was probably Kapon, who showed that MBE growth in a V-shaped groove of a particular crystallographic orientation is, for some materials, equal at the base of the V relative to the sides of the V, and for others it is enhanced at the base of the V.[21] In that case the V can be sharpened by the first material, and a wide crescent deposited at the bottom using another, lower bandgap material. Others have performed MBE regrowth on etched mesas with various preexisting facet planes.[22,23] A similar approach is to be pursued here, except that the mesas are formed *in situ*, and the intended final structure is not the same.

The earliest work on the possibility of selective area epitaxy of faceted nanostructures of various types using masked substrates was by MOCVD.[24,25] In an important step, Arakawa used selective area MOCVD to create V-grooves *in situ*.[26] The present work was inspired by an early paper on selective area epitaxy by Asai *et al.*[27] in which facets were formed in GaAs by

MOCVD. With this observation, it seemed likely to us that selective area CBE could be used to fabricate a sharp inverted-V, and a crescent-shaped structure could be grown atop it. The mechanism of MOCVD growth is very different from CBE, however, and from these experiments conclusions should not be made regarding CBE.

The growth process used in this work, and described in more detail below, will be to fabricate mesas with (111) sidewalls and overgrow them, forming a crescent at the tip. Koshiba *et al.* [28] published such an inverted-V growth technique before us,[29] but since it was MBE, it relied on growth on mesas, and had the side effect of growth between mesas. Very early work in MOCVD showed that V-shapes of InP could be grown, but they could not be overgrown with InGaAs.[30,31] Suguira explored CBE growth of InGaAs:InP on such mesas, but found for the growth conditions he considered that there was no overgrowth of narrow band gap material.[32] In fact, for InGaAs:GaAs this appears not to be the case, [29] and as will be seen in Chapter 3 for InGaAs:InP this is not always true.[33]

Particular Faceting Technique

The basic idea for exploiting faceting (i.e. the formation and persistence of particular crystalline planes during growth) is to begin with a substrate pre-patterned for selective area epitaxy. For these experiments both GaAs and InP

substrates were used, masked 100 nm thick with silicon dioxide, with windows opened lithographically. Because of the high yield and relative ease most patterns were made photolithographically, however some samples were prepared by e-beam and ion beam lithography.

The windows were prepared in a variety of shapes, however, the most important to this study are the long, narrow windows (centimetres in length, and microns in width). These windows were aligned with the major or minor flats of the wafer, and thus aligned to the [011] or [01$\bar{1}$] crystallographic direction. In order to correctly align these patterns it is crucial to be aware of wafer conventions. (See Appendix B, page 172). The mask breaks the translational symmetry of the epitaxial surface perpendicular to the long axis of the window.

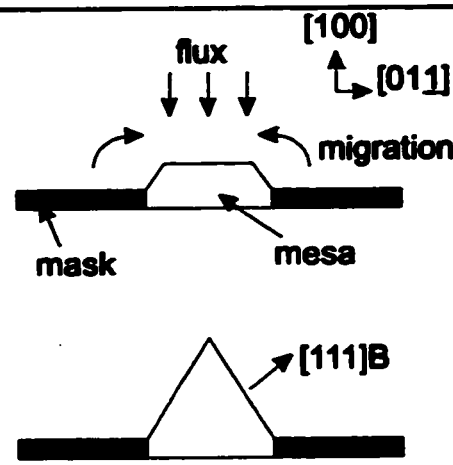


Figure 8. Schematic of growth process

The initial cross section of a sample is shown at the top. Mesa growth due to the incident flux occurs in the unmasked area. Material migrates off slow growth planes, forming facets such as the (111)B planes, shown at the bottom.

Once epitaxial growth begins, the presence of the mask causes normally buried crystallographic planes to be exposed. In the case of [011] oriented

windows, group V rich B-planes are exposed. For [01$\bar{1}$] oriented windows, group III rich A-planes are exposed. The specific planes will be discussed in Chapter 2, the important point here is that different crystallographic planes are exposed. Because of their different crystalline structures, they can be expected to grow at different rates, and slow growing planes tend to persist. If there exist slower growth planes parallel to the alignment of the mask, but not in the plane of the mask, the tendency will be for the sidewalls to develop into these slow growth planes, and for these planes to grow inward at the expense of the top surface. The result is that the flat top surface is made narrower than the initial window width. A schematic of this growth process is shown in Figure 8.

This effect allows one to use conventional lithography to define a mask, and to narrow the mask *in situ*. Taking this to its extreme limit, it is possible to try to make the top surface small enough that, if topped by a narrow bandgap material, geometrical quantum effects might be observed.

Other Fabrication Techniques

Above, one method of fabricating nanostructures has been described. The specific technique used in this work was presented, and it was placed in the context of similar techniques which rely on faceting.

To provide further context, in this section some of the more important techniques which have been investigated to date are discussed. A complete review of the effort to fabricate semiconductor nanostructures would be lengthy, and given the immaturity of the field, necessarily incomplete. It should be emphasized that every technique so far has serious drawbacks, and the ideal technique for the fabrication of quantum wires or quantum dots has probably not yet been devised.

The ideal nanostructure would have strong confinement (>100 meV) to a line in the case of a quantum wire, or to a point in the case of a quantum dot. The width in the confining directions should be very small (<10 nm). Quantized levels would then be well spaced, and the density of states would show its characteristic dimensional dependence.[34] For technological applications they should be easily reproduced, so that a wafer can be filled with structures, and so that the structures are highly uniform. Interfaces between barrier and active layers should be of the highest quality.

One of the most successful techniques so far has been to grow a two dimensional electron gas, in which changes in doping during epitaxial growth are used produce a sheet of carriers. These carriers are then confined laterally using electrostatic gates.[35,36] Heterojunctions, in which a composition change changes the bandgap, are also often used. However, lateral confinement is

weak (~1 meV), and the structures are large (~1 μm^2) so that milliKelvin cryostats must be used to observe confinement effects. Confinement effects are observed in conductivity experiments, but they are for the most part too weak to be seen optically.

Among purely epitaxial techniques, one of the first proposals was a T-shaped intersecting quantum well.[37] Quantum wire confinement to the intersecting area is obtained from geometrical effects which are discussed in Chapter 4. Such structures have been fabricated by growing a quantum well, cleaving *in situ*, and regrowing on the cleaved surface.[38] Interesting optical properties are observed for these structures, but it seems a very impractical fabrication technique.

Another possibility which has been investigated has been to use the increased probability of incorporation near step edges in epitaxy to gain some lateral control over growth on terraces.[39] (Such steps and terraces are obtained, for example, when a substrate is polished to an angle slightly away from the (100) plane.) Some difficulties include the roughness of the structures, and the experimental results so far are controversial.

Recently, using strain to disrupt layer by layer epitaxial growth and produce ~10 nm diameter self-assembled islands of low bandgap material has become popular.[40,41,42] This technique is attractive because the growth process

is very simple compared to other techniques, and the optical properties are quite good. The disadvantages of the growth technique at present is that size and composition fluctuations are very large. Also the distribution of such structures is random, and the density is not independent of the structural parameters.

By comparison, since the faceting technique used here creates the structure entirely *in situ* it does produce high quality interfaces. It allows for many wires to be produced across the sample, with any desired location. The lateral scale of confinement for samples investigated here will be seen to approach the quantum size regime. The confinement could be strong if the thickness or composition of the overgrown well is changed significantly. In the following chapters the fabrication of real inverted V-mesas will be demonstrated, and their properties described.

Conclusion

Chemical beam epitaxy, the growth technique used in this work, has been described and placed into the context of other, more conventional growth techniques. In themselves, epitaxial growth techniques allow control of deposition down to the atomic layer, but do not provide any lateral control. Selective area epitaxy was developed in response to scientific and technical

interest in laterally patterned semiconductor structures. In conventional selective area epitaxy an upper bound on feature size of ~0.1 micron is imposed by the current limitations of lithographic technologies. However, by exploiting crystalline faceting it is possible to decrease lateral sizes from lithographic dimensions down to nanometre scales.

The first step towards realizing this type of structure is to study faceting behavior. In the next chapter the fabrication of real gallium arsenide based structures will be described. It will begin with a description of basic issues for growth, discuss structural issues of faceting and roughness, and conclude with a description of optical properties.

Chapter 2

REAL GaAs BASED STRUCTURES

Introduction

If nanostructures could be fabricated routinely today there can be little doubt that they would find an eager market, both in academe and in the electronics industry. There are, however, many technological hurdles associated with fabrication on these scales. In this chapter the issues related to the faceted selective area growth technique will be described.

Using gallium arsenide substrates, the growth conditions required for optimized facet formation will be presented. A simple model is used to rationalize the faceting behavior. Once facets can be produced, it becomes possible to grow inverted V-shaped structures. After it is shown that such structures can be produced, attention will be turned toward characterization. Photoluminescence spectra will be related to growth dynamics and structural properties.

Growth Conditions

InGaAs:GaAs mesa structures grown by CBE were found to be very sensitive to growth conditions, in a number of different ways. First, the selectivity of the growth was strongly dependent on temperature. Second, the rate of incorporation of indium in InGaAs quantum well layers was a strong function of temperature. Finally, the sidewall shapes were very sensitive to V/III ratio. This will be discussed in detail below.

Selectivity

There is extensive literature on selective area epitaxy by MOCVD[43], and good references are available for CBE.[44,45] Although this work does not concern selectivity *per se*, the first stage in this type of growth is to establish selectivity, and so the most important considerations will be highlighted here. First and foremost we found the selectivity conditions to be very sensitive to the preparation of the surface silicon dioxide. As an example, a poorly developed resist was stripped and repatterned before etching the silicon dioxide. This seemingly inconsequential action resulted in the first pattern showing up after growth as micron scale GaAs non-selective globules atop the silicon dioxide. More generally, a less than ideal process can raise the selectivity temperature (see below) for an entire sample to a prohibitively high temperature, or worse, make selective growth impossible. The facet growth

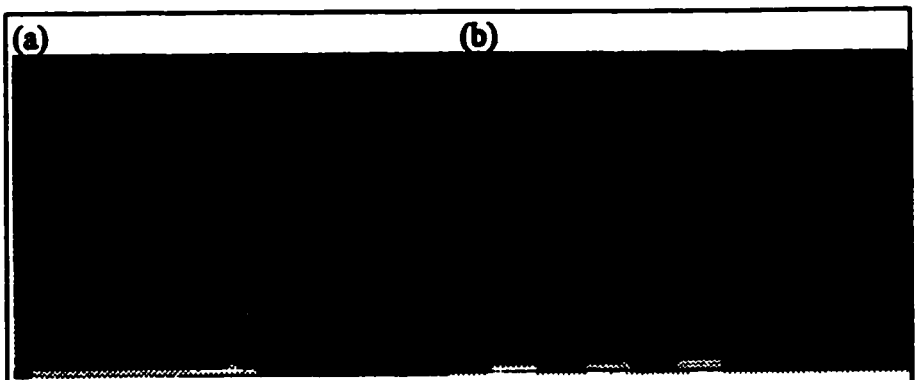


Figure 9. Selective area epitaxy

Electron micrographs of selective area growth. In (a) epitaxial growth of GaAs has occurred along the dark line to the right of the micrograph. The rest of the surface, masked with silicon dioxide, shows submicron specks of non-selective growth. A cleaner surface or higher growth temperature will show better selectivity. For example, (b) shows perfect selectivity. Again, on the right there is an unmasked GaAs mesa. The rest of the sample is masked by silicon dioxide. No nucleation has occurred on the mask. The white rectangular marker at the bottom of (a) is five microns. The white markers in (b) are each one micron in length.

process is tolerant to some non-selectivity, a fact that will be exploited later. However, if there is too much non-selective growth on the mask it can interfere with growth in the windows.

For a given growth rate, and to a lesser extent V/III ratio, selectivity occurs above a certain temperature (the "selectivity temperature").[45] This is because desorption of an element from the silicon dioxide mask in its simplest description is a temperature activated chemical process. (Desorption on the unmasked part of the substrate is weaker because the exposed semiconductor surface catalyzes incorporation into the surface.) In our CBE system we found the selectivity temperature for GaAs growth with TEG to be ~540 °C.

The effects of growing just below the selectivity temperature and just above it are shown in Figure 9. At too low a temperature, sub-micron, irregularly shaped clusters are deposited on the mask. This growth can be said to be partially non-selective or phase-selective.[44] Such non-selectivity is nucleated by the group III element. Consequently reducing the growth rate reduces the selectivity temperature.[45] (The growth rate is usually proportional to group III flux.)

Indium Incorporation

Indium is a group III element, lying beneath gallium in the periodic table. As such, it is a larger atom, and its outer three electrons are more weakly bound. In combination with arsenic it can form a direct gap zinc-blende semiconductor structured crystal, having a larger lattice constant and smaller bandgap than gallium arsenide. As a narrower gap material it can be used to produce quantum confined structures on GaAs substrate. However, there are two considerations which limit its usefulness.

One issue is strain. The lattice constant of InAs is about 7% larger than that of GaAs, a direct consequence of the larger size of the indium atom. This much strain implies a critical thickness of a little over one monolayer, at typical growth temperatures, for which the material may be grown without relaxing and creating many defects. Using a ternary (i.e. $In_xGa_{1-x}As$) the strain

is reduced, resulting in a larger critical thickness. Quantum size effects can be observed for structures in InGaAs:GaAs of less than about 10 nm. There is, as a result, a practical limit of ~20% indium concentration, above which the critical thickness is too small to produce good quality heterostructures.

Another issue, particularly important for selective area growth, is that the weaker bond strength of indium implies that it desorbs readily at high temperatures. In MOCVD and MBE, this has the well known consequence of indium segregation, in which the indium atoms do not incorporate readily into the lattice, but instead float on the growth front.[46]

For CBE growth at substrate temperatures between about 400°C to 550°C, the incorporation of indium into InGaAs layers on (100) GaAs is efficient.[47,48] As the temperature is increased, desorption of indium becomes increasingly important, until, at about 600°C almost no indium is incorporated into the quantum well layers. As will be shown from the PL, indium desorption from the (100) plane puts an upper bound (≈570°C) on the usable temperature range for growing InGaAs wire structures. However, this usable temperature range is extended by patterning, because indium migrates to the (100) surface from the (111)B at a rate comparable to the desorption of indium from the (100) as shown by the PL, below (Substrate Temperature Dependence, p68).

There is an obvious conflict then, between the low temperature required to prevent indium from segregating, or worse desorbing, and the high temperature required to ensure selectivity, specifically, to ensure that gallium desorbs from the mask. This problem is inherent to the particular material system. One way to manage the problem is to recognize that the growth typically consists of a very thick GaAs layer (~1 μm), one or more thin (<10 nm) InGaAs quantum wells, and a cap of GaAs (~100 nm). Then, a high growth temperature can be used for most of the growth, with a brief drop to low temperature, and slight loss of selectivity for the growth of InGaAs. Most of the GaAs based InGaAs samples grown as part of this work exploited this technique.

V/III Ratio

Variation of the V/III ratio of the supply gases for conventional (100) growth usually does not change the composition of the crystal significantly, because group III and group V elements usually settle in sites on different sub-lattices. It does affect impurity incorporation and surface morphology. However, in selective area growth, planes other than the (100) are exposed, and the V/III ratio is very important in selecting which facet planes grow more quickly than others, and therefore which facet planes persist. The V/III ratio is such an important parameter that a substantial part of this chapter will

be devoted to describing the effects of the variation, and trying to explain some of that variation.

Detailed Growth Conditions

Substrates were two inch diameter *n*-type chromium doped (100)±0.2° oriented GaAs wafers patterned with SiO_2 as described in Chapter 1, p20. Between patterning and introduction to the CBE machine, no cleaning was performed. The substrates were cleaved into quarters, and mounted on two inch silicon wafers by melting indium on the silicon wafer, placing the sample on the silicon wafer and allowing it to cool ("indium mounting"). The silicon wafers were then held by tungsten clips to hollow molybdenum blocks ("molyblocks"). The blocks were placed into a vacuum chamber. After reaching UHV conditions ($<10^{-8}$ torr) the sample was heated to 200°C for one hour to desorb any remaining organic molecules, such as those that may still remain after etching the masks. The sample was then introduced to the growth chamber. A second deoxidation step was performed, with the sample heated to 620°C for 10 minutes under a high flux of cracked arsine (30.0% full scale). The flow of cracked arsine prevents the surface quality from deteriorating due to desorption of gallium or arsenic at this elevated temperature. The growth could then proceed.

Table 1. Typical growth conditions for InGaAs:GaAs mesas.

Description	*Duration*	*TMI*	*TEG*	*AsH_3*[*]	*T(°C)*
GaAs buffer	65 min	–	40	20	565
T ramp	2 min	–	40	20	ramp
InGaAs well	20 sec	10	40	20	530
interrupt	10 sec	–	–	20	530
T ramp	2 min	–	40	20	ramp
GaAs cap	5 min	–	40	20	565

This represents a baseline only. Fluxes are in percent of full scale. To convert fluxes into flow rates see Appendix A. Dashes indicate no flux. Substrate temperature did not always follow a ramp and ranged from 510°C to 575°C. TMI flux actually ranged from 5 to 20%, depending on the desired composition. Arsine flow rates were varied from 5% to 30%, depending on desired V/III ratio, ranging from ~0.34 to ~2.0.

Substrates were rotated with a period of several seconds to ensure beam uniformity on average. Typical growth conditions for GaAs and InGaAs on GaAs crescent structures are tabulated in Table 1.

[*] Arsine was delivered by the As1 line (see Appendix A)

Figure 10. Faceted growth on a square window.

Indium phosphide was grown on a 1.0 μm×1.0 μm square opening in a silicon dioxide mask patterned by e-beam lithography. The growth exposes crystalline facets that were not expressed on the (100) substrate. In the center there are planes tilted away from the (100). At top, bottom, left and right are (111) planes. Four other planes are seen on the diagonals. The marker represents one micron. (All other SEM micrographs in this chapter are of GaAs samples.)

Structural Characterization I: Faceting

One of the most beautiful aspects of crystalline materials is that the atomic scale properties are reflected in the macroscopic form that the material ultimately takes. For example, indium phosphide grown by CBE on a flat (100) substrate with a square mask forms a variety of facets (Figure 10). (Gallium arsenide would also show facets, however, no square e-beam masks

were used for GaAs growth in this project.) Because chemical beam epitaxy is highly non-equilibrium, the specific shape the crystal takes depends greatly on the initial conditions and on the process by which it was formed.

In this work, windows of various micron scale widths and macroscopic lengths were etched into silicon dioxide. Layers of a micron or more in thickness were grown in these windows by chemical beam epitaxy. After growth the lines were cleaved in cross section and placed into a scanning electron microscope (SEM), usually without staining. The resulting growth habit and morphology were photographed. The facet structure was observed to vary dramatically with mask orientation and with V/III ratio. The resulting cross-sections are presented below (Figure 11).

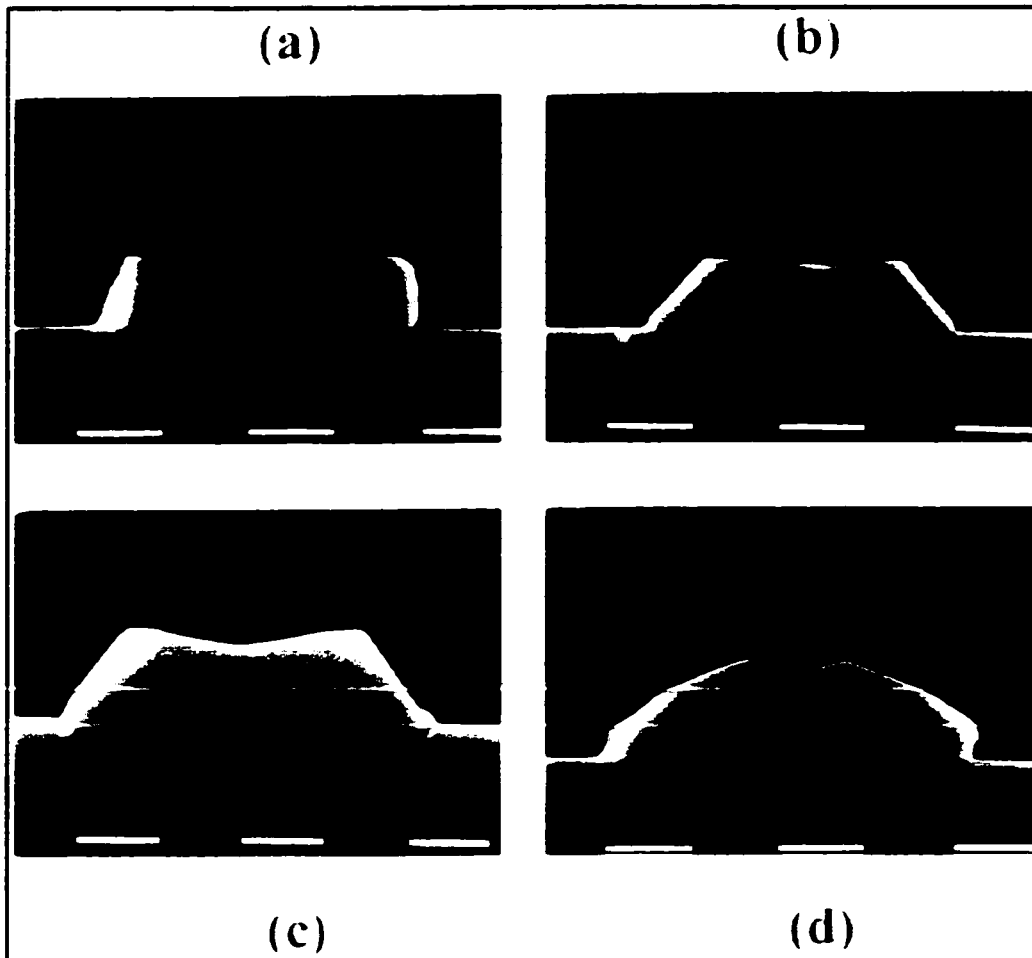


Figure 11. Mesa cross-section as a function of V/III ratio

Scanning electron micrographs of [011] and [01$\bar{1}$] oriented mesas for different V/III ratios at high (570°C) temperatures: (a) and (b) Low V/III ratio (≈ 0.5), (c) and (d) High V/III ratio (≈ 2.0). (a) and (c) [011] oriented mesas, (b) and (d) [01$\bar{1}$] oriented mesas. White markers represent one micron.

Mask Orientation and V/III Ratio

Referring to Figure 11, the effects of varying the V/III ratio and mask orientation can be seen. The V/III ratios used for conventional planar growth of GaAs are close to unity. Low V/III ratio (≈ 0.5), [011] oriented mesas (Figure 11a) show vertical sidewalls with very little lateral growth. The

corners are slightly rounded, but otherwise the surfaces are flat. In the orthogonal mask orientation, low V/III ratio, [01$\bar{1}$] mesas (Figure 11b) show (111)A sidewalls, with slight overgrowth in the lateral direction. For high V/III ratio ($\approx$ 2.0), the growth is quite different. The [011] oriented mesas (Figure 11c) show flat (111)B oriented sidewalls. The (100) surface is no longer flat. Wider mesas show that the (100) surface is flat in the middle, but curves upward at the edges, with a length scale of nearly one micron. As will be shown in the next chapter, this sort of curvature is characteristic of the diffusion of material, in this case from the slow growth (111)B plane to the (100) plane. The [01$\bar{1}$] orientation (Figure 11d) shows irregular growth in the lateral direction. Among the many facets expressed, the roughly (311) oriented facet is the smoothest. The (100) is also curved, but with a shorter length scale than the orthogonal mesas. These samples were grown at high temperature (570°C). The growth rate, as measured at the center of the large mesas, was slightly less for the low V/III ratio mesas, despite the fact that the TEG flux was held fixed. Material appears to have migrated more on the high V/III ratio samples.

Mask Width

There are, as a result of the variation in facet expression with respect to the growth conditions, at least two apparent possibilities for the CBE growth

of wires. One is to grow on (111)A faces with low V/III ratio, the other to grow on (111)B faces with high V/III ratio. We have chosen the latter, mainly because, as will be shown below, the morphology of (111)B facets was smoother than (111)A facets, and because the inverted V-shape formed under growth conditions closer to the normal optimum growth conditions for (100) InGaAs:GaAs. The (111)A faceting growth conditions are unusual in that the V/III ratio is rather extreme and such faceting was observed in conjunction with non-negligible desorption (c.f. the growth rate reduction discussed in the previous section).

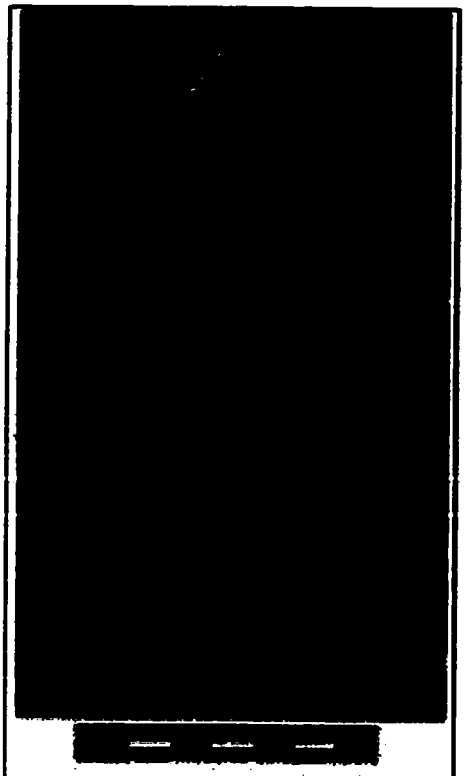


Figure 12. Various width [011] GaAs mesas in cross-section.

From bottom to top, (111)B sidewalls come closer together, forming a V for the narrowest mesas. Note the curved profile on the top (100) near the (111)B sidewalls. Irregular globular structures are non-selective growth. The white rectangular markers represent one micron.

Cross-sectional SEM photographs of [011] oriented mesas of different widths from ~2 μm to 10 μm are shown in Figure 12. Clearly defined (111)B facets are formed on all the mesas. The widest mesas show normal (100) growth at the center, far from the sidewalls.

However, near the sidewalls there is a slight growth rate enhancement, and a fairly smooth decay to the usual (100) growth. It will be shown in Chapter 3 that this shape arises naturally from the diffusion of group III species in particular, from the (111)B facet to the (100) surface.

An important point here is that the distance along which the surface deviates from (100) before becoming a (111)B facet is only ~0.8 μm. This curvature will be seen to extend further for similarly prepared indium phosphide structures.

Moving up Figure 12 from wide mesas at the bottom to narrow structures at the top, it is seen that the curved growth fronts associated with diffusion merge when the mesa becomes wide enough, increasing the growth rate locally at the center of the mesa. Eventually, however, the mesa becomes so narrow that the (111)B facets overlap, producing a sharp V. These will be

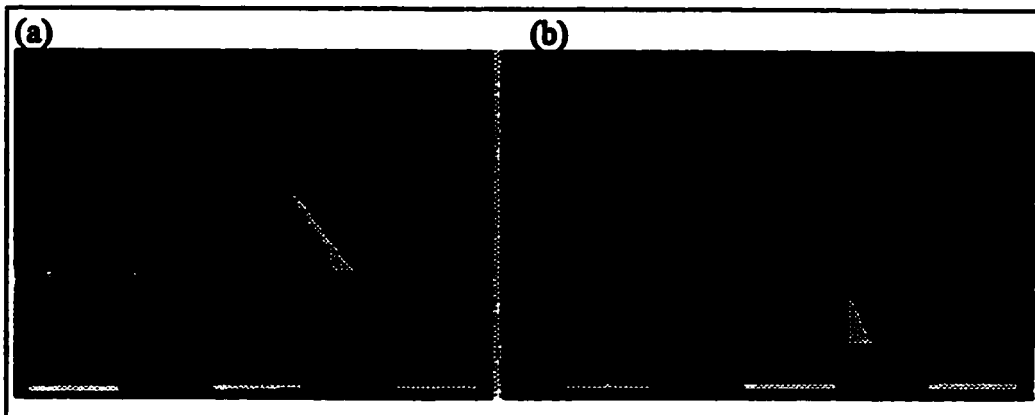


Figure 13. Conformal growth of completed mesas.
Growth for (a) 45 minutes, (b) 85 minutes. White markers represent one micron.

referred to as V-mesa structures. The V is so sharp (in fact the measurement of the sharpness is limited by the instrumental resolution here) that if there is overgrowth on this V-shape it is a good candidate for the fabrication of quantum wires.

Figure 13 shows one growth which has lasted just long enough to complete a V, and another which has lasted much longer. Overgrowth is conformal, meaning that it is uniform across the surface.

Tip Sharpness

If one is to fabricate true quantum structures, it is important to determine the sharpness of the tip during growth. SEM is inadequate for this purpose, so transmission electron microscopy (TEM) was used to examine a mesa in cross-section (Figure 14).

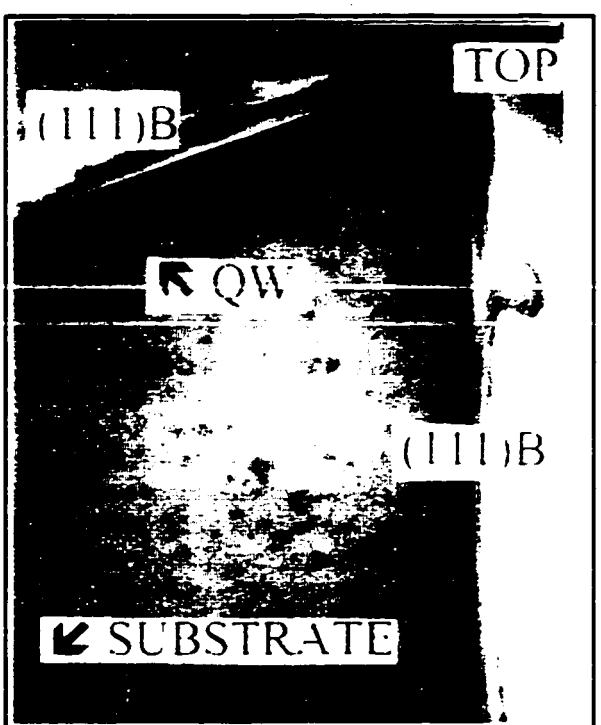


Figure 14. Transmission electron micrograph of the tip of a faceted mesa structure.

The patterned substrate is lower left and not seen in the photograph. A side quantum well, not visible in reproduction here, runs parallel to the top (111)B surface, at the location marked "QW". The tip is curved, with a radius of curvature of ~25 nm.

TEM measurements show that the tip is curved, with a radius of curvature of approximately 25 nm - rather broad to observe quantum confinement effects (except at very low temperatures), but much smaller than the original mask, and as small or smaller than can be prepared by other methods. The tip shown here is not exactly circular, but rather is itself faceted on very small scales, showing what are probably (311) planes and a very nearly (100) segment which is possibly less than 10 nm - leaving the possibility of the observation of quantum confinement effects.

Limits of Faceting

To produce quantum structures the mesa tip needs to be as sharp as possible. Some factors required for a sharp tip are evident. First, the initial mask roughness must be minimized, since slight fluctuations can become large relative to structure size. Perhaps most importantly, the diffusion process, which will be described in more detail later, must be controlled. Unlike deposition, which is very well controlled, migration along a facet and between facets is strongly dependent on the growth conditions, and differs for each plane. This could be complicated by desorption as well - not only from the facets but from the narrow regions joining them.

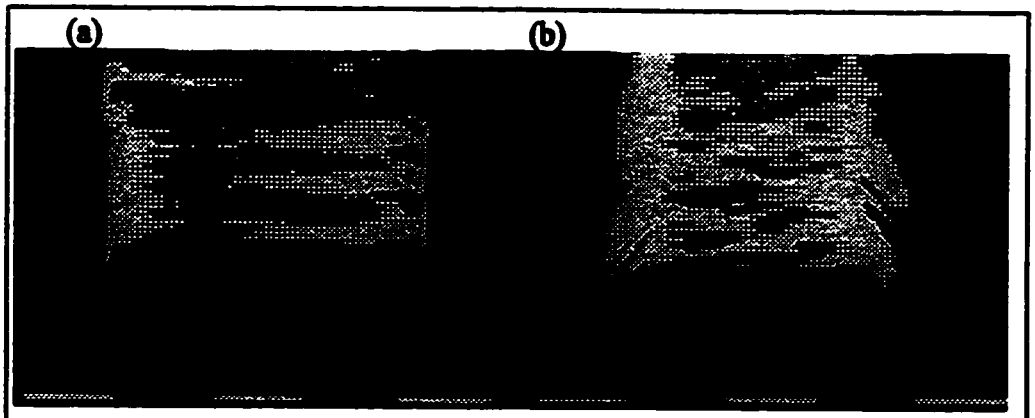


Figure 15. Surface defects at low V/III

The top surface is viewed from a steep angle to bring out the roughness. For GaAs at low V/III the [011] oriented mesas (a) show [01$\bar{1}$] directed (i.e. perpendicular) striations along the top (100) surface. The [01$\bar{1}$] oriented mesas (b), show rounded defects. White markers are one micron.

Structural Characterization II: Morphology

Variation in the morphology of the various facet surfaces, some but not all of which was systematic, was seen with variation in the growth conditions. Most work was focused on fabricating a sharp V shape as seen in cross-section, but these variations in morphology can involve dramatic departures from the intended structure that are not obvious in cross-section. The roughness observed can be quite complicated, involving deviations from a flat surface that can only be quantified in three dimensions. This section, is intended simply as a catalog of some of the morphological peculiarities observed.

Wide mesas

As a first example, the (100) surface for wide mesas shows a different sort of roughness depending on stripe orientation, even for simultaneously grown samples. At low V/III ratio (~0.5), for [011] oriented mesas, the surface is striated in the transverse direction, as can be easily seen at grazing incidence (Figure 15). For the same growth conditions, [01$\bar{1}$] mesas show circular defects on a micron scale, reminiscent of the oval defects common in MBE grown (100) GaAs. That these samples were grown simultaneously and show such different characteristic defects means that they must be a result of an interaction between the top surface and the sidewalls. Interestingly, the same sort of defects, though less pronounced, appear in InP based samples as well.

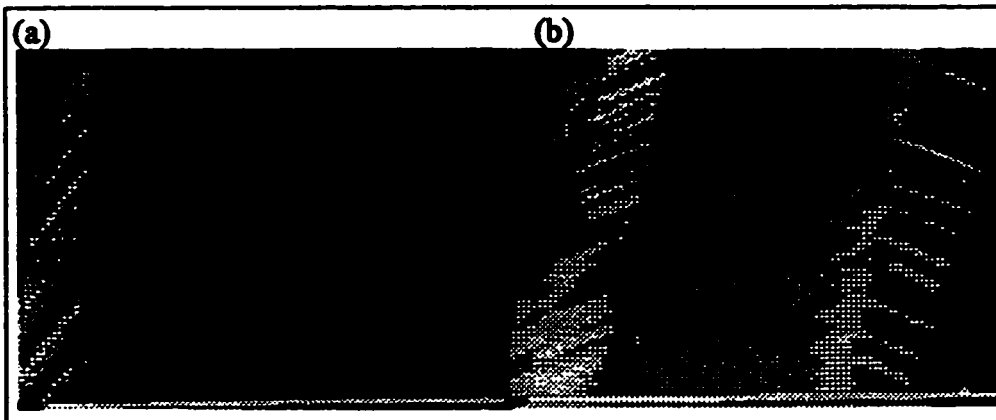


Figure 16. Surface defects at high V/III
(a) [011] oriented and (b) [01$\bar{1}$] mesas. The mesas are viewed at a shallow angle of incidence with the (100) plane at the top, center. The white markers indicate 10 μm.

Figure 17. Surface morphology of narrow mesas

the wider mesas, they are marred by a sub-micron roughness corresponding to a slight tilt of planes in a longitudinal direction.

Narrow [011] mesas

If wire structures are to be fabricated from [011] oriented mesas using (111)B planes, the roughness must be minimized in these structures. To that end, the kinds of structural problems that arise specifically for these narrow mesas are catalogued.

First, even on the smoothest mesa, triangular defects the width of the facet itself were observed (Figure 18). These are strikingly similar to the pyramidal defects observed by others under certain conditions for growth on planar (111)B substrates.[49]

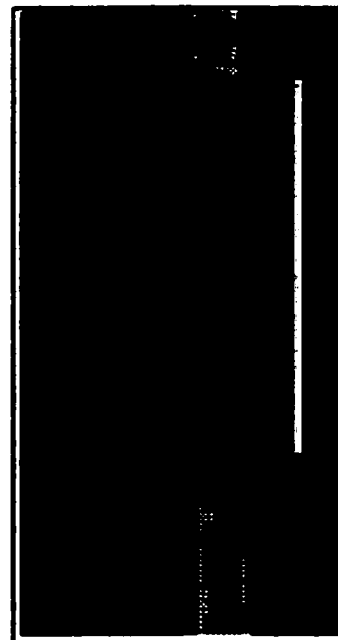


Figure 18. Pyramidal defects in an [011] mesa

An electron micrograph of a narrow [011] directed mesa viewed from the [100] growth direction. The sidewalls are (111)B planes, but they show triangular defects. The marker represents 10 μm.

Another problem encountered while growing V-mesas was the occasional asymmetry seen in the shape of the V (Figure 19). Sometimes one (111)B sidewall would grow at the expense of the other. If the winning side alternates over the length of the structure the mesa becomes irregular. Occasionally one

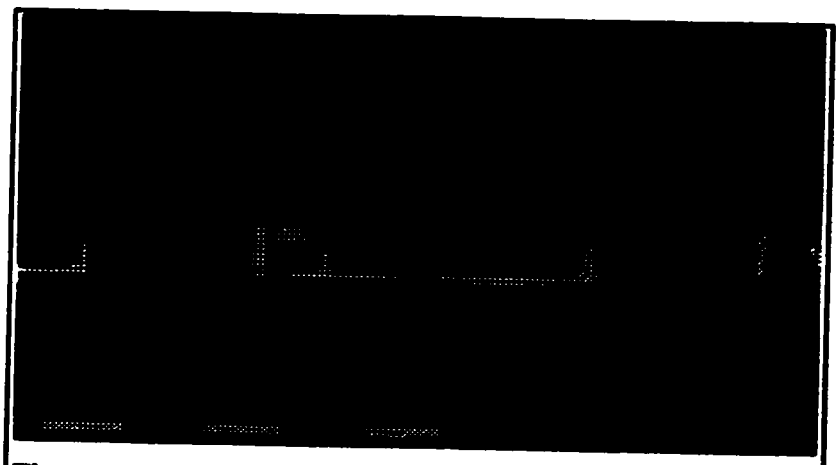


Figure 19. Tilted V mesas

The mesa on the left shows one side of the V winning out over the other, resulting in a non-symmetric shape. The mesa on the right remains symmetric, however. White markers represent 1 μm.

side was seen to win consistently over another. Clearly if the flux is not normal to the (100) surface and the sample is not rotated a shadowing effect will result in one side growing faster. However, some samples show this irregularity while others do not. This work cannot provide a systematic explanation of the asymmetry.

The overall pattern of V-mesas in GaAs typically involved abrupt shifts in the lateral position of the vertex of the mesa (Figure 20). The roughness of these surfaces proved to be unpredictable.

Figure 20. Roughness of narrow [011] mesas

In the [011] directed mesas the (111)B sidewalls have come together to a form a V. However, the (111)B sidewalls show some variations in that they are angled back and forth, and hence the line where the two (111)B planes join together to shift abruptly on the scale of ~1 μm. The white marker represents 10 μm.

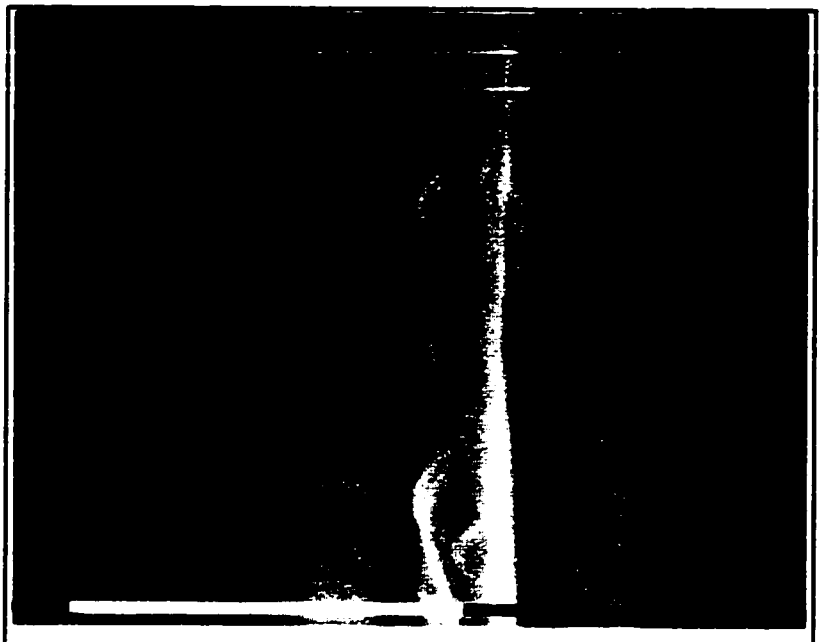


Figure 21. Wide mesas vs. narrow mesas

Two very wide mesas separated by a narrow space and a narrow mesa. The wide mesas show smooth (111)B sidewalls, while the narrow V-mesa shows irregular growth. The white marker represents 10 μm.

Wide mesas vs. narrow mesas

Control of the migration processes is necessary if roughness is to be eliminated in narrow mesas. Wide mesas can show excellent morphology while their narrow counterparts may be very rough. Figure 21 underscores this problem. The central V-shaped mesa shows irregular growth which the wide mesa does not. (Generally, most V-mesas were not this rough.) It seems likely that this problem is avoided for a very wide mesa, because there is always a wide (100) surface which can act as a reservoir for diffusion and the diffusion conditions are very similar over time. For a narrow mesa on the other hand, conditions are rapidly changing on the top surface as it becomes narrower, ultimately switching to essentially (111)B growth only. Of course, given that wide mesas are typically so smooth, the fabrication technique could be altered such that the intent would be to fabricate a narrow wire at the intersection of the wide facet such as the (100) plane with a sidewall such as the (111)B. But if the technique depends on a wide facet, it would seem impossible to ever produce a high density of wires.

All of these types of roughness might make one pessimistic of the prospects of producing quantum wires, but there are still a few reasons for optimism. First, some samples showed less roughness than others, moreover, some regions of a given sample would be smoother than others, suggesting

that with better control of the process, smooth wires could be fabricated. In principle, the structure need only be smooth on a very small scale for quantum effects to be observed locally. Also, there is room for improvement in the process - we have done little to try to improve the initial patterning step. Some possibilities include aligning to a cleaved edge instead of a wafer flat, since a wafer flat is up to 1° off the intended crystallographic plane, or, more importantly ensuring smoother edges by using a higher quality mask or shorter wavelength photolithography. But the final, and perhaps strongest reason for optimism is that many of these types of roughness become much less important in other material systems, most notably indium phosphide, which is readily accessible to CBE and will be investigated in Chapter 3.

Phenomenological Theory of Faceting

A phenomenological approach can provide an explanation for the formation of facets. To try to understand faceting as a function of V/III ratio an approach used by Sugiura[50], which follows generally from earlier work by Asai[51], will be applied and generalized. The basic observation is that the growth rate tends to be proportional to the bond density on a surface. The object then is to calculate that bond density for the surface in question, and

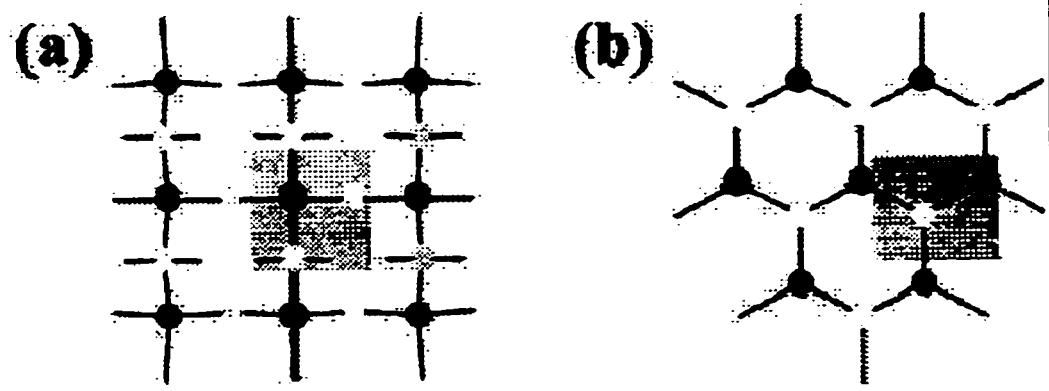


Figure 22. High symmetry crystallographic surfaces.

(a) The (100) surface. (b) the (111) surface. The surface is (111)A if the terminating plane (i.e. top layer) is gallium, (111)B if it is arsenic. The surface unit cell, which is orientation dependent, is highlighted in gray.

assume that facets form because the bond density and resulting growth rate are low on that surface.

Facets at high V/III

The work of Sugiura *et al.*[50] explains the facet growth of InP at high V/III ratios (~2). At high V/III ratios, it is assumed that surfaces are terminated by the group V element, in this case arsenic. Unreconstructed surfaces are considered throughout, for simplicity. The growth rate is assumed to be proportional to the bond density. To determine the bond density, the number of bonds available per surface unit cell is counted, and divided by the area of that unit cell.

There are three special crystallographic planes which require consideration. The first is the (100) surface, illustrated in Figure 22a. Each

surface unit cell has two atoms, one gallium and one arsenic. The arsenic has two free bonds available for new gallium atoms. The gallium has no unsatisfied bonds available, being completely surrounded by arsenic.

Next consider the (111)B plane. It is arsenic terminated, and so each arsenic atom has only one free bond available for gallium. There is one arsenic atom per surface unit cell, so there is one bond available for gallium for each unit cell. (Note that the unit cells in the (100) and (111) planes are of different dimensions, though.)

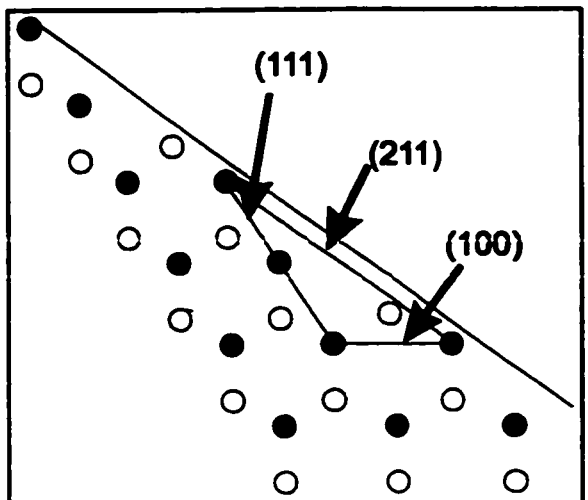


Figure 23. Constructing all (*hkk*) planes.

Any (*hkk*) plane between (100) and (111) can be constructed of (111) steps and (100) terraces. Here, schematically, a (211) plane is constructed of a (100) terrace and two (111) steps.

Finally consider the (111)A plane. It is normally gallium terminated. However, considering only high V/III ratios, the model requires all surface dangling bonds to be arsenic related bonds, so an arsenic atom is attached to each of these surface gallium atoms. In that case there are three remaining free bonds available for gallium on each (111)A unit cell.

Other planes

This model can be extended to all planes between (111) and (100) by plotting the positions of atoms for the unreconstructed (*hkk*) planes with *h,k* integers, and *h>k* (Figure 23). Such a construction was used by Yamada *et al.*[52] to explain the relative growth rates of certain crystallographic planes in migration enhanced epitaxy. Any such plane can be made of (111) steps and (100) terraces. For example, the (311) surface can be made up of one (111) step for each (100) terrace. More generally, it can be shown[†] that an (*hkk*) plane can be constructed of *m* (100) terraces and *p* (111) steps according to the relation

$$\frac{m}{p} = \frac{1}{2}\left(\frac{h}{k} - 1\right). \qquad (1)$$

A slight complexity is introduced here, since the minimum number of steps and terraces should be used. This results in two cases depending on whether *h-k* is even or odd. If *h-k* is even, $m = (h-k)/2$ and $p = k$, however, for *h-k* odd $m = h-k$, $p = 2k$.

Each segment contributes its bonds to the surface. If α represents the number of bonds available for gallium on a (100) terrace, and β represents the number on a (111) step, then the number of bonds per 2D unit cell, B, is

[†] the author's derivation

$$B = p\left(\alpha\frac{n}{p}+\beta\right). \quad (2)$$

In the case of high V/III ratio discussed above, for A oriented surfaces α=2, β=3, and for B oriented surfaces α=2, β=1.

Finally the area of the surface unit cell can be calculated as the volume of a unit cell, divided by the thickness of a monolayer. The monolayer thickness can be calculated by finding the lattice vector which minimizes the dot product of itself and the surface normal. That dot product can then be defined as a monolayer thickness. For an (*hkk*) plane where h and k have been reduced to have no common factor greater than one, this results in a thickness t_{hkk},

$$t_{hkk} = \frac{a}{\mu_{h-k}\sqrt{h^2+2k^2}}. \quad (3)$$

where μ_{h-k}=1 for h-k even, μ_{h-k}=2 for h-k odd, and a is the lattice constant.

Defining the angle θ as the angle away from the (100) so that

$$\cos\theta = \frac{\frac{h}{k}}{\sqrt{\left(\frac{h}{k}\right)^2+2}}, \quad (4)$$

the bond density as a function of angle, is given by

$$\frac{Bt}{V} = \frac{1}{a^2}\left[2\alpha\cos\theta+2\sqrt{2}\left(\beta-\frac{\alpha}{2}\right)\sin\theta\right]. \quad (5)$$

The result is plotted in Figure 24, for both A and B oriented planes, for both high V/III ratio, as above, and low V/III ratio, described below.

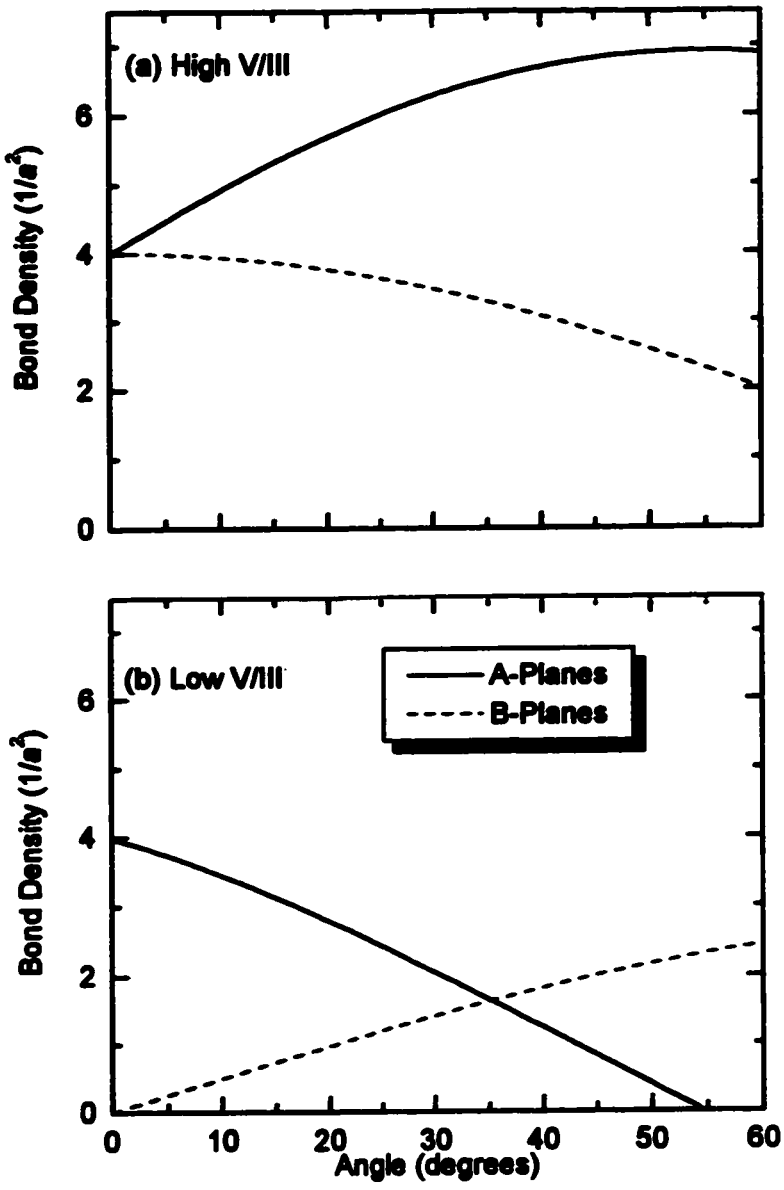


Figure 24. Bond density as a function of angle

The bond density for A and B oriented (*hkk*) planes, between (100) and (111) for (a) high V/III ratio and (b) low V/III ratio. The low V/III bond densities are different at zero angle because the B-plane density has been calculated for a lower V/III ratio.

Low V/III

It is necessary to extend the model to low V/III ratio surfaces. As the V/III ratio is lowered not all arsenic sites will be filled and weakly bound arsenic atoms will be the most likely to be emptied. This can be modeled simply by eliminating all arsenic atoms bound by only one bond, and counting the remaining bonds available for gallium. For very low V/III ratios, all arsenic atoms bound by two bonds can be eliminated, counting the bonds available for gallium. Of course the arsenic still needs to be incorporated, but the assumption is that the incorporation of gallium becomes a rate limiting step[‡], and arsenic only participates in the way that it affects the incorporation of gallium. Site occupation of arsenic varies in a continuous way from nearly complete occupation through to nearly complete absence as the V/III ratio is varied. This has been modeled by Asai[51] who included temperature activation and V/III ratio for specific crystallographic directions. Here these effects will not be included explicitly, and only the above simplified cases of very high and low V/III ratio will be considered. For low V/III ratio then, the only modification in the growth rate is to set β=0 for A oriented planes. For very low V/III ratio both α and β are zero for the A oriented planes, while α becomes zero for B oriented planes, but β is unchanged (i.e. is one). The

[‡] Relative rates of different facets are what determines the shape, not the absolute rate.

resulting bond densities, and hence relative growth rates as a function of angle away from the [100] direction are shown in Figure 24. Facets will be seen for the low growth rate directions.

Comparison of Model With Growths

Most of the features of the growth habit can be explained with the above simple model, the consequences of which are summarized by Figure 24. At high V/III ratio the [011] oriented mesa shows a (111)B facet sidewall, while the [01$\bar{1}$] mesas show irregular lateral growth. The (111)B facets form because that is the slow growth direction for B oriented planes. At lower V/III ratio the (111)A growth rate tends to zero, and hence the [01$\bar{1}$] oriented mesas show (111)A sidewalls. In contrast, the growth rate for B oriented planes is higher than the vertical (i.e. [100]) growth rate, so the [011] oriented mesas do not show (111)B planes. However the growth rate of B-oriented planes on [011] oriented mesas is not so high as that of the A-oriented planes on [01$\bar{1}$] oriented mesas grown at high V/III.

The model can be extended to planes between the (111) and the (011) by expanding in terms of those planes.

The above growths were performed at high temperatures (~570°C). For lower temperatures (~540°C) , the (111)B facet is expressed for lower V/III ratios. In the context of the model, the reduced substrate temperature results

in a higher coverage of arsenic for fixed V/III ratio.[53] Hence the high V/III ratio analysis, which assumed complete arsenic coverage, is more appropriate for a wider range of V/III ratios at lower temperatures.

Instability near (311)A

One problem with [011] oriented mesas grown at high V/III (~2) is the extreme irregular lateral growth, showing tilted planes near (311) among others. Part of the explanation lies in the high V/III ratio bond density graph (Figure 24a). The B-planes always lie below the A planes in bond density, so an A-plane can reduce its bond density, and so reduce its growth rate, by tilting away from the A direction to the B. This explains the tendency to tilt, and partly explains the roughness described earlier.

The (311) plane is special in our model too, because it has equal numbers of (100) and (111) segments, and so marks the transition between (100) and (111) type planes. As a result it is the "roughest" plane. Until now only bond densities have been considered, not their distribution. It is possible that this special distribution of bonds on the (311) plane is connected to its expression at high V/III ratio. In particular since substantial migration is seen at high V/III it is possible that this "roughness" influences migration in such a way as to express a facet (i.e. to allow a facet to persist.) However, the importance

of surface reconstruction should not be underestimated. Lying outside the model, this alone may be responsible for the expression of such a plane.

Photoluminescence

Nanoscale structures can be fabricated by other means. The reason that the previously described fabrication technique has been devised is that we wish to fabricate nanoscale structures which are optically active. On this level the project has been successful: when excited by visible laser light, all of the structures are optically active. In what follows, the emission spectra of these structures will be examined in some detail.

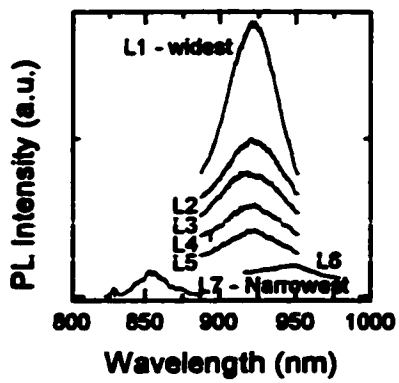


Figure 25. Photoluminescence for various GaAs mesa widths

The sample is a thin InGaAs layer between a GaAs buffer and GaAs cap grown on windows in a silicon dioxide masked gallium arsenide substrate. Windows range from ~20 μm (L1) to ~2μm (L7). Spectra were taken at 4.2K.

Mesa Width

Since PL from an unpatterned mesa is fairly straightforward to interpret, it is of interest to study the PL as the mask width is decreased. The samples, in

a cryostat at 4.2K, were excited by a HeNe laser of 3 mW attenuated by a 20dB neutral density filter focused to a ~100 μm spot, and the emission dispersed by a spectrometer and detected by a CCD array. Figure 25 shows the spectra of a high indium content (~18% In) wide well (~10 nm), grown on an [011] oriented masked substrate.[6] The peaks are extremely broad for two reasons - one is the unusually high temperature (for CBE) at which the InGaAs was grown here (~570°C). This was required to ensure good selectivity, but because of the high temperature, impurities are incorporated and the peak is broadened. Another reason for the breadth is that at such high indium concentrations the layer is likely relaxed (this is suggested by x-ray observations). Under optimum conditions on unpatterned substrates it was possible to grow InGaAs layers with narrow PL at a longer wavelength (~930 nm) than most peaks seen in Figure 25, however. Even with these broad peaks a change is seen with decreasing mask width.

As the mesa width is decreased (i.e. moving from L1 at the top, down to L5) the PL peak position stays constant, close to 920 nm. But when the mesa gets still narrower the PL first shifts to a still longer wavelength, almost 950 nm. Finally, for the narrowest mesa the PL appears at a short wavelength, near 855 nm.

[6] That these numbers are consistent with the PL can be determined with reference to Appendix C

At first this seems difficult to interpret, but it can be explained with reference to Figure 12, and becomes especially clear when viewed in light of data presented later in this thesis for similar InP based structures. The important feature that can be seen in Figure 12 is that the (100) growth front is curved only for the nearest ~0.8 μm to the (111)B facet edge. This curvature arises from the migration of group III material in particular, off the (111)B plane to the nearly (100) surface. There is no reason to assume that indium does not migrate as well as gallium. In fact, because of the weaker binding energy of indium it is likely that it migrates preferentially. A quantum well will be wider then, and may well contain more indium for this curved region. The PL of the widest five mesas is predominantly from the (100) top, but for the second to narrowest mesa, only the curved edges remain, with thicker wells and more indium, thus the PL peak shifts to a longer wavelength. Finally for the narrowest mesa, no (100) surface remains, just a crescent and thin (111)B quantum wells. We identify the short wavelength peak with the (111)B quantum wells. (The peak at 830 nm is due to a carbon acceptor.)

The mesa width dependence of the orthogonal orientation is more complicated. This is not too surprising, given the multifaceted sidewalls, as can be seen at right on Figure 17, for example. The photoluminescence for

this case shows additional lines, which are not straightforward to interpret. Since this orientation will not be used to try to fabricate mesa wire structures, its PL will not be discussed further in this work.

Substrate Temperature Dependence

Since growth at high temperatures was seen to be detrimental to the material quality, in terms of the width of PL peaks, the growth temperature

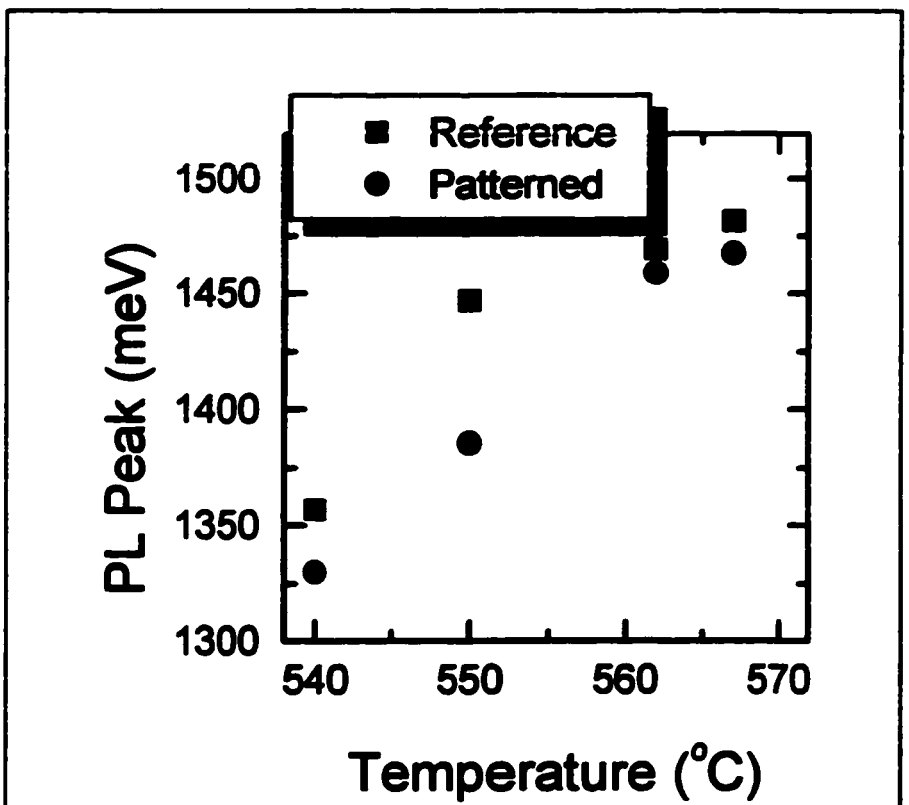


Figure 26. Temperature dependence of principal PL peak

For unpatterned reference quantum wells and for narrow V-shaped mesas. The widths of the V-shaped mesas at their base was <2 μm.

was reduced. This meant that indium desorption was reduced and the indium composition of the layers was increased. Below about 540°C the silicon dioxide mask became coated by micron scale globules of non-selective growth, therefore the procedure of growing the buffer at high temperature and dropping the temperature only for the well layer was used.

Photoluminescence was measured for a series of samples which differed only in substrate temperature during growth. In this case, samples that were held at a temperature of 5K were excited by a 10 mW Krypton laser focused to a very tight (~10μm) spot. The emission was dispersed by a spectrometer and detected with a germanium *pin* photodiode. For fixed fluxes, the unpatterned substrate quantum well peak shifted to long wavelength, from overlapping the carbon peak (~1490 meV) at high growth temperature (570°C), to 1356 meV for low growth temperature (540°C). (Figure 26.) From a growth rate calibrated by x-ray diffractometry - to a superlattice grown for comparison to the unpatterned sample - the well width is estimated at 6.5±0.5 nm. For the unpatterned samples the composition can be deduced from the wavelength (in the manner of Appendix C), ranging from 6% indium at high temperature to about 20% at low temperature.

The principal peaks of the narrow mesa structures are in every case at longer wavelength than the unpatterned reference quantum well. Even at high

temperatures, the quantum wire area luminescence indicates a significant amount of indium has been incorporated.

For the V-shaped mesas, the actual composition of the quantum wire structures is less certain, because the shape of the confining potential is difficult to determine. The morphology of the wire area is not as simple as a quantum well, and the distribution of indium in the quantum well is unknown. There are two approximations that can be made for the structure. One is to presume that the well width is unchanged, and determine the resulting composition. The redshift of the patterned sample peaks can be explained by an effective composition increase of ~4% indium on average for the patterned samples. The other approximation is to determine the well width required to

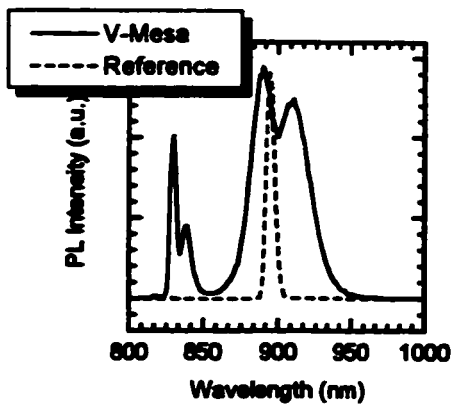


Figure 27. V-mesa quantum well spectra

A thin InGaAs layer is grown over a V-shaped GaAs. It shows four peaks, a carbon acceptor, a short wavelength peak probably associated with (111)B sidewall quantum wells, and two long wavelength peaks on either side of the reference quantum well peak, associated with the crescent. Here the reference PL was obtained with HeNe laser excitation, the V-mesa PL with a Ti:Sapp tuned to 773.5 nm. The temperature was 5 K. For most of the sample area the unpatterned reference sample showed an unusual broad tail centered near 950.0 nm.

maintain the same composition. In this case the shift in PL spectra requires an increase in well width of, on average, over a factor of two.

Narrow Mesa Spectra

PL with better signal to noise ratio than the previous set was obtained using a photomultiplier tube as the detector for a double grating spectrometer. Both helium neon (HeNe) and (tunable) titanium sapphire lasers (Ti:Sapp) were used for excitation. The HeNe laser had a power of ~3 mW and ~0.1 mm diameter spotsize, while the Ti:Sapp had a power of ~10 mW and spotsize ~0.5 mm in diameter. The sample was cleaved such that only one type of structure was observed at one time - here in particular, the narrow V-shaped structures, and later an unpatterned reference sample.

If excitation intensity is reduced to this power, typically a good quality narrow V-shaped mesa shows several peaks in its spectra (Figure 27). First, all our samples showed a carbon related peak at 830 nm. Typically there is a short wavelength peak with no counterpart in the unpatterned PL at ~840 nm. A broad principal peak is seen at a longer wavelength than the unpatterned quantum well peak. At low excitation powers for some of the samples this broad peak splits into two peaks, one of which was always at longer wavelength, another of which was at shorter wavelength, as compared to reference quantum wells.

Since they always appear in the same relation to the unpatterned (100) peak, we assume they must be associated with a region of quantum well confinement (and perhaps additional wire confinement) of similar width and composition to the unpatterned sample. The only place this seems likely on a patterned V is near the top of the crescent. It seems unlikely these peaks could be associated with the (111)B sidewalls alone since even if the growth rate were the same there, the well would be only 57% as thick owing to geometrical considerations. Also, as will be shown quantitatively for InGaAs on InP, the growth rate is typically lower there, even for a complete mesa.

Since these (111)B sidewalls appear to have a low, but non-zero growth rate of InGaAs, and since the unpatterned samples lack this peak altogether, it seems reasonable to associate the peak at 840 nm with thin, shallow side quantum wells.

Comment on the difficulty of excitation spectroscopy

Since the samples emit light efficiently, it would seem that excitation spectroscopy would be of interest, since it would access the states above the ground state, providing much more information. In photoluminescence excitation spectroscopy (PLE), the pump wavelength is scanned from a short wavelength out to a long wavelength peak while the detector monitors only the long wavelength peak. Excited states associated with the monitored peak

will show up as peaks. Unfortunately, apart from showing the bulk band edge, PLE on these samples did not even show the high energy tail of the monitored peak, let alone any band edge.

The failure of PLE is in a way a consequence of the success in making small structures. When normal PL is performed, the pump laser excites the entire crystal, and carriers decay efficiently into the lower quantum structure levels. In PLE, however, the intent is to excite carriers to levels within the structure. The barrier layers are very poor absorbers of photons in PLE because the excitation energy is beneath the barrier layer bandgap. The mesas produced here were of order 2 μm^2 in cross-section. The wire structures were certainly smaller than 50 nm×50 nm (~3×10^{-3} μm^2), perhaps even 6 nm×6 nm (~4×10^{-5} μm^2). The excitation cross section is one one-thousandth to as little as one one-hundred thousandth as small in excited state PLE as in PL.

This suggests that, in order to successfully detect a PLE signal, either very sensitive detectors are required, or there must be a very high filling factor of very uniform quantum structures.

V/III Ratio

Growing samples which differ only in arsine flux resulted in quite dramatic changes in the structure of the patterned sample PL, but little change, apart from broadening and a larger carbon acceptor peak in the unpatterned PL.

Multiple peaks in the same general region as the unpatterned peak shifted and broadened as the V/III ratio was varied. Unfortunately because the peaks were quite broad, and the V/III ratio was changed in large steps, it was not possible to deduce the physical meaning of these shifts. More study into the effect of V/III ratio on migration of quantum well material between (111)B and (100) facets is required.[54]

Conclusion

Fundamentally, the most important result of this chapter has been the demonstration that the fabrication technique works. Using faceted growth in selective area epitaxy it has been possible to decrease the feature size from a couple of microns to at least tens of nanometres. Furthermore, although the photoluminescence was broad, the structures were optically active, and certainly no worse than structures fabricated by other, arguably more complicated fabrication techniques. However, it has also been seen that there is considerable room for improvement. In particular, the structures appear unnecessarily rough. Moreover, the InGaAs:GaAs system has caused some problems, such as the tendency for indium to desorb at the elevated temperatures preferred for selective area growth of gallium with TEG, as well as the detrimental effects of strain inherent to this material system on the

optical properties. In the next chapter, it will be seen that many of these problems can be alleviated by changing to a new material system.

Chapter 3

REAL InP BASED STRUCTURES

Introduction

This chapter is an attempt to apply the knowledge gleaned from the gallium arsenide system to the indium phosphide based system. First some previous publications which relate directly to this work are described. After describing some of the advantages of indium phosphide based growth for this project, optimal growth conditions for faceting are described. The ability to grow thick layers of lattice matched InGaAs means that scanning electron microscopy can reveal much more about the growth process in this case. (On GaAs substrates the thickness of the InGaAs layer is limited because it quickly becomes defective due to strain. Also, the contrast between InP and InGaAs is excellent whereas for InGaAs on GaAs it may not be.) A simple diffusion model is applied to describe cross-sections observed in mesa structures. Multiple, relatively narrow photoluminescence peak widths make the optical studies more meaningful and systematic peaks shift associated with

narrowing the mesa width can be seen. Finally, inverted V-shaped mesas are grown, and their photoluminescence described.

Background

Before the advent of CBE, MOCVD was the best technique for growing indium phosphide based heterostructures. Since MOCVD, like CBE, uses molecular group III sources, the growth can be selective. Kayser *et al.*[55] describe low pressure MOCVD of structures like the wide mesas considered here. That work was later extended to CBE.[45] Some previous work describes the faceting behavior of wide InP mesas, typically with the intent of eliminating facets for device applications.[56,57] Faceted growth by CBE on micron wide mesas has also previously been studied,[50] but in that case the V-shaped mesas were not overgrown to form crescent shaped structures.

InP vs. GaAs

The samples grown on GaAs substrates had a number of problems. First, since the ideal growth conditions for conventional (100) GaAs and InGaAs were only quickly optimized, unintentional doping was, even for many of the unpatterned samples, excessive. Second, there was a conflict between selective area growth and incorporation of indium (the well material). Third,

probably due to the short migration length of species, there was usually serious microroughness associated with the mesa structures. Finally, the lattice mismatch due to the larger size of indium compared to gallium limited the range of compositions accessible, and so limited the depth of the potential wells. (An approximate upper limit [see Strain and Lattice Matching, p83] of ~20% indium corresponds to a conduction band well depth of ~0.14 eV while a lattice matched InGaAs:InP conduction band well has a depth of ~0.23 eV, and the valence band well is still deeper.)

All these problems are alleviated to some degree on InP substrates.

Material Quality

Since the CBE machine is primarily used for the growth of InP based laser structures, considerable effort had already been expended by others in optimizing the growth of InP. The mobility is widely accepted as a measure of the quality of grown material, and, in fact, the machine we used had previously been used to set the record of 311 000 $cm^2/V{\cdot}s$ at 50 K for mobility of InP.[58] Though the samples studied here would not have been as high quality as that, they would certainly be as pure as can be reasonably expected.

Selective Area Epitaxy

The conflict between SAE and indium incorporation for GaAs substrates is reduced for InP substrates. Desorption is a temperature activated process.[59] This is why selective area growth on dielectric masks fails at low temperatures. If the temperature is too low, the growth material, rather than desorbing from the mask material, sticks, and upon doing so nucleates further growth. To prevent gallium from sticking to the SiO_2 mask, it was necessary to increase the substrate temperature slightly from about 520°C to about 550°C. However, indium, which is a larger group III atom with a therefore weaker bond strength (in particular for methyl and ethyl groups[60], and by inference in the more general situation), for similar reasons tended to desorb from the growth surface. The result was that improving the selectivity of the silicon dioxide mask meant excessive desorption of quantum well material, and hence a lower potential depth than intended. This could be compensated by an increase in the flow rate of TMI, but at the expense of material quality.

For reasons related to indium desorption, the optimum substrate temperature of InP growth is low - typically ~440°C. Indium sticks so poorly to the mask that even at these temperatures almost perfect selectivity is obtained during the growth of InP. There is no problem of gallium desorption in the range of temperatures that produces good quality InP. It might be

surmised, however, that gallium would stick to the mask at this temperature - and it might appear that the situation is the same as for GaAs-based substrates.

Fortunately this is not the case. Although there is some loss of selectivity, it is by no means as bad as was seen on the GaAs substrates. Before explaining why, it may be useful to tabulate the selectivity behavior of gallium and indium described above (Table 2). The unwanted behavior for these growths is in bold type. For the normal growth of GaAs, at ~520°C, the problem is that gallium sticks to the mask. For the normal growth of InP, at ~440°C the problem is, again, gallium sticking to the mask. But in the GaAs based samples, gallium is supplied for the thick GaAs buffer layer, as well as through the quantum well and cap. For the InP based samples, gallium is only grown through the thin, narrower bandgap (InGaAs) layer. The supply of non-selective sources is substantially reduced for InP based samples, especially if only a thin quantum confined layer of narrow bandgap material is grown.

Table 2. Selectivity behavior of group III species at various temperatures on semiconductor and masked surfaces.

Growth Conditions	Selectivity of species			
	on semiconductor (GaAs or InP)		*on mask (SiO_2)*	
Substrate T (°C)	Ga	In	Ga	In
~570	sticks	**desorbs**	desorbs	desorbs
~520	sticks	sticks	**sticks**	desorbs
~440	sticks	sticks	**sticks**	desorbs

The group III species either sticks to the surface, or desorbs. Behavior that is a problem for these growths is in bold. The ideal GaAs growth temperature is ~520°C, while the ideal InP growth temperature is ~440°C.

In fact, there may be even another effect - one of surface chemistry - working in favor of selectivity for the InP based samples. Above, the group III sources have been treated as though they were elemental. However, in CBE the group III sources arrive at the surface with some organometallic bonds still attached. The difference in sticking between semiconductor and mask surfaces arises in large part because the group III elements arrive with these organometallics attached, and the semiconductor surface can catalyze their breakdown. Since the decomposition of the source molecules is a temperature activated process[61], for the lower temperature (~440°C), decomposition of the source TEG is less likely than at ~520°C. If the chances of desorbing are negligible at both temperatures, continuing to reduce the

temperature will only reduce the likelihood of decomposition. Thus, even though growth at low temperatures is non-selective, this problem could be less pronounced at very low temperatures than at intermediate temperatures.

It should be emphasized, again, that a clean mask surface is essential for selective area epitaxy, no matter whether the substrate is GaAs, InP or any other material. Slight defects or traces of process chemicals can nucleate non-selective growth.

Roughness

There was always a microroughness associated with GaAs based mesas. Qualitatively, roughness did depend on the mask quality and whether or not the substrate was rotated. Though the mesas remained terminated in inverted V-shapes, the V would shift sharply horizontally, and slightly vertically resulting in a jagged top. The scale of this roughness was less than about a half of a micron. This must have contributed to the breadth of the photoluminescence peaks.

The first growths of InP based mesas turned out to be much smoother on these scales. This might be attributed to the weaker bond strength[60], and longer diffusion length[62] of indium as compared to gallium. Field emission SEM photos show the (111)B sidewalls of InP structures to be considerably

smoother than GaAs structures. The roughness of InP structures will be described more thoroughly later in this chapter.

Strain and Lattice Matching

The last advantage of using InP compared to GaAs as a substrate material is that with InGaAs it is possible to grow material lattice matched to indium phosphide. This is possible because the lattice constant of InP (0.58687 nm at room temperature) lies between the larger InAs (0.60583 nm) and smaller GaAs (0.565325 nm) [63], and since the ratio of indium to indium plus gallium can be varied by changing flow rates, the average lattice constant of the InGaAs can be made the same as the InP. Lattice matching occurs for material of composition $In_{0.53}Ga_{0.47}As$. Layers with very little average strain (~0.1%) can be produced.

The drawback is that the formation of these crescent shaped narrow band gap regions is quite complicated, with migration of source material from other facets as well as incident flux contributing to the overall growth. As a result, lattice matching is probably not maintained throughout the structure. Moreover, unlike InGaAs:GaAs, since the indium concentration may be higher or lower than the lattice matched composition, both compressive and tensile strains are possible - perhaps even simultaneously in different regions of the structure.

Growth Conditions

Substrates were two inch diameter *n*-type sulfur doped (100)±0.2° oriented InP patterned with SiO_2 as described in Chapter 1, p20. The growth process was analogous to the GaAs substrates described in Chapter 2. Again, between patterning and introduction to the CBE machine, there was no cleaning. Quarter wafers were indium mounted on two inch silicon wafer which were in turn held by tungsten clips to a hollow molyblock. The blocks were placed into a vacuum chamber and brought to UHV conditions. Again, the sample was heated to 200°C for one hour to desorb unwanted organic molecules and then introduced to the growth chamber. A deoxidation step was performed, with the sample heated to 500°C for 10 minutes under a high flux of phosphine (20.0% full scale) to prevent the surface quality from deteriorating.

Table 3 Typical InP based growth conditions.

Description	*Duration*	*TMI (%)*	*TEG (%)*	*PH_3 (%)*	*AsH_3(#3) (%)*
InP buffer	15 min	40	–	10.0	–
In ramp	2 min	40 to 12.2	–	10.0	–
delay	65 sec	–	–	10.0	–
	1 sec	12.2	23.0	–	–
InGaAs layer	30 sec	12.2	23.0	–	15.0
	15 sec	–	–	–	15.0
interrupt	5 sec	–	–	–	–
In ramp	1 min	12.2 to 40	–	10.0	–
InP cap	4 min	40	–	10.0	–

Flow rates are given in percent of full scale.* A dash indicates no flow. The delay after ramping is to allow the flow to traverse the line from the mass flow controller to the sample and stabilize. Interruptions allow the growth surface to become smooth. Specific switching sequences between gas flows are intended to optimize the interface qualities.[64] Actual growth conditions change from sample to sample based on composition and growth rate calibrations by x-ray, SEM and PL. Arsine is delivered on the AS3 line. (See Appendix A)

Typical growth conditions for InGaAs on InP crescent structures are tabulated in Table 3. Substrates were rotated with a period of several seconds so that the beam would be more uniform, on average. One complexity arises from the fact that only one indium line is available†, and therefore switching between materials requires a ramp in TMI flow rate, since the growth of InP and InGaAs have different optimum TMI flow rates. For the growth of multiple quantum wells it was advantageous to maintain the same TMI flow

* See Appendix A for conversions of flow rates into physical parameters

† The limitation to one indium line is not fundamental. An extra line could have been added to the CBE machine used for this work (at extra expense and downtime.)

rate for the InP barrier layers (between InGaAs well layers) as for the InGaAs well layers themselves. This was because changing the flow rates would require an excessively long growth interruption between wells. In this case, since the TMI flow rate was less than a third for InGaAs as for the usual InP, the barrier layers were grown more slowly than usual. Also in this case, it would have been necessary to reduce proportionally the flux of phosphine in order to maintain constant V/III ratio for the entire growth of InP.

Structural Characterization

Growth Habit and Morphology

Initial tests of the growth of InP on line patterns showed that facets formed in a similar fashion to GaAs. For lines oriented in the [011] direction

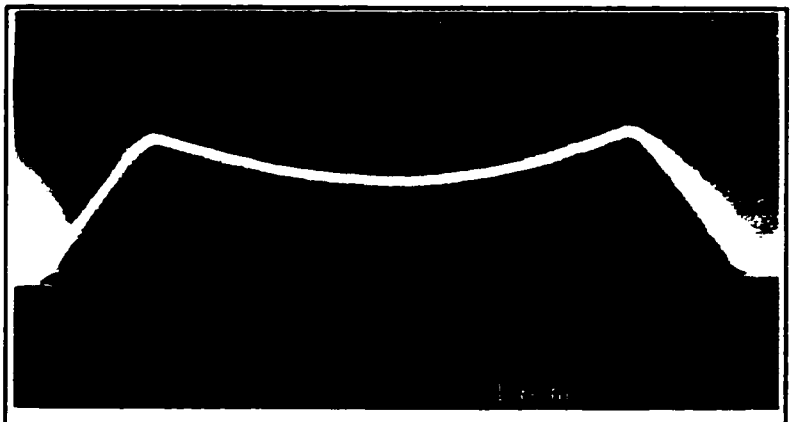


Figure 28. An [011] oriented InP mesa in cross-section.

The InP is dark. A faint light line just below the top is an InGaAs quantum well. The sidewalls are (111)B facets. The top surface has a curved exponential shape. The marker represents 1 μm.

(Figure 28), the growth habit was almost identical: (111)B facets, and a curving exponential decaying away from the (111) facet edge to the (100) surface. For [01$\bar{1}$] oriented lines (Figure 29) the situation was similar to GaAs in that there was substantial lateral overgrowth, however, the facet planes were less jagged than the GaAs counterparts, showing more clearly defined high index (*hkk*) planes. The formation of these planes was irregular, however, and apparently unstable. In general, the planes were smoother on InP samples than GaAs samples. Like the GaAs mesas, the (111)B planes were especially smooth before completion of the top V as shown in Figure 30.

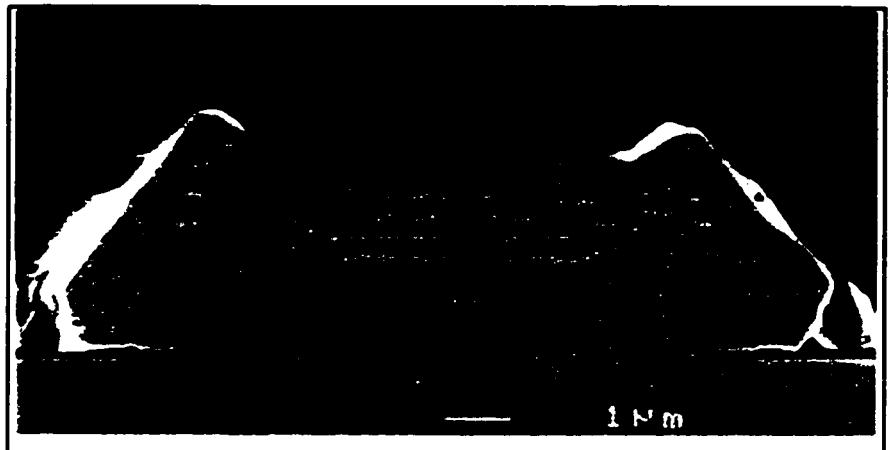


Figure 29. An [01$\bar{1}$] oriented InP mesa in cross-section.

Not all mesas showed such clear bilateral symmetry. There is significant lateral overgrowth. The light lines, nominally 110 nm thick are InGaAs marker layers. There are several clearly defined facet planes, as well as curved regions. The slightly darker InP regions on either side are probably doped unintentionally relative to the central part. The marker represents one micron.

Upon completion of the V-shape, the structures show some roughness but still much less than their GaAs counterparts, in general (Figure 31). This completed mesa shows three distinct kinds of roughness. First, there is an overall slight tilt of the (111)B planes on the scale of microns. Second, some localized areas of the (111)B planes show irregular, rounded growth, which appears to have started around the time of completion of the V. Third, on very small scales (<0.1μm) at the vertex of the V, most of the samples showed a distinct transverse hatched microroughness. Finally, for wide mesas, the morphology of the (100) plane is quite similar to the GaAs for either orientation. As with GaAs, the [01$\bar{1}$] directed lines show rounded defects. The [011] directed lines show striated transverse defects, though they are broader than their GaAs counterparts (Figure 32).

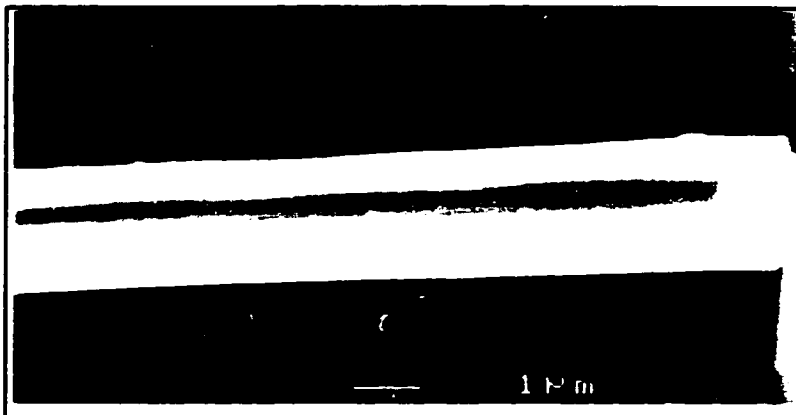


Figure 30. Roughness of an [01 1] mesa just before completion.

An [011] mesa seen from above. Notice the slight non-selectivity on the mask surface. The (111)B and (100) planes are smooth. The marker represents one micron.

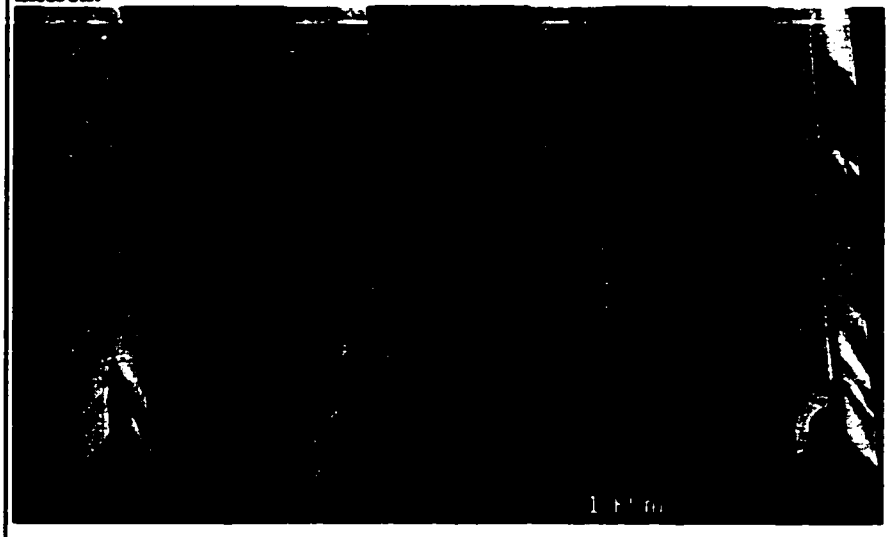


Figure 31. Roughness of a complete InP [011] mesa.

At least three different types of roughness can be seen (see text). The marker represents one micron.

Bulk InGaAs:InP

Because InGaAs can be lattice matched to InP, it becomes possible to grow thick (>50 nm) layers of InGaAs without the layers relaxing. Furthermore, FE-SEM is able to distinguish between InP and InGaAs even if the layers are as thin as five to ten nanometres, with no stain required - just a good cleave. It is possible, then, to grow thick tracer layers of InGaAs in InP to deduce some of the growth dynamics by viewing the mesas in cross-section.

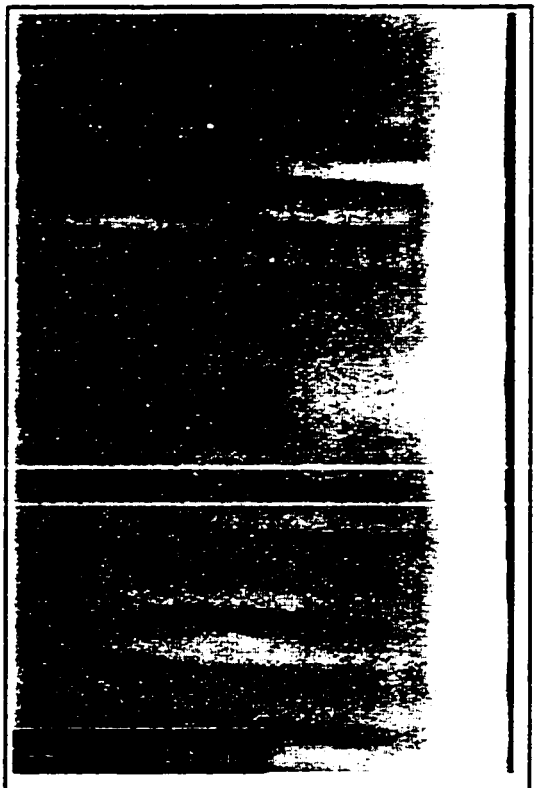


Figure 32. Wide mesa (100) roughness.

The dark band at the right is the masked area, and the white is the (111)B sidewall. The broad area taking up most of the photograph is the (100) surface, showing transverse striations. The gray patch in the middle is due to charging from the electron beam. The marker represents one micron.

Wide Mesa

A cross-sectional micrograph of a wide mesa (Figure 33) with marker layers leads to a number of observations. A summary of the growth conditions for this sample (#96058) is shown in Table 4. The sample consists of an InP layer grown at the higher growth rate, followed by a

Table 4. Growth parameters for sample 96058.

Layer	*Duration (min)*	*TMI (%)*	*TEG (%)*	*PH_3 (%)*	*AsH_3 (%)*
InP	60	40	-	10	-
InGaAs	5	12	23	-	15
InP	24	12	-	10	-

The growth temperature was 440°C. The second and third layers were repeated five times. Growth interruptions separated each layer. Arsine is delivered by the AS3 line. (Appendix A)

nominally lattice matched five period multilayer of 0.11 μm InGaAs and 0.19 μm of InP, with the multilayer InP grown at the low growth rate.

A straight (111)B facet is seen on the left of Figure 33. On the (111)B facet, starting from the edge of the first InGaAs layer, which marks the beginning of the slow growth rate layers, 0.51 μm of InP was grown. Perhaps

Figure 33. A wide InP mesa in cross-section.

Clear (111)B facets have formed on the left. To the right the mesa becomes flat (100). The light InGaAs marker layers show the curvature of the growth front. The top light surface of the mesa is overexposed and should not be misconstrued as an InGaAs layer. The marker represents 1 μm.

a more direct comparison between facet growth rates is between growth rates in the same direction. That is, the growth rate, usually defined normal to the surface, is instead measured with respect to a specific direction. For example, the growth rate on the (111)B surface can be measured in the [100] direction. This can be done by dividing the normal growth rate by the cosine of the angle between the [111] and [100] directions. The thickness of the InP layer grown on top of the (111)B facet plane inferred to exist at the time of the growth of the first InGaAs marker layer corresponds to 0.72 μm. Apart from very thin layers on localized regions (<1 μm) of the (111)B facet almost no InGaAs was grown. This compares to the (100) surface, far from the (111) facet, for which 0.6 μm of InGaAs was grown, and 1.0 μm of InP was grown. Therefore approximately 28% of the InP has migrated or desorbed from the (111)B, and all of the InGaAs has migrated or desorbed.

The amount of InP that desorbed was in fact too small to be measured accurately. Most of the InP migrated to the (100) surface, resulting in the curvature seen in the InGaAs marker layers. To within 20%, all of which can be attributed to measurement

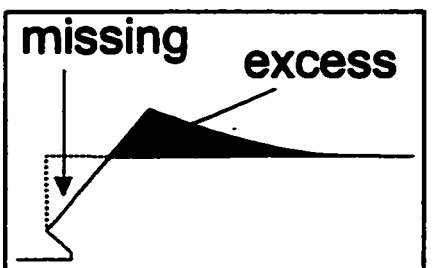


Figure 34. Missing and excess material.

The edge of a wide [011] mesa is shown above. Missing material has migrated off the sidewall to form an excess over and above the usual (100) growth.

error associated with the difficulty in defining the boundaries exactly, the "missing" InP from the (111)B plane is made up for by the area defined by the lowest curved InGaAs layer, the initial (111)B plane, and the (100) surface (Figure 34). Both areas are ~1μm^2, leading to a area of 150 nm^2 viewed in cross-section which migrates in one second, under the high growth rate conditions.

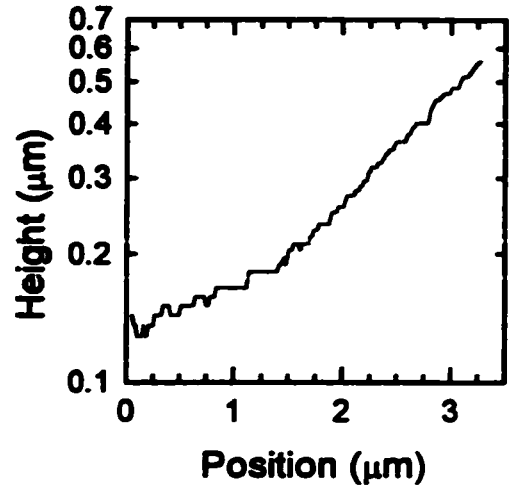


Figure 35. Semi-logarithmic graph of growth front

A straight line fits the logarithm of the height of the growth as a function of distance away along the substrate, away from the facet edge, which is at the 3 μm position (i.e. the right hand side). The characteristic length is 1.9 μm.

The curvature associated with the growth front strongly suggests that the migration is simply diffusion in response to a concentration gradient in available lattice sites. A similar FE-SEM micrograph of another sample (#96032) grown under the same conditions was digitized and converted to graphical data using a freely available image processing package.[65] The

InGaAs marker layer positions fit a straight line very well when plotted on a semi-logarithmic graph, meaning that the height of the growth front grows exponentially toward the (111)B facet (Figure 35). This is suggestive of diffusion because the exponential represents an elementary solution to the diffusion equation with one boundary being a source of material (see Diffusion Model, p 96).

The characteristic length in the plane of the substrate from the facet edge to the point where the InGaAs marker layer appears to be flat (100) is 1.7 μm for fast growth InP, and 3.4 μm on the slow growth InP, or an increase of 200% for a reduction in TMI flux by 30%, and no change in flux of phosphine. This suggests that the diffusion is in response to a gradient in the group III element occupation from the (111)B to the (100).

The normal (100) InP growth rate is very nearly linear in TMI flux, so it might appear that the increase in diffusion length stems from the 30% growth rate reduction on the (100). However, since the phosphorus supply was unchanged while the indium supply (Table 4) was reduced, in fact the V/III ratio was reduced. This is qualitatively similar to the GaAs samples for which material seemed more inclined to redistribute itself at high V/III.

It should be noted that the top (100) surface seems to have nearly reached an equilibrium in diffusion, since the beginning of the flat area on successive marker layers is very nearly parallel to the (111)B facet.

Narrow Mesas

Of particular interest with regard to quantum wire structures are the growth dynamics of completed V mesas. Figure 36 and Figure 37 show the same growth run considered above, but on narrow mesas. One striking feature is the sharp triangular shape of the final (100) growth which completes the ridge. Qualitatively, this indicates just how much more preferable growth on (100) is for InGaAs as compared to growth on the (111). This feature will be discussed in more detail later.

The nominal thickness of the InGaAs on an unpatterned mesa is 0.11 μm, which is also the thickness far from the facet edge on patterned substrates. The close-up on the sidewall InGaAs (Figure 37) reveals that, for a completed V, the nominal (111) thickness is ~18 nm (though the uncertainty is quite large and difficult to estimate owing to the resolution of the FE-SEM). For the (111)B, this corresponds to ~31 nm if measured in the [100] direction, instead of along the [111] normal - a reduction in growth rate to 28% of the unpatterned rate.

On the other hand, for a completed V-mesa there appears to be no reduction in the growth rate of InP. Though many of these layers are asymmetric, on average the completed (111)B InP layers are about 0.10 μm thick, which corresponds to 0.18 μm in the (100) direction. This agrees well with the nominal 0.19 μm (100) InP growth on unpatterned areas. For the most part, then, InP does not desorb from the (111)B for these growth conditions.

Diffusion Model

An exponential profile develops naturally for diffusive processes, and

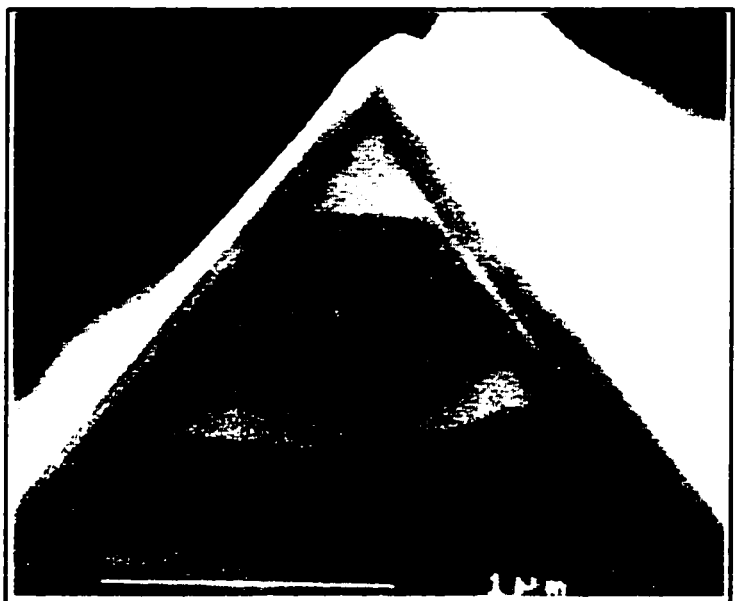


Figure 36. A completed V-shaped mesa
InGaAs is light and InP is dark. The triangular light region is perhaps most striking. Note also the appearance of InGaAs on the (111)B sidewalls once the V is complete. The marker represents one micron.

specifically for the type of faceting process considered here. Here a simple description of the diffusion of material from the (111)B plane to the (100) is fully derived. [66,67]

First, it is necessary to be clear about exactly what is diffusing. In epitaxy, the underlying crystal is rigid. In molecular beam techniques there is no gaseous boundary layer. However, the top fraction of a monolayer is covered with species which have not yet been incorporated into the lattice. It is these species which diffuse. The only difference between normal surface diffusion in epitaxy and the present case is that the diffusion takes place on more than one facet.

Consider first a one dimensional diffusion along a flat surface. Using a continuum model with only one species for simplicity, let x be the coordinate along the surface, and $u(x)$ be the number density (per unit length in 1D) of species on the surface. The rate of change of this occupation is equal to the flux arriving, less the amount incorporated into the underlying crystal, less the amount desorbed, plus the amount diffusing to a point x. The flux density depends on the cosine of the surface normal with the incident beam angle, and can be written as $F\cos\theta$ with F the normal flux and θ the angle. The amount incorporated per unit time into the underlying crystal is proportional to the amount of material on the top layer, and can be written u/τ_{inc} where τ_{inc} is the

incorporation time, a rate constant. The amount desorbed per unit time is described in a similar fashion, and can be written u/τ_{des} where τ_{des} is the desorption time. Finally, there must be a diffusion term, with a surface diffusion constant D, which can be defined by $D=L^2/\tau_{diff}$ where L is a characteristic surface diffusion length, and τ_{diff} is a diffusion time. (In general the diffusion "constant" D need not be constant at all, in which case L and τ_{diff} must vary as well. For example, it may depend on position or the angle of the surface normal. Here only the simplest case of constant D is considered.) Balancing all these terms,

$$\frac{\partial u}{\partial t} = F\cos\theta - \frac{u}{\tau_{inc}} - \frac{u}{\tau_{des}} + D\frac{\partial^2 u}{\partial x^2} \qquad (6)$$

for an arbitrary surface. Note that the incorporation and desorption terms can be combined if a single characteristic time is defined as the inverse of the sum of the inverses of these times.

Next, consider the (111)B facet. For this facet $\cos\theta=1/\sqrt{3}$. Assume that the diffusion length is very long compared to the length of the facet, so that it has uniform density. Defining τ as above

$$\frac{1}{\tau} = \frac{1}{\tau_{inc}} + \frac{1}{\tau_{des}} \qquad (7)$$

and consider only the steady state. Then

$$u = \tau F \cos\theta = \frac{\tau F}{\sqrt{3}} \equiv u_0 \qquad (8)$$

a constant.

Next, consider the (100) surface, for which $\cos\theta=1$, in the steady state. Under normal growth conditions, there is little or no desorption and $\tau_{des}\to\infty$. Then (6) becomes

$$0 = F - \frac{u}{\tau_{inc}} + D\frac{\partial^2 u}{\partial x^2} \qquad (9)$$

the flux term can be combined with the others by using comoving coordinates with $v=u-\tau_{inc}F$. Then (9) becomes

$$0 = -\frac{v}{\tau_{inc}} + D\frac{\partial^2 v}{\partial x^2}. \qquad (10)$$

Suppose that at $x=0$ this (100) facet is met by an infinite (111)B facet like the one described above. At this boundary the u must match. Then $u(0)=u_0$, or $v(0) = u_0 - \tau_{inc}F \equiv v_0$. Let the (100) facet also be infinite. Then, as $x\to\infty$, if there are no other sources, diffusion must die off, and $v\to 0$. (i.e. the occupation u depends only on the flux term.) It should be emphasized that x is measured along the surface, not along a Cartesian coordinate, so it need not necessarily be parallel to the (100) plane.

The only solution to (10) satisfying the above boundary conditions is a decaying exponential in x:

$$v = v_0 e^{-x/L}. \tag{11}$$

Writing this in terms of u and substituting terms,

$$u = F\tau_{inc}^{100}\left[1+\left(\frac{1}{\sqrt{3}}\cdot\frac{\tau^{111}}{\tau_{inc}^{100}}-1\right)e^{-x/L}\right]. \tag{12}$$

where the superscripts are to make explicit the crystallographic orientation in question. Recall that the material is incorporated at a rate proportional to u, with the proportionality the incorporation constant, so if the height is small enough so that the Cartesian distance from the vertex is close to the distance x along the surface (true for small enough time t), the height is, where F and u must be interpreted as heights per unit time, t.

$$h = tF\left[1+\left(\frac{1}{\sqrt{3}}\cdot\frac{\tau^{111}}{\tau_{inc}^{100}}-1\right)e^{-x/L}\right]. \tag{13}$$

Provided desorption is not too strong the (111) incorporation time is much longer than the (100) and so the quantity in parentheses should be positive. In this way, diffusion gives rise to an exponential profile.

For the real growth, the vertex between facets was observed to shift steadily towards the (100) surface. Assuming that it is perfectly linear in time, the above solution works as well with the substitution $x \to s = x - ct$. In that case,

the solution is steady state in the comoving s coordinate, but of course cannot be steady state in x. Substitution of such a solution back into the diffusion equation (6) shows that the diffusion length L is different, given by

$$L = \frac{c\tau}{2}\left(\sqrt{1+\frac{4D}{\tau c^2}}-1\right). \qquad (14)$$

as $c \to 0$ the usual $L = \sqrt{D\tau}$ is recovered. As $c \to \infty$ it is seen that the diffusion length L tends to zero. Generally, L decreases with increasing c. So the effective diffusion length is reduced by the facet movement. This is really just the facet vertex movement catching up with the diffusion.

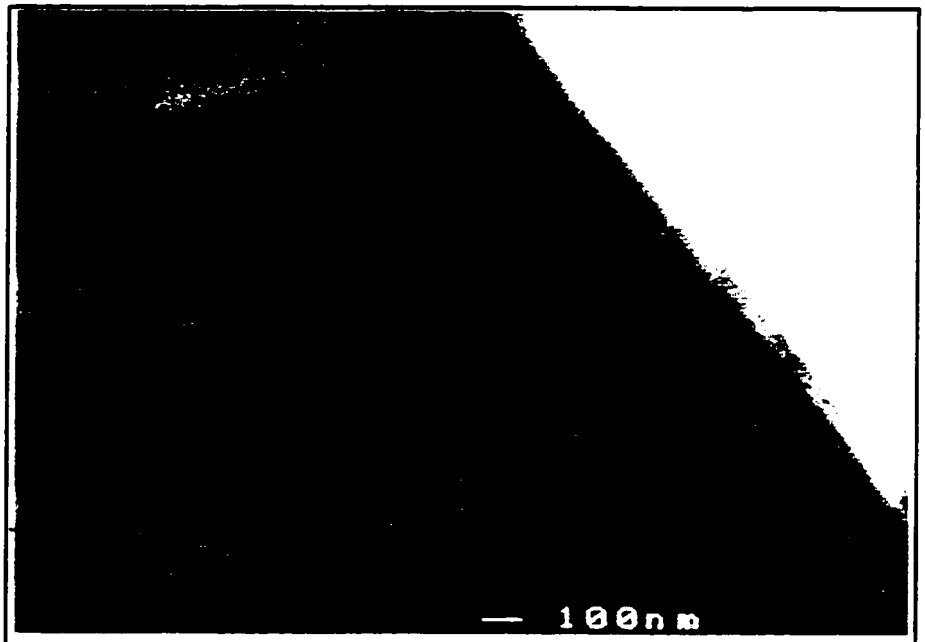


Figure 37. Close-up on the side of a completed mesa.

Note that the (111)B sidewall InGaAs (light area) grows only when the top "V" is completed. A very short (111)B tail is seen even on incomplete V layers. The white marker represents 100 nm.

A more complete solution would account for the finite facet length and finite diffusion length on the (111)B. Finally, the curvature of the surface should be accounted for, since the distance x along the surface is longer than the Cartesian distance measured along the exact (100) surface. The inclusion of the curvature gives rise to a highly non-linear problem which is difficult to solve in a consistent manner, and for which a simple solution is difficult to obtain.

Facet Mesa Structures

There are at least three separate types of mesa structure that can be fabricated based on what has been shown so far (Figure 38). First, the simplest, for a mesa which is wide relative to the facet length (or its height), and wide relative to the diffusion length on the (100) plane, there is conventional (100) growth far from the mesa edge. Provided the facet length is not too long relative to the diffusion length on that facet, there will be little growth on the (111)B plane, and migration will imply curvature near the edges of the nominally (100) surface, and thicker wells near the edges.

The other extreme is a mesa which is narrow relative to its height, such that a complete V is formed. In that case the remaining (100) surface is much smaller than the diffusion length along that surface. There is (111)B growth of InGaAs and InP, because there are very few of the more favored (100)

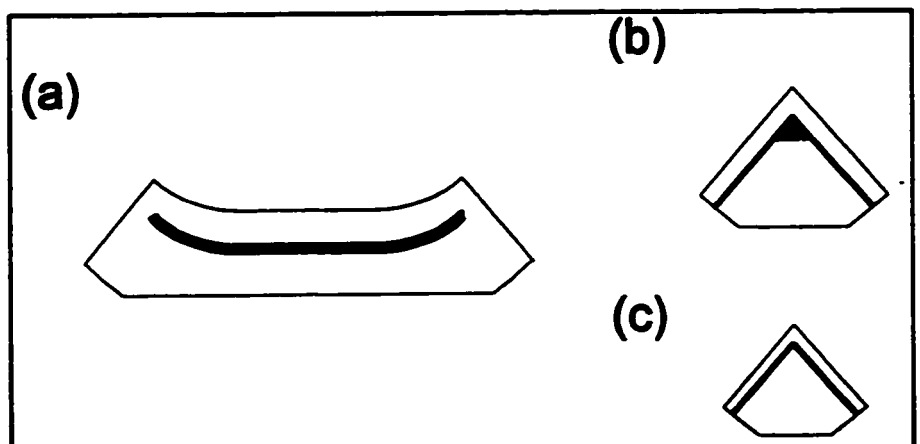


Figure 38. Three types of mesa structure.

There are at least three distinct cases for [011] oriented mesas, depending on whether the well material is grown (a) on a broad (selective) area (b) during completion of the V-shape or (c) on a complete V-shaped mesa.

sites available, and they are readily filled. The InGaAs desorbs more readily from the (111)B plane than the (100) and so the InGaAs growth rate is reduced. A sharp V (~20 nm at the tip) can be formed in this region, but for the well material its composition and thickness are difficult to control because they are more influenced by the migration of species from the (111)B than the incident flux. This is the type of wire growth that is pursued here.

A third case occurs from the time that the (100) facet is small relative to its diffusion length, but just before it is at its narrowest. Here, as in the first case there is a lot of migration of material from the (111)B to the (100), and hence the growth rate is increased dramatically on the (100). But because the narrow band gap material sticks so much more readily on the (100) compared to the (111)B, sharp facet planes are maintained. This raises the prospect of

growing triangular cross section wires bounded by (100) and two intersecting (111)B planes.

Though the third case leads to a more attractive wire structure in principle, in practice it is very difficult to grow quantum structures of this nature. To have an observable level spacing due to quantum effects, the structure has to be of linear dimension less than 10 nm[‡], in practice this amounts to approximately 20 seconds of growth. The entire mesa structure is grown in roughly 90 minutes. Even knowing the facet growth rate, and the exact width of the initial pattern, it is very difficult to time the growth of the narrow band gap material to occur just before the completion of the V. It becomes impossible, when variations in the width of mesas are included, to imagine forming this type of structure over a known area. *In situ* monitoring of the growth on nanometre scales could in principle eliminate the timing problem. Another possibility is to grow segments of many widths, with neighboring segments differing by ~10 nm in width. Then one could be assured that at least one line would show the type of growth seen in the third case on a quantum mechanical scale - but then the particular one must be

[‡] This can be seen using the infinite square well eigenstates to Schrodinger's equation using the quantum well's reduced effective mass. If the level spacing is to be greater than a few meV, the structure must be a few 10's of nm or less in width. A finite square well will have to be still narrower.

found. This may be of interest scientifically, but it is not a practical way to mass produce wire structures.

Photoluminescence

Photoluminescence provides an optical probe of the electronic structure of the sample. PL peak energies can be used to put constraints on the composition of compound semiconductors and the width of thin quantum wells. PL peak widths also provide an idea of the quality of the structure. It is also highly sensitive to dopant incorporation and is widely used as a tool for this purpose in the analysis of bulk semiconductors. Details on the relation between peak energies and structural parameters are given in Appendix C, p173.

Table 5. Mesa Widths

Mesa (#)	*Width (μm)*
1	19.6
2	9.8
3	7.2
4	5.0
5	4.1
6	3.0
7	1.8

Nominal width of opening in silicon dioxide mask.

In an effort to understand the growth behavior on the scales required for quantum wire fabrication it is of interest to study the PL of mesa structures as their width is decreased. The wide mesas are essentially conventional (100) structures, with faceted sidewalls, for which the PL was quite similar to the conventional growth, while the narrowest structures are dominated by the facets, finally becoming inverted V-mesas, which have very different PL. A

Layer	*Duration (min)*	*TMI (%)*	*TEG (%)*	*PH_3 (%)*	*AsH_3 (%)*
InP	25	40	–	10	–
InGaAs	0.5	12.2	23	–	15
InP	1	ramp	–	10	–
InP	3	40	–	10	–

Table 6. Growth parameters for sample 96089.

The growth temperature was 440°C. Arsine is delivered by the AS3 line.

mask with seven different nominal line widths, given in Table 5, was used to grow a thick InP buffer, a single wide quantum well, nominally 10.5 nm thick, and a cap layer. Specific growth conditions are tabulated in Table 6.

PL spectra for each mesa width were taken at low power, using a 3 mW HeNe laser, attenuated with a neutral density filter by a factor of one hundred for all mesas, except the narrowest, which was attenuated only by a factor of ten. The spot size was ~100 μm in diameter. The PL was collected by means of a converging lens and the spectra obtained by means of a Fourier transform interferometer (FTIR). The samples were held at a temperature of 4.2K unless specified otherwise.

[011] oriented mesas

Spectra for [011] oriented mesas of all the widths are shown in Figure 39. InGaAs quantum wells were grown at near lattice matching conditions under the growth conditions specified in Table 6. Most mesas show three peaks, in comparison to the single peak from an unpatterned reference quantum well. The most striking feature is the continuous shift of the PL to long wavelength, starting gradually for the wider mesas (e.g. a redshift of 16 meV for the first three mesas), but becoming extreme for the narrowest mesa. (There is a difference of 111 meV between the widest and the narrowest mesa.) The energies of peak emission are tabulated in Table 7.

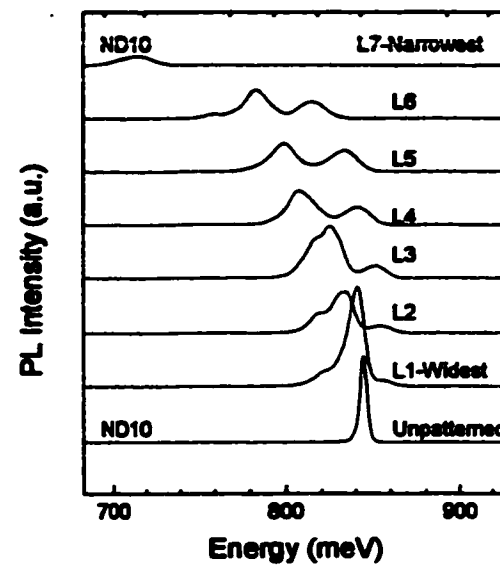


Figure 39. Photoluminescence of [011] oriented mesas of various widths at 4.2K

A 3.0 mW HeNe laser was used, with a 20 dB neutral density filter, except for those spectra marked ND10, for which a 10 dB filter was used. The unpatterned PL has been reduced to appear a similar size as the others. The other spectra have the same normalization.

Although this experiment alone is not enough to identify with certainty all the peaks, assignments can be made based on clues provided by this figure alone.

Table 7. Photoluminescence peak wavelength at 4.2K

Mesa	*Peak 1 (meV)*	*Peak 2 (meV)*	*Peak 3 (meV)*
Unpatterned	—	843.9	—
Widest	824.8	840.7	854.5
Second	818.6	833.5	855.0
Third	818.2	824.8	850.8
Fourth	817.7		851.1
Fifth	809.1		840.7
Sixth	757.8[§],782.8		815.3
Narrowest	713.9		—

For [01$\bar{1}$] oriented InP mesas with (111)B sidewalls. A thin (nominally 10.5 nm) InGaAs (nearly lattice matched) layer has been grown between InP barriers. Peaks 1 and 2 merge for narrow lines, hence the single column for peaks 1 and 2, below the third line.

First, the unpatterned sample shows strong PL at one wavelength. It is expected that the widest mesa should show a similar PL spectrum. Indeed, the widest mesa shows a very strong peak shifted slightly to longer wavelength. This peak is found in the middle of two others for the widest four mesas, and gradually disappears into the other peaks as the mesa gets narrower. Because of the proximity in energy of this dominant peak for the widest lines to the unpatterned wafer peak, it seems reasonable to attribute it to the (100) quantum well. The narrowness of the unpatterned peak (5 meV) along with the absence of any other peak means it must attributed to the free exciton of the InGaAs quantum well layer.

[§] Apparently a new peak unrelated to the wider mesas' third peak.

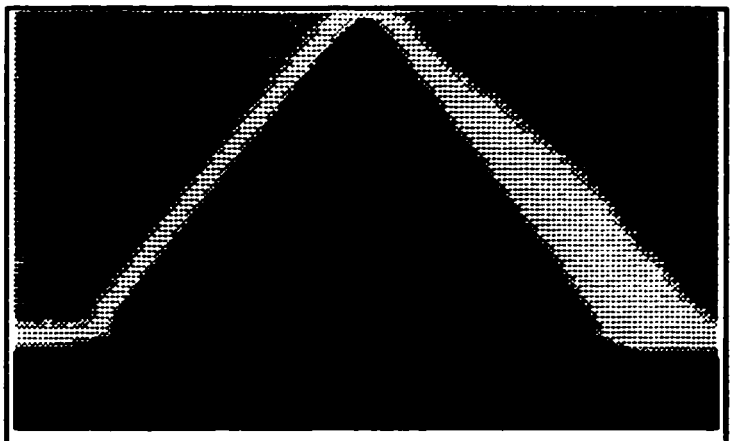


Figure 40. The narrowest [011] InP mesa.

The thin light horizontal layer is an InGaAs quantum well. The quantum well layer is flat, not grown in a V shape. The marker represents 100 nm.

Another feature of this spectrum seen on the widest four lines is that the two outside peaks shift gradually together to lower energy, and remain equally spaced (≈30 meV) throughout.

Scanning electron micrographs (Figure 40) reveal that the narrowest mesa has grown to an inverted V-shape, however the InGaAs was grown before the V-mesa was complete. It is a flat (100) well, about 330 nm across, and is probably too thick to show strong confinement effects (~ 20 nm thick). There is no evidence of (111)B planes of InGaAs growth in the SEM photo.

Examining Figure 39, for the fourth through sixth widest lines, it is seen that the dominant peak of the wider mesas disappears, apparently into the low energy peak. SEM photographs reveal that as this is occurring, the flat (100) region is being replaced by the curved sidewall areas which are moving

together and overlapping. Identifying the low energy peak with the curved edge slightly off the (100) plane provides an explanation for the merging of peaks, since the two regions, curved near (100) surfaces, and flat (100) surfaces, merge together as the top of the mesa becomes narrower than the diffusion length.

The extra peak in the second narrowest mesa at low energy is likely an acceptor related impurity transition, separated as it is by roughly 30 meV from the peak associated with the curved top. Observation of this impurity peak suggests that there is essentially no quantum confinement in the vertical direction, a conclusion which is consistent with the growth rate enhancement in this region. (See e.g. Figure 36.)

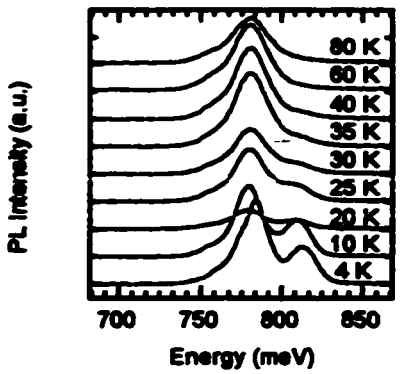


Figure 41. Temperature dependent photoluminescence
For the third widest line of an [011] oriented mesa.

At this point, it may be useful to quantify some of the shifts seen in the PL peaks. The unpatterned well has an estimated thickness of 10.5 nm. From the tables in Appendix C, this means the composition is 45.8% Ga - a little over one percent indium rich compared to lattice matching. For the second widest

line, the dominant peak has shifted to 833 meV, which would correspond to an increase of 1% in indium content, or a width enhancement of ~30%. SEM micrographs of wide InGaAs layers suggest that the thinnest InGaAs regions of the wider mesas do not differ measurably in width from the unpatterned sample. However, for the narrow mesas, there can be a substantial growth rate enhancement. This is seen at its most extreme in Figure 36, where it amounts to an almost threefold growth rate enhancement. Looking at the sixth line, the peak has shifted to 809 meV, which would correspond to 41.5% Ga if the well remained the same width. However, this is very close to the bulk bandgap of 805 meV for InGaAs with a composition of 45.8% Ga. Thus the shift to this point can be explained entirely by an increase in the well width. The last line, at 713 meV is shifted so far, however, that it must be concluded that the indium concentration has also increased by perhaps as much as 15%. In general it would appear that some of the shift is due to a increase in indium content, and some due to an increased well width. The (111)B facets supply gallium and indium to the (100) plane, but it would appear that indium is supplied at a higher rate. This is consistent with the longer diffusion length usually associated with indium.

The temperature dependent PL (Figure 41) of these levels reveals that the highest energy peak disappears very rapidly with an increase in temperature. This, at first sight seems unusual, since one might expect, instead, a lower energy impurity peak would disappear with temperature as the impurities become ionized. But if the high energy peak is related to the curved area tilted slightly away from the (100) plane and the middle peak is related to the

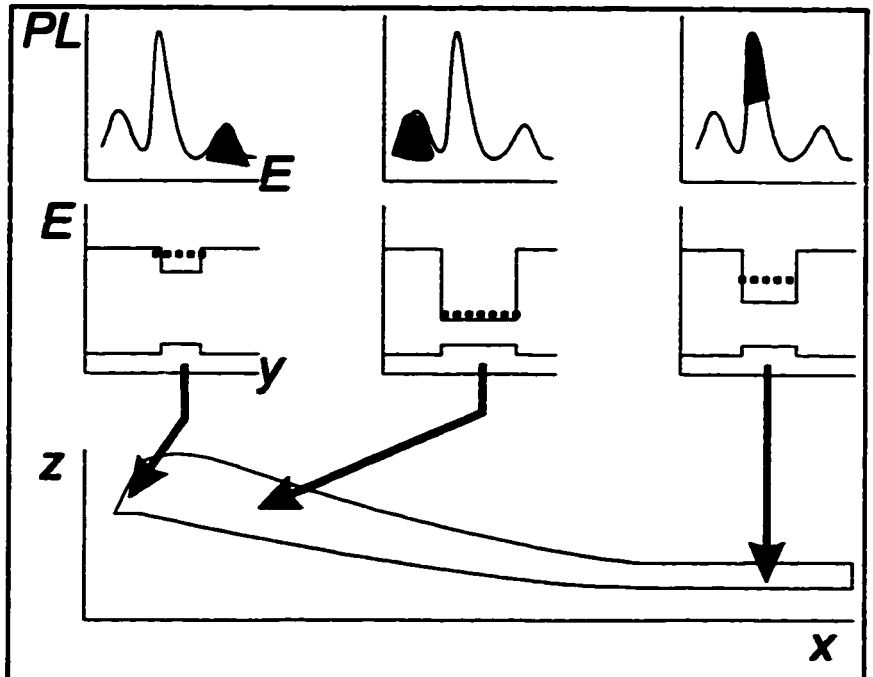


Figure 42. Model of [011] mesas.

Features in the optical spectra are related to structural parameters. The top schematically represents the PL spectra. The peak in question is dark. Below each spectrum is a graph of the energy band structure with depth into the structure. The narrower gap region is InGaAs. An arrow points to a schematic of the mesa InGaAs region showing where the transition takes place. The *z*-direction is the growth direction, the *x*-direction is measured perpendicular to the mesa stripe, but the *y*-direction in the energy band diagrams represents the perpendicular to the quantum well plane (usually close to the *z*-direction).

(100) plane this effect can be understood.

In this case, it takes only a small amount of thermal energy to send excitons from the curved portion to the flat top. Since for the widest mesa the flat top is large, recombination can occur there without difficulty. In fact, the high energy peak is all but gone at 30K, which corresponds to about k_BT=3 meV. By comparison the two largest peaks are separated by 28 meV.

The above discussion of the PL data together with the SEM micrographs is summed up by Figure 42, wherein each PL peak represents a distinct lateral region of the mesa well.

Inverted V-Shaped Mesas

For narrow mesas, which had grown to a V under the growth conditions of Table 6, results were similar to growths of V-shaped mesas InGaAs on GaAs. Photoluminescence revealed two overlapping broad peaks shifted to low energy (Figure 43). The reference quantum well grown on an unpatterned substrate corresponds to nearly lattice matched (47% Ga) with a width of ~8.6 nm. The low energy peak was shifted down by 90 meV and the intermediate energy peak shifted down by 41 meV. The high energy peak corresponds to an increase in indium composition (to 40% Ga) or an increase in width to 19.0 nm. For the lowest energy peak it appears that the

composition must have shifted, because even for the extreme case of bulk InGaAs the composition would have to be ~42% Ga.

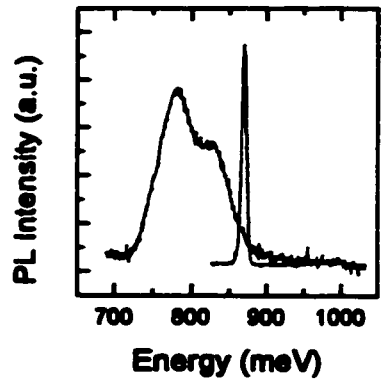


Figure 43. Photoluminescence of an InP Inverted-V Crescent

The sharp peak at 870 meV is a reference quantum well. The broad double peak at 829 meV and 780 meV comes from the InGaAs crescent structure on InP.

The true structure does show a bulge at the vertex of the V. This can be seen from an SEM photo of a similar growth (Figure 44), for which thick layers of InGaAs were grown, rather than thin quantum wells. (For a thin quantum well layer the InGaAs layer would be very difficult to observe.) In contrast to layers grown while the V still had a bit of a flat top, which can show substantial growth rate enhancement, the maximum thickness of the crescent on a completed V is typically less than the nominal (100) well width. The (111)B sidewall growth, while somewhat asymmetric, is about a third as wide as the bulge of the crescent. The radius of curvature of the indium phosphide layer in this image is ~30.0 nm. The top radius of curvature of the InGaAs layer is thinner.

There is a lot of variability between layers, but it may be instructive to use Figure 44 to predict the structure of the much thinner InGaAs layer. The

nominal (100) growth for the thick layers of InGaAs was 110 nm. From the middle crescent in Figure 44, the maximum bulge is 76 nm, and the left and right sides 26 nm and 44 nm, respectively for an average of 35 nm. The radius of curvature is 30 nm. Scaling this down to the nominal 8.5 nm thick (100) quantum well for which the PL is presented in Figure 43, the result is ~6.0 nm bulge and 2.7 nm minimum. The radius of curvature should remain unchanged, at 30 nm.

One model consistent with the PL would be that the low energy peak comes from the bulge area, and the second peak comes from the side quantum wells. However, the (111)B quantum wells are usually quite thin, and it seems unlikely that the observed PL comes from them, unless they are indium rich. Note, however, the right hand side quantum well of the bottom crescent in Figure 44, which suggests some (111)B wells may be wide enough to be the origin of the second peak. In general there was substantial roughness in the InGaAs crescent layer shape. Another model consistent with the PL could be that the peaks are radial (normal to the epilayer) modes, and that their breadth comes from a large number of very closely spaced angular (parallel to the epilayer) modes.

Conclusion

In this chapter, it has been demonstrated that indium phosphide based growth has many advantages for selective area epitaxy. Although the main goal had been to fabricate quantum wire structures, along the way it was shown that bandgap shifting of over 100 meV results from decreasing the mesa width in the micron regime. Though indium phosphide structures are smoother than their gallium arsenide counterparts on a broad scale, it has been shown that some roughness exists for the thinnest mesas. The explanation may be that growth conditions were chosen with the (100) surface in mind, and the formation of facets. However, best growth conditions may be different for overgrowth on those facets. In particular, the growth the V-shaped mesas is more similar to growth on unpatterned (111)B substrates than (100) growth, and ideal growth conditions may be different.

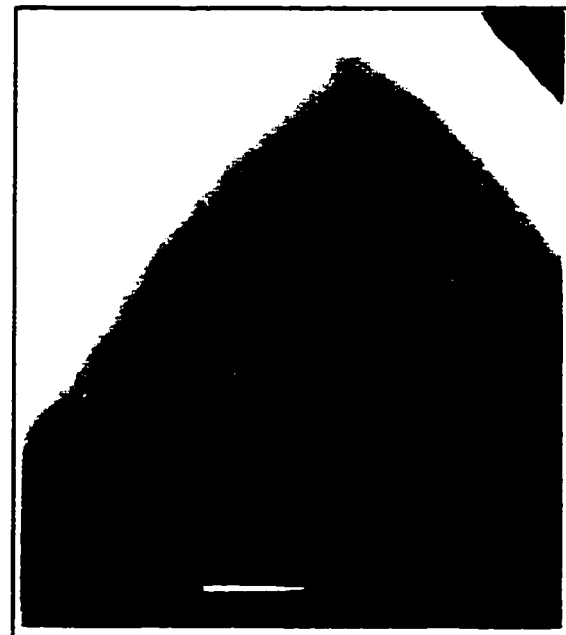


Figure 44. InGaAs V-shaped crescent structures on InP

The top is light because it is near the surface and overexposed. The light V-shaped crescents are InGaAs, and the dark areas are InP. The mesa is [011] oriented. The white marker, bottom, center, represents 100 nm.

SEM analysis of thick layers was used to estimate the size of quantum wire structures. The final wire structures are very small, and may be small enough to exhibit lateral confinement. However, there may be a limit to the sharpness of the facet intersections, which would limit the usefulness of the technique for fabricating nanostructures. The tips of the crescents always appeared curved with a radius of curvature of greater than 10 nm. If this is a fundamental limit, the structures will only show very weak confinement, even at 5K. Transmission electron microscopy or another high resolution technique might enable a more direct determination of structural properties. But photoluminescence, which has been used above quite extensively, is very sensitive to small scale structural parameters.

Ultimately the crescent photoluminescence is quite similar to that of the gallium arsenide based structures. However, unlike for simple heterostructures for which spectra can be related to structural parameters (for example, see Appendix C), we lack the analytical tools to make such an identification for crescent structures. In the next chapter this problem is addressed by developing a simple analytical model for geometrical quantum confinement.

Chapter 4

THE THEORY OF QUANTUM WIRES

Introduction

The ability to fabricate structures which can meaningfully be called quantum wires is very recent[23-29,35-43] and, compared to the bulk or to quantum well based heterostructures, quite primitive. Many important properties, some of which will be highlighted in the next section, have only been theoretically studied in the past five to ten years.

The treatment of conventional epitaxial heterostructures is well established. At its simplest, the problem of calculating energy levels and matrix elements of optical transitions for a narrow bandgap epilayer reduces to the solution of the Schrodinger's equation for a constant potential well. The wavefunction can be expressed analytically and is used, for example, as the starting point for investigating device concepts. A model of this type would be helpful for the characterization and application of crescent structures. It turns out that there is only one such model which is separable and the main purpose of this chapter is to describe it.

Background

The progress in the early years of semiconductor physics, and indeed solid state physics in general, can be attributed in large measure to the simplification of the mathematics which arises from translational and rotational symmetries. The bulk properties of semiconductors are understood in terms of the energy dispersion curves which arise from the solution of Schrödinger's equation subject to these symmetries. Even impurities, which break these symmetries, are usually understood as perturbations to the symmetric solutions. Part of the difficulty in calculating the properties of quantum wire structures is that by definition they involve breaking these symmetries strongly along two axes.

A partial answer to the question of how to deal with quantum wires can be realized by analogy with quantum wells. Since their inception in the early 1970s, quantum wells have been understood in terms of the effective mass approximation (EMA) - also known as the envelop function approximation (EFA)[68]. In the EMA the quickly and slowly varying parts of the wavefunction are separated. The quickly varying part gives rise to bulk-like dispersion curves. The slowly varying part gives rise to an effective mass Schrödinger equation with constant potential wells corresponding to the difference in bandgap between barrier and well materials. (The total difference

must be apportioned between conduction and valence bands.) The continuity equations for the derivative of the wavefunctions are modified to include the effective mass for each layer.

Within this approximation there are at least three distinct means of laterally confining carriers. The first and most obvious, but in practice most difficult to fabricate, is to have a lateral bandgap variation alone. Another means - the primary means here, also used to great effect in 0D[40,41,42] - is to have a local thickness fluctuation, such that a particle is confined in a bulge. The last is quite subtle. A particle can be confined in a bend.[69] All these effects may be combined in a true wire structure. The analytic model to be described below includes the second and third effects, though it is not possible to separate the contribution of each. The model could be extended to include changing potentials, but that is not pursued here. It should be emphasized that the bulge and bend confinement mechanisms are purely quantum mechanical in origin, in that they rely on the particle satisfying a wave equation. Only the lateral bandgap variation has a classical analog. It is worth elaborating on these effects.

The bulge confinement is readily understood. Loosely speaking the Schrödinger equation implies that energy eigenstates increase with increasing curvature of the wavefunction. In a quantum well, for example, the

wavefunction is forced to bend in such a way as to disappear in each barrier layer. If the well thickness is increased the wavefunction does not have to curve as much and so the energy is lower. If there are local thickness fluctuations, for example in a quantum well, it is expected that the wider sections would therefore be more likely to contain the particle. This is the bulge confinement.

This effective bandgap approximation was apparently first used and extended by Kapon as a means of estimating energies in crescent structures.[70] To make this more quantitative, an elementary description of an effective binding energy resulting from the bulge will be used.[71] Consider an infinite square well of thickness a. It will have the ground state of $E_a=\hbar^2\pi^2/2ma^2$. If the well is made thicker by δa the ground state will drop to $E_{a+\delta a}=\hbar^2\pi^2/2m(a+\delta a)^2$. Suppose the well is narrow except for a small region which is thicker. Then a particle can be bound to the thicker region (in an approximate sense) by

$$E_{bulge} = E_{a+\delta a} - E_a = \frac{\hbar^2\pi^2}{2m}\left[\left(\frac{1}{a+\delta a}\right)^2 - \left(\frac{1}{a}\right)^2\right] \approx -E_a \cdot 2\frac{\delta a}{a} \qquad (15)$$

where the approximation assumes small δa. Note that a T-shape or cross shape is represented by (15) with $\delta a \to \infty$.

The bound state in a bend is a consequence of the Helmholtz equation which was apparently not fully appreciated until the late 1980s. Consider a quantum well that is deformed slightly so that its width (measured perpendicular to the surface tangent) remains constant. There is no bulge. Such a structure has been proved to support a bound state.[69] The proof was beautifully generalized by Goldstone and Jaffe[72] showing that, "A bulge corresponds to attraction, and in one dimension an attraction no matter how weak suffices to produce binding." (Even before these papers were published it was known that a broad, smooth curve could be mapped into a tangential coordinate system wherein the curve could be represented by an attractive potential scaling as the inverse of the radius of curvature squared. An elegant proof is given in ref. [73]. More recent work extends this approximation to more general geometries.[74])

A limiting case might be a right angle, though it should be recognized that the distance between the vertices of the potential barriers is wider (by $\sqrt{2}$) than the perpendiculars to the tangent, and so includes a bulge in some sense. In that case, numerical simulations and experiments in the microwave analogy[75] show that the ground state energy lies 7% below the lowest quantum well level. Therefore, in approximate terms, the maximum bend binding energy is

$$E_{bend} \approx -0.07 \cdot E_g \qquad (16)$$

where E_g is the quantum well ground state.

A comparison of (15) and (16) suggests that although the bend and bulge binding energies can be of similar magnitudes, the bulge binding energy can be made even an order of magnitude stronger.

A number of approximations are given in the references cited above. It would be useful if an exact analytic solution could be found to represent geometrical quantum confinement. What is needed is an analytic solution on a constricted geometry. As it turns out, there is one and only one coordinate system (Figure 45) with a constriction (i.e. a local increase in width within

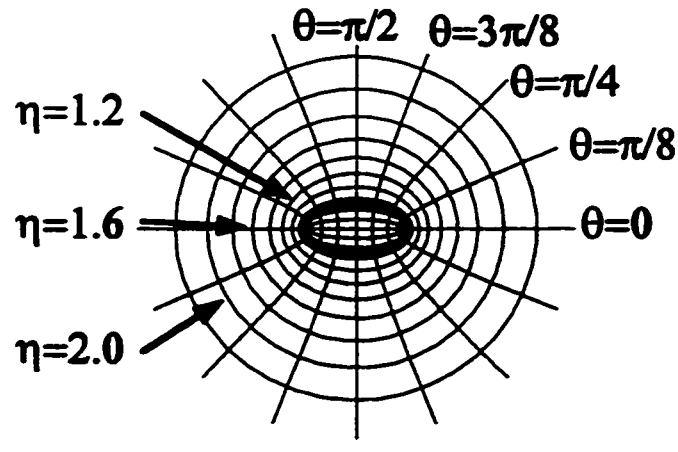


Figure 45. The elliptical cylindrical coordinate system.

The *xy* plane is covered by the quasi-radial coordinate η, and the quasi-angular coordinate θ. The lines of constant η are ellipses with of interfocal distance $2s$. Lines of constant θ are hyperbolae. The *z*-axis, out of the plane of the page, is the usual Cartesian one.

narrower width regions) for which the Schrödinger equation separates: the elliptical cylindrical coordinate system. The main difficulty in dealing with the solution is that the eigenfunctions are not calculated in most computational packages.[**]

After considerable effort was expended on this solution the author learned that this coordinate system had been used once to represent quantum wire structures.[76] However, in that case the constricted geometry was not used - instead only a flattened structure was investigated - and it is difficult to meaningfully define geometrical confinement in that case. Furthermore, a table lookup method was used for the eigenfunctions, which limited the calculations which could be performed to a handful of energy levels for a few specific shapes. Still, features of the solution presented below can already be seen in ref. [76], it demonstrated the usefulness of using a fixed cross-sectional area for the calculation, and it represents a special case of the following model.

In what follows the elliptical cylindrical solution will be developed fully to represent geometrical quantum confinement.

[**] It may seem peculiar to suggest that computational problems limit the usefulness of an analytic solution. What is meant is that the solution can be expressed entirely in terms of well known special functions. However, most computational libraries do not compute those special functions. The analytic form is still useful because the properties of the special function are well understood.

Mathieu Function Model

In the case of a stepwise constant potential, as in the standard EMA treatment of a quantum well, the solution of Schrödinger's equation reduces to that of the Helmholtz equation

$$\nabla^2\Psi + \chi^2\Psi = 0 \qquad (17)$$

where $\chi^2=2m(E-V)/\hbar^2$ with m being the carrier effective mass, E the eigenenergy, and V the potential, a piecewise constant function of coordinates. All coordinate systems in which the Helmholtz equation separates are known.[77] In only one such coordinate system in two dimensions do the coordinate lines form a constriction. This is the elliptic cylinder coordinate system (Figure 45), defined by[78]

$$\begin{aligned} x &= s\cosh\eta\cos\theta \\ y &= s\sinh\eta\sin\theta \\ z &= z \end{aligned} \qquad (18)$$

where x and y are the usual Cartesian coordinates, $s>0$ is a constant, $\eta\in(0,\infty)$ is a quasi-radial coordinate, and $\theta\in(0,2\pi)$ is a quasi-angular coordinate. Lines of constant η are ellipses of interfocal distance $2s$ and ellipticity $\varepsilon^2=1-\tanh^2\eta$. Lines of constant θ are hyperbolae. The crescent quantum wire can be represented by half of the ellipse, taking the bulge at $\theta = \pi/2$ to be the wire

region, and truncating the ellipse along the $\theta=0,\pi$ plane. That is, the crescent wire is approximated by half of the elliptical cylinder.

Writing $\Psi(\eta,\theta)=\Theta(\theta)N(\eta)Z(z)$, the separated ordinary differential equations are

$$\partial_\theta^2\Theta+(a-2q\cos 2\theta)\Theta=0 \qquad (19)$$

$$\partial_\eta^2 N-(a-2q\cosh 2\eta)N=0 \qquad (20)$$

$$\partial_z^2 Z+K^2 Z=0 \qquad (21)$$

where $q=\chi^2 s^2/4$ is the in-plane eigenenergy up to a scaling factor, K, the longitudinal momentum (due to the out-of-plane eigenenergy), and the separation constant $a(q)$ is a continuous function of this energy. At any given q there is an infinite sequence of possible $a(q)$. Eq. (19) is known as the Mathieu equation, and Eq. (20) the modified Mathieu equation. Their solutions are the Mathieu and modified Mathieu functions, respectively.

The general solution to the Mathieu equation is a sum of periodic and aperiodic parts.[79] For concentric elliptical shells of constant η, since the angular coordinate must have 2π periodicity, we may ignore the aperiodic solution, leaving

$$\Theta_m(\theta)=A_m\,\mathrm{se}_m(\theta,q) \qquad (22)$$

or

$$\Theta_n(\theta) = B_n \, \mathrm{ce}_n(\theta, q) \qquad (23)$$

where se(z, q) and ce(z, q) are the sine elliptic and cosine elliptic Mathieu functions, respectively, and the sets A_m and B_m are constants. These functions can be calculated by expanding them in Fourier series. The Fourier coefficients are the eigenvectors of a tridiagonal matrix equation linear in q, with eigenvalues $a_n(q)$, the separation constant described above.[80] Details of the calculation are given in Appendix D.

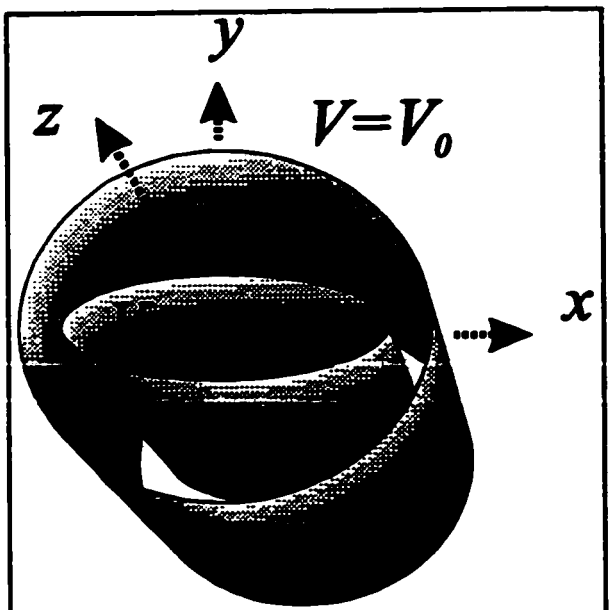


Figure 46. A crescent wire in the elliptical coordinate system.

A double crescent is bounded by constant elliptical coordinate lines in the plane, which should be viewed as an slice of an infinite cylinder coming out of the plane. The potential is zero between hardwall barriers (here $V_0 \to \infty$). The crescent can be made by putting a hardwall barrier along the x-axis.

The quasi-radial solutions to the modified Mathieu equation are

$$N_m(\eta) = C_m \, \mathrm{Sey}_m(\eta, q) + D_m \, \mathrm{Gey}_m(\eta, q) \qquad (24)$$

or

$$N_n(\eta) = F_n \, \mathrm{Cey}_n(\eta, q) + G_n \, \mathrm{Fey}_n(\eta, q) \qquad (25)$$

where in Eq. (24), Sey(z, q) and Gey(z, q) are the first and second modified Mathieu functions corresponding to the sine elliptic quasi-angular solution, and the sets C_m, D_m, F_m and G_m are constants. Eq. (25) is the analogous solution corresponding to the cosine elliptic quasi-angular solution. These functions can be expressed completely in terms of the Fourier coefficients used for the (unmodified) Mathieu functions and Bessel functions of positive integer order.[81]

The simplest possible model for geometrical quantum confinement is a cylindrical shell bounded by concentric ellipses of constant η, with zero potential between the ellipses and infinite potential elsewhere (Figure 46). If the inner ellipse is defined by $\eta=\beta$, and the outer ellipse by $\eta=\rho$, then the boundary condition that the wavefunction vanish on both these ellipses is satisfied, for the sine elliptic quasi-angular solution, when the following function vanishes

$$Z_m(\beta,\rho,q) = \mathrm{Sey}_m(\rho,q)\mathrm{Gey}_m(\beta,q) - \mathrm{Sey}_m(\beta,q)\mathrm{Gey}_m(\rho,q) \qquad (26)$$

provided Sey$_m$(z, q) is nonzero on the boundary. The cosine elliptic solution is completely analogous with the substitutions se→ce, Sey→Cey, Gey→Fey, so it will not be considered separately.

The solution is a function of three geometrical parameters: s, a scaling factor which defines the coordinate system through Eq. (18), β defining the

inner ellipse and ρ defining the outer ellipse. The corresponding inner and outer bounding ellipticities are $\varepsilon_i=\varepsilon(\beta)$ and $\varepsilon_o=\varepsilon(\rho)$. The complete solution can be reduced to two parameters by noting that under an isotropic scaling $x\rightarrow\alpha x$, $y\rightarrow\alpha y$, the Laplacian $\nabla\rightarrow(1/\alpha^2)\cdot\nabla$, so that given a solution $\Psi(x, y)$ to the Helmholtz equation (17) with eigenvalue χ^2, $\Psi(\alpha x, \alpha y)$ is a solution of the isotropically scaled boundary conditions with eigenvalue χ^2/α^2. Therefore by fixing s, the entire parameter space can be explored by varying $\varepsilon_i \in (0,1)$ and $\varepsilon_o \in (0, \varepsilon_i)$. The factor α can be chosen such that the zero potential region has a constant area, A. (Ref. [76] is the case $\varepsilon_i\rightarrow 1$, for sine-like states.) Such a choice is

$$\alpha^{-2} = \frac{\pi \cdot s^2}{A}\left(\frac{\sqrt{1-\varepsilon_o^2}}{\varepsilon_o^2} - \frac{\sqrt{1-\varepsilon_i^2}}{\varepsilon_i^2}\right). \qquad (27)$$

This simplification can be generalized to finite potentials, with the potential height scaled by a factor α^{-2}.

For infinite potentials, the result of such a choice is that the eigenenergies are given by

$$E = \frac{2\hbar^2\pi}{mA}\cdot q \cdot \left(\frac{\sqrt{1-\varepsilon_o^2}}{\varepsilon_o^2} - \frac{\sqrt{1-\varepsilon_i^2}}{\varepsilon_i^2}\right). \qquad (28)$$

The eigenenergies as a function of ε_o for various ε_i are shown in Figure 47 for the sine-like solutions of Eq. (22) and Eq. (24). The different levels

are denoted [m, k] where m refers to the index in Eq. (22) and (24). In the limit of zero ellipticity, in which the Mathieu functions become Bessel functions, m can be viewed as a quasi-angular quantum number, while k can be viewed as a quasi-radial quantum number. This relation can be established by comparing the eigenenergies of the elliptical potential (Figure 47), in the limit as the outer ellipticity goes to zero, with the eigenenergies of a circular drum (Figure 49). However, it is not strictly correct to assign a purely quasi-angular or purely quasi-radial character to either quantum number, since a change in either quantum number affects both quasi-radial and quasi-angular behavior. Figure 48 shows eigenenergies as a function of ε_o for $\varepsilon_i=0.90$ for the cosine-like solutions of equations (23) and (25).

The idea of geometrical confinement becomes meaningful when bulge eigenenergies are referred to the normal one dimensional quantum well energy levels. The thick line on Figure 47 and Figure 48 represents the 1D quantum well ground state for a well of thickness equal to the narrowest space between inner and outer ellipses. A level can be said to be geometrically confined if it lies below this ground state. It is seen that the larger ellipticities have more geometrically bound levels. At $\varepsilon_o \approx 0.15$ the lowest wire level becomes a geometrically bound state. As this ellipticity is increased, one by one more levels become geometrically bound.

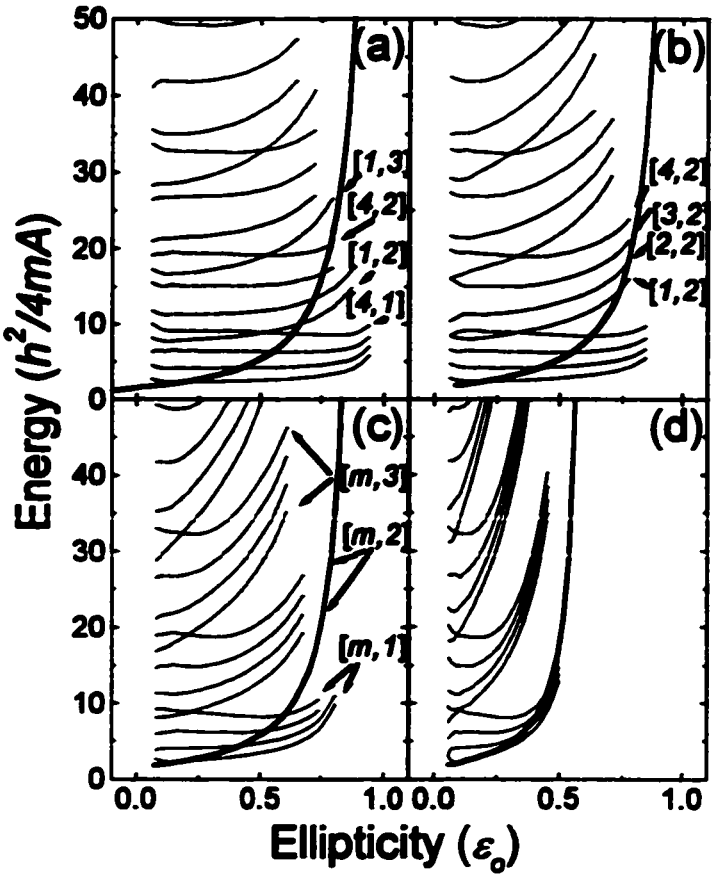


Figure 47. Sine-like eigenenergies of the hard wall elliptical shell potential.

The eigenenergies are plotted for various inner ellipse eccentricities (a) ε_i=0.999, (b) ε_i=0.975, (c) ε_i=0.900, (d) ε_i=0.600. The energy, in terms of the quantities defined in the text, is plotted against the outer shell ellipticity. Levels are enumerated by $[m,k]$ where m is the quasi-angular quantum number, and k is the quasi-radial quantum number. The thick black line represents the ground state of the infinite square well of thickness corresponding to the narrowest distance between inner and outer ellipses.

The energy levels plotted are cut off at high outer ellipticity because the calculation proceeds up to a fixed q. The constant area factor α^{-2} results in the maximum energy calculated being reduced as the ellipticities approach one another. They have also been cut off at very low ε_o for convenience because as $\varepsilon_o \to 0$, the corresponding quasi-radial coordinate $z \to \infty$ and the numerical calculation loses precision.

The sine-like solutions are always nodal at $\theta=0$ and $\theta=\pi$, therefore they also represent solutions of a crescent structure (i.e. half of the elliptical cylinder structure, cut along $\theta=0$ and $\theta=\pi$). Consider the high inner ellipticity sine levels (Figure 47a). The left of the abscissa corresponds to a circular outer surface, the right corresponds to a highly elliptical structure. For small ε_o the solution corresponds to that of a semicircular potential. The quasi-

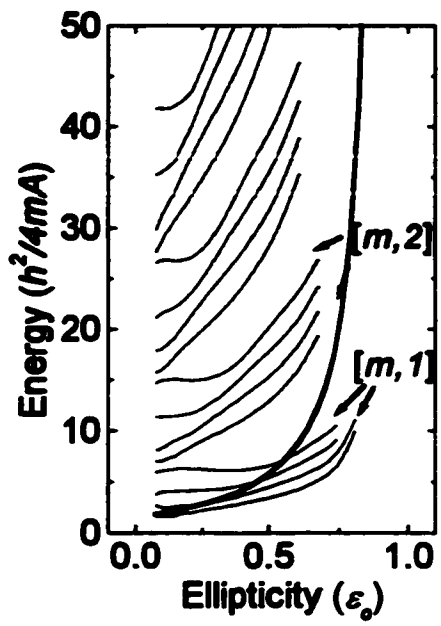


Figure 48. Cosine-like eigenenergies of the hard wall elliptical shell potential.

The eigenenergies are plotted for s inner ellipse eccentricity $\varepsilon_i=0.900$. Here m is the quasi-radial quantum number, and k is the quasi-angular quantum number. The thick black line shows the ground state of the infinite square well of thickness corresponding to the narrowest distance between inner and outer ellipses.

radial and quasi-angular modes are well spaced. As $\varepsilon_o \rightarrow \varepsilon_i$, the quasi-angular m-modes bunch together, losing their distinction, while the quasi-radial k-modes become more strongly separated. The constant area constraint pushes all the energy levels up.

One consequence of the change in level spacing is that some levels must cross. In fact, all but the lowest levels appear to show crossings. For example, Figure 47a shows the $[4,1]$ and $[1,2]$ levels crossing, as seen in the limiting case.[76] More generally, the $[m,k]$ and $[m-3,k+1]$ levels appear to cross. Moreover, they cross at nearly the same ellipticity, in this case, slightly above $\varepsilon_o=0.5$. Such geometrical band crossings are a generic feature of two or more dimensions of confinement: if one dimension of confinement can be controlled independently of the other, a band crossing can be expected as one direction of confinement comes to predominate over another. However, these band crossings occur only in geometrically unbound levels in this model.

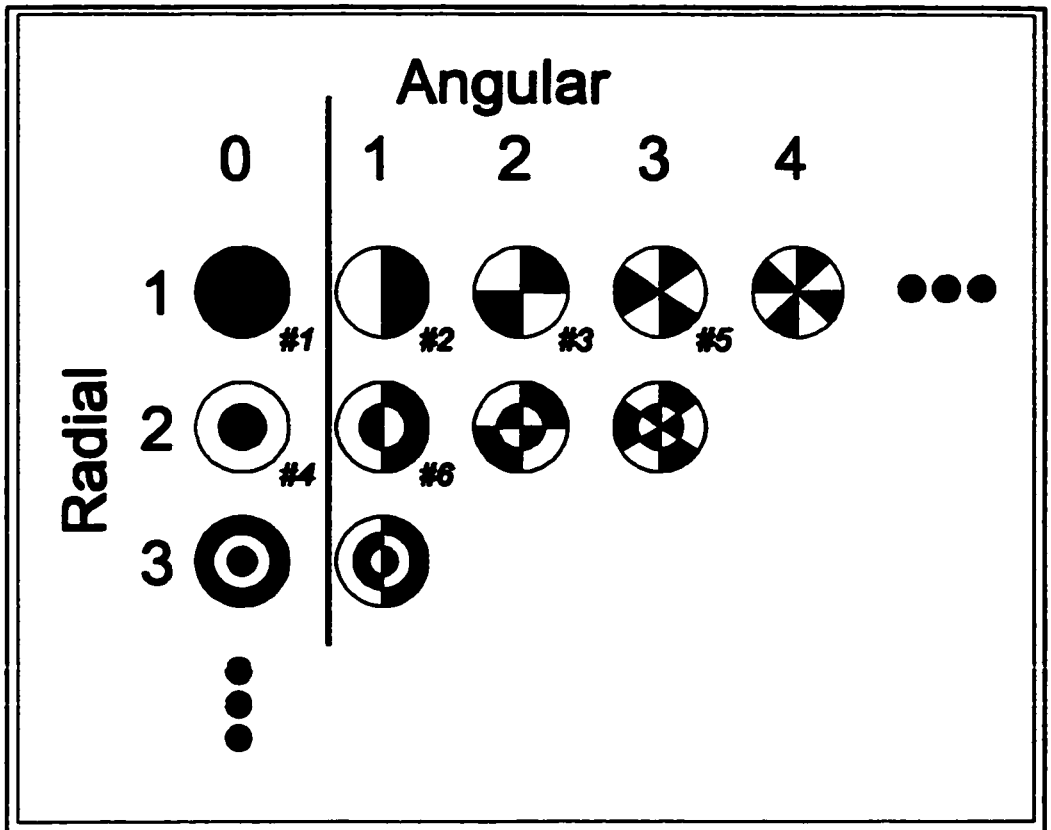


Figure 49. Resonant modes of a drum

The solutions to the Helmholtz equation for a fixed circular boundary are simply the vibrational modes of a drum.[82] Here are shown the first few angular and radial modes, enumerated along the horizontal and the vertical, respectively. The regions of positive displacement (i.e. wavefunction) are dark, negative displacement, light. For the quantum wire solution these represent the extreme limit of zero ellipticity. The vertical line separating the first column is to emphasize that the zero node angular solution is not allowed, since in this case the boundary conditions require that the wavefunction be zero along the θ=0,π plane. The modes are enumerated in order of the magnitude of the energy eigenvalues (italics, below and to the right of the circles). In this way, eliminating the first column, we can identify the first three quantum wire modes as different angular quantum numbers, followed by a different radial quantum number.

As the inner eccentricity is reduced (Figure 47 b,c,d), the same bunching of quasi-angular m-modes can still be seen. The same bands cross one another. However, this crossing occurs at a lower outer ellipticity. At small outer ellipticities, all levels depend only weakly on the inner ellipticity. This

occurs because, as the outer eccentricity approaches zero, the wire outer boundary approaches a circular shape. For fixed area, the coordinate system is such that any ellipse of larger ellipticity (not equal to one) becomes vanishingly small compared to an ellipse for which $\varepsilon_o \to 0$. Therefore, effectively the wire cross-section is circular, with a central, small point of infinite potential.

As with the sine-like solutions, each successive quasi-angular cosine-like mode has additional zeros. However, these zeros are not at multiples of π and therefore none of the cosine-like solutions are solutions of a crescent structure described above. In a sense these cosine-like modes are the bonding states between two crescents while the sine-like solutions are the anti-bonding states with zeros between the two crescents.

The cosine levels show asymptotic degeneracies with the sine levels. If the lowest cosine level is enumerated by $m=0$, while the $[m, k]$ cosine and $[m, k]$ sine levels are degenerate at small ε_o, the $[m-1,k]$ cosine and $[m, k]$ sine levels become degenerate as $\varepsilon_o \to \varepsilon_i$. The cosine levels therefore show a similar level crossing behavior to the sine levels.

Comparison to Experiment

It is of interest to compare the predictions of the idealized model to experiment. Energy levels of crescent quantum wire structures are readily measured by photoluminescence experiments, and an independent estimate of the potential profile may be obtained, for example, using transmission electron microscopy.

Since perfect elliptical structures have not been fabricated, there is some necessary ambiguity mapping the parameters of the experimental structure into this simplified three parameter model. The best mapping would be to digitize the observed interface positions and use a means test such as a least squares fit to find best fit ellipses for the inner and outer walls in the wire area. Given the simplicity of the model, however, a rougher approximation is justified.

Comparison to others

A good test structure for this model may be the crescent quantum wire structure grown by Koshiba et al.[28] and recently modeled with a discretized solution by Inoshita and Sakaki.[83] A 64 meV difference was found between the quantum wire transition and a transition associated with material of the same composition that was not patterned into a wire. The maximum thickness of this crescent was 8.0 nm, and minimum, 4.6 nm. A straightforward

mapping might be to use the maximum and minimum thickness obtained in the elliptical model. The minimum well thickness (l) and maximum well thickness (L) are respectively

$$l = s\left(\frac{1}{\varepsilon_o} - \frac{1}{\varepsilon_i}\right) \tag{29}$$

and

$$L = s\left(\sqrt{\frac{1}{\varepsilon_o^2} - 1} - \sqrt{\frac{1}{\varepsilon_i^2} - 1}\right). \tag{30}$$

The model of Inoshita and Sakaki is a two parameter model while the Mathieu function model is a three parameter model, so there remains one parameter to fix. A good choice, since it represents a minimum interfocal distance might be to take s=10.0 nm, since it is roughly analogous to the width parameter b=10.0 nm in Inoshita and Sakaki. Then the ellipticities are 0.66 and 0.94 respectively, and the energies are calculated in the manner of Figure 47. Using low temperature GaAs effective masses for the electrons and the holes of 0.0665m_e and 0.5m_h gives E_e≈106 meV and E_h≈14 meV. This compares to the calculated E_e+E_h≈80 meV which can be read off Figure 5 of Inoshita and Sakaki. It is the effect of finite barriers that is the primary reason for the discrepancy, not the geometry.

Rather than work through the full finite potential problem, the effect of finite barriers might be estimated by comparing the ground states of infinite

square wells and finite square wells of similar geometry and material parameters, to the one considered here. For example, taking a barrier height of 0.3 eV, effective mass m=0.07, and width 10.0 nm, the Ben Daniel-Duke quantum well has the ground state energy 34 meV [84]. The same structure, for an infinite well has a ground state of 53 meV. The finiteness of the barriers thus amounts to a 64% correction. The same correction applied to our calculation gives 77 meV, in good agreement with Inoshita and Sakaki. To reconcile either model with the measured PL peak energies it is necessary to include excitonic effects.

Comparison to InP V-Mesas

Having seen how this simple model compares to others' work, it can be applied to samples like those described in Chapter 3. The parameters estimated from the SEM photographs of a crescent were for a nominally 8.5 nm (100) growth, 6.0 nm for the bulge, 2.7 nm for the minimum width and 30 nm for the radius of curvature. It should be emphasized, again, that there was a lot of variability between layers, and most certainly, along the mesa. This structure is narrower, but has a larger radius of curvature than the one discussed above. Applying the same mapping as used in the previous section, these parameters imply ellipticities of 0.926 and 0.855. The area of the crescent is 308 nm^2. Using for the InGaAs an electron and heavy hole

effective mass of 0.04 and 0.36 m_e, respectively, results in a ground state of E_e=266 meV and E_{hh}=29 meV which must be added to the bandgap of the InGaAs. To fit the lowest PL peak of 780 meV the composition would have to be ~94% indium. As shown above, though, finite bands would give a more realistic (smaller) indium content.

The model provides more information, which can explain some features of the PL. First, the radial modes are widely spaced, with the second radial mode 0.79 eV higher. Only the first two radial modes have geometrically bound states. Although the spacing is too large, as expected because of the infinite potential walls, this is consistent with two main peaks in the PL. Also, between each radial mode are dozens of angular modes, the average spacing of which is about 25 meV. Allowing them to be even closer together, because the potential is finite, the many angular modes are probably part of the explanation of the breadth of the PL. (Other factors include the obvious roughness, and possible band filling.)

Conclusion

Above, the only separable solution for geometrical quantum confinement was described. The machinery required to calculate energy levels and wavefunctions was also described. The model was compared to realizable

structures, and seen to compare well to non-analytic models currently in use. The model overestimates energies, particularly for narrow structures, because of the hard wall barriers. But since this model is easy to use, it is of interest to apply it to calculate quantities which are more difficult to obtain in other models. For example, for the photoluminescence of quantum wells it is necessary to account for the substantial redshift (*~10* meV) caused by the exciton binding energy. This binding energy can only be increased by the additional quantum wire confinement here. It is worthwhile to go beyond the cylindrical wire approximation used by others [83,85] and estimate this in the elliptical-cylinder coordinate system. In fact, such a calculation is crucial for very small structures since, as explained below, in 1D the exciton binding energy diverges as the radius of confinement goes to zero. The machinery developed in this chapter will be used in what follows to better account for the exciton.

Chapter 5

EXCITONS IN QUANTUM WIRES

Introduction

Having developed a simple model of crescent structures, this chapter is a first step towards making the model more realistic. There are many ways to improve the model, for example, to include dispersion in k-space, instead of considering only flat band conditions, to include multiple bands, not just the conduction and valence band, and to allow them to interact, and to consider finite wells. These are important to characterization, however, they involve the introduction of many parameters, and would not likely show effects which are distinctly one-dimensional. Another possibility is to look at the exciton in these structures. In real heterostructures excitons are typically dominant in the optical spectra. Moreover, their binding energy can decrease significantly with increasing confinement. In fact, in one dimension, as the radius of confinement drops to zero, the binding energy diverges. By studying the exciton the model can be made more realistic, but at the same time, we are focusing on a key problem in one dimension. In this chapter, then, the model

developed in the previous chapter is applied to estimate the exciton binding energy.

Background

A single particle picture is not sufficient to explain the observed spectra of even bulk semiconductors. It is necessary to consider the effects of collective excitations, whereby through particle-particle interactions the carriers are able to reduce their total energy. Probably the simplest and most important such excitation is the exciton. Though it should more correctly be viewed as a true collective excitation - i.e. involving many carriers, it can be mapped onto a two particle picture.[86] In this picture the exciton is viewed as an electron-hole pair bound by the Coulomb interaction. Although in III-V semiconductors the lowest conduction band level is derived from an *s* shell and the highest valence band is derived from a *p* shell, a simple hydrogenic model for the exciton of the electron hole pair taken from these bands has been used in the bulk case with considerable success, for example in predicting excitonic binding energies.

The application of such a model to confined systems is more difficult. The difficulty stems from the fact that the Coulombic term in the Hamiltonian has a spherical symmetry in the interparticle separation. The confining potential

has no such symmetry. Even for a simple quantum well, for which there is confinement only along one axis, it becomes necessary to resort to approximate methods.

Perhaps the simplest way of accounting for the exciton is to use the variational method. A variational calculation provides a means of estimating the excitonic binding energy without having to know the exact solution. In fact, insofar as heterostructures are concerned, variational calculations are the method of choice for estimating binding energies.[85,87,88,89] The calculations always give a lower bound on the ground state energy. The closeness to the actual ground state energy is determined by how close the trial function is in the abstract Hilbert space to the true ground state. By changing a variational parameter in the trial solution a trajectory is followed through this space, and of course, the lowest energy point on this trajectory represents the best solution. With more than one variational parameter, a subspace thought to approximate the exact solution may be explored. Fortunately, even a fair approximation to the wavefunction tends to give a good approximation to the ground state energy.[90]

Having already proposed a basis for the single particle states, below they will be used to build the variational solution for the exciton. First, however, a

digression regarding the Coulomb interaction in one dimension highlights the importance of such a calculation.

Coulomb Interactions in One Dimension

If the exciton is an essential spectral feature in bulk and in two dimensions, it is even more important to treat realistically in one dimension. In the extreme limit of true one dimensional confinement - that is, solving the Schrödinger equation in one spatial dimension only, the Coulomb interaction gives rise to a singularity implying that the exciton has *infinite* binding energy.[85] In a real structure there can only be quasi-1D confinement, meaning the wavefunction has a finite extent in the confined directions. This alone is enough to remove the singularity, and some simple reasoning will be used to explain the problem, and its resolution, below. However, this underscores the fact that the exciton binding energy has the potential to be quite large as the confinement becomes stronger and stronger.

The Coulomb interaction energy, since it varies inversely with interparticle spacing, is singular for zero interparticle spacing. Treated semi-classically, this singularity can be avoided by allowing the two mutually attractive particles to orbit one another - they never collide with one another, so the singularity is avoided. Of course, such orbits may be possible in three or two

dimensions, but if the particles are constrained to travel along a fixed axis (i.e. if they are one-dimensional), they can only fall together and upon colliding, infinite Coulomb attraction ensures they never separate.

The purely quantum mechanical resolution to the problem arises because the probability densities are smeared out and physical parameters must be obtained as integrals over that probability density. In fact the probability density of electron and hole are allowed to overlap in space, but the mathematical requirement to integrate over space eliminates the singularity. The Coulomb interaction still contributes a $1/r$ term to the Hamiltonian, but integration in 3D gives rise to a Jacobian factor r^2 which eliminates the singularity. In 2D a Jacobian factor of r eliminates the singularity. In 1D it persists.

In both pictures it is clear that for quasi-1D confinement, that is confinement to a tube of narrow cross-section, the singularity is avoided. However, it should also be clear that as the cross-section of a wire is reduced, the excitonic binding energy should increase until, in the limit, it is infinite. This underscores the importance of correctly accounting for this energy for quantum wires.

Non-interacting Exact Solution

In the absence of any Coulomb interaction, for hard wall confinement the electrons and holes separately satisfy the single particle Schrödinger equation. For example, the electronic wavefunction is an eigenstate of the Hamiltonian

$$H_c = \frac{\mathbf{p}_c^{\,2}}{2m_c} + V_c(\eta,\theta). \tag{31}$$

Here the subscript c represents the conduction band, $\mathbf{p}$ is the momentum operator *for three dimensions* and V represents the potential profile. In the case of hard walls, the potential is zero inside the crescent, infinite outside, and implies that integration over all space can be replaced by integration over the crescent only (because any wave function must be zero outside the crescent). For hard wall potentials the solution was seen in the previous chapter to be the product of the Mathieu and modified Mathieu functions that vanishes on the hard wall boundary. The eigenstate corresponding to the above Hamiltonian is then, inside the crescent

$$\Psi_c = \frac{1}{C}\Theta_c(\theta_c)N_c(\eta_c)e^{ik_z z}, \tag{32}$$

and zero outside the crescent. A positive z traveling solution has been chosen, and C is a normalization constant such that the integral of the product of the quasi-angular and quasi-radial solutions is unity. It is not meaningful to

normalize the z dependent factor since it is a traveling wave. The hole solution is treated in a similar fashion here.

Excitonic Hamiltonian

Turning on the Coulomb interaction, the total Hamiltonian of the excitonic calculation is the sum of kinetic energy terms for the electron and hole, potential energy terms, and the Coulomb attraction term.

$$H_{exciton} = E_{gap} + \frac{\mathbf{p}_c^2}{2m_c} + V_c(\eta,\theta) + \frac{\mathbf{p}_v^2}{2m_v} + V_v(\eta,\theta) - \frac{e^2}{\kappa|\mathbf{x}_e - \mathbf{x}_v|} \qquad (33)$$

Here κ is the static relative permeability of free space, and e the elementary charge. The units are such that $4\pi\varepsilon_o=1$. Let the z-axis represent the translational symmetry axis of the wire. Then it is convenient to write the Hamiltonian explicitly in terms of this axis. Considerable simplification obtains from the use of relative and center of mass coordinates, instead of separately locating each particle (see e.g. Bastard[91]). Defining z and Z through

$$z = z_c - z_v \qquad (34)$$

$$(m_c + m_v)Z = m_c z_c + m_v z_v, \qquad (35)$$

and for notational convenience setting

$$\Delta^2 = (x_c - x_v)^2 + (y_c - y_v)^2 \qquad (36)$$

$$\mu = \frac{m_e m_v}{m_e + m_v} \tag{37}$$

$$M = m_e + m_v \tag{38}$$

Then the Hamiltonian may be rewritten

$$H_{exciton} = E_{gap} + \frac{\mathbf{p}_{e,2D}^2}{2m_e} + V_e(\eta,\theta) + \frac{\mathbf{p}_{v,2D}^2}{2m_v} + V_v(\eta,\theta) - \frac{e^2}{\kappa\sqrt{\Delta^2 + z^2}} - \frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial z^2} - \frac{\hbar^2}{2M}\frac{\partial^2}{\partial Z^2} \tag{39}$$

where we have written the z-direction momentum operators explicitly in terms of derivatives. Having written out the Hamiltonian in this form the trial solution can be constructed.

Trial Solution

The estimate of the variational energy will only be as good as the trial solution. Therefore the trial solution should be picked with some care. However, the main purpose here is to demonstrate the utility of the simple solution, and therefore the simplest reasonable trial solutions will be used.

In the previous chapter the solution to the non-interacting part of the Hamiltonian was already found. The solutions were shown to be expressible as Mathieu functions. It is natural then to use these products in the excitonic wavefunction. If the excitonic wavefunction has the single particle solutions

as factors then it is assured that the excitonic wavefunction vanishes at the hard wall barriers.

Physically, because of the translational symmetry along the z-axis, the exciton can be expected to move freely in this direction. Therefore a traveling wave dependence on the center of mass Z coordinate can be used.

Since the Coulomb term is attractive, it seems reasonable to assume that the excitonic wavefunction gradually increases as the hole and electron are brought closer together. If there were no potential variation the exciton would form a hydrogenic state, with the wavefunction exponentially decreasing from zero interparticle separation. A reasonable trial solution might include such a hydrogenic factor in interparticle separation. However, for simplicity, a separable solution in z might be used. With these considerations the trial solution may be represented by

$$\Psi = \frac{1}{C}\Theta_c(\theta_c)\Theta_v(\theta_v)N_c(\eta_c)N_v(\eta_v)f(z)e^{iKZ}, \qquad (40)$$

where C is a normalization constant, chosen such that the probability integrates to unity (excluding the traveling wave Z which, by definition, cannot be normalized), K is the center of mass momentum and $f(z)$ will be defined below.

In the case of a quantum well the analogous separation does very well for narrow structures.[92] It is important to remember, however, that with this

choice the in-plane Coulombic interaction is not explicitly part of the variational solution. For wide structures variational solutions which represent this interaction will be more accurate. However, the separation allows for an additional step of integration to be made analytically, and so reduces the computational difficulty. For the purposes of this work, computation time was reduced to minutes from hours because of the extra dimension of analytic integrability.

It remains to choose a suitable $f(z)$. It must vanish at infinity in both directions. Since the potential is symmetric under reflection $z \rightarrow -z$, the same symmetry should be expected of the wavefunction. Finally, since the Coulombic attraction increases as the relative separation between electron and hole decreases, it seems reasonable to assume that the function is peaked at zero interparticle separation, decreasing continuously away from this peak. One may draw inspiration from the exponential form of the hydrogenic wavefunction and try

$$f(z) = e^{-|z|/\lambda} \qquad (41)$$

where λ is in a sense a Coulomb interaction length, and can be used as a variational parameter. As will be shown below, this choice requires some caution due to the discontinuity in the z momentum across zero interparticle

separation (Appendix E). A simpler choice meeting the criteria above is a Gaussian form

$$f(z) = e^{-z^2/\lambda^2}, \tag{42}$$

where λ again is a variational parameter.

Variational Calculation

The task is to calculate

$$E[\lambda] \equiv \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geq E_0, \tag{43}$$

where E_0 is the exact ground state.[93] Obviously it is the closest E to the exact ground state which is of interest, so λ is chosen to minimize the vector product. It is important to note here that the inner product implies integration over all coordinates, and so if elliptical cylindrical coordinates are used, the appropriate Jacobian must be introduced to the integrand as follows:

$$\int dA = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy = \int_0^{\infty} d\eta \int_0^{2\pi} d\theta \; s^2 \left(\cosh^2 \eta - \cos^2 \theta \right). \tag{44}$$

Using the separable trial solution with the excitonic Hamiltonian in the variational calculation the result is

$$E[\lambda] = E_{gap} + E_{c,2D} + E_{v,2D} - \frac{e^2}{\kappa}\left\langle \frac{1}{\sqrt{\Delta^2 + z^2}} \right\rangle \Big/ \langle \Psi | \Psi \rangle \qquad (45)$$
$$+ \left\langle -\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial z^2} \right\rangle \Big/ \langle \Psi | \Psi \rangle + \frac{\hbar^2 K^2}{2M}$$

Here the single particle eigenenergies $E_{c,2D}$ and $E_{v,2D}$ have been used. The fourth term, representing the Coulomb interaction, will be examined below. The fifth term represents the kinetic energy of the electron and hole as they are forced together. It can be evaluated without difficulty for the Gaussian trial solution, but requires caution for the decaying exponential. The last term is the center of mass dispersion relation. Only the fourth term (implicitly, through the wavefunction), and the fifth term are functions of the variational parameter. Rather than minimize the total energy, equivalently the exciton binding energy (at $K=0$)

$$E_{binding}[\lambda] = \frac{e^2}{\kappa}\left\langle \frac{1}{\sqrt{\Delta^2 + z^2}} \right\rangle \Big/ \langle \Psi | \Psi \rangle - \left\langle -\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial z^2} \right\rangle \Big/ \langle \Psi | \Psi \rangle \qquad (46)$$

can be maximized.

Evaluation of the Exciton Binding Energy

For brevity, define the (boundary condition dependent) integral operator

$$\int dA\phi^2 \equiv \int\int dA_c dA_v \Theta_c^2(\theta_c)\Theta_v^2(\theta_v)N_c^2(\eta_c)N_v^2(\eta_v). \qquad (47)$$

The exciton binding energy will be expressed in terms of this integral operation.

Exponential Solution

Using integral tables[94] the expectation value of the Coulomb term can be integrated with respect to z, and the denominator's integral with respect to z is trivial. The kinetic term, despite appearances, requires some care. The result for the kinetic term is expressed here, and its derivation is left for Appendix E.

$$\langle \Psi | \Psi \rangle = \frac{1}{C^2} \int dA \phi^2 \int_{-\infty}^{\infty} e^{-2|z|/\lambda} dz = \frac{1}{C^2} \lambda \int dA \phi^2 \tag{48}$$

$$\left\langle \frac{1}{\sqrt{\Delta^2 + z^2}} \right\rangle = \frac{1}{C^2} \int dA \phi^2 \int_{-\infty}^{\infty} dz \frac{e^{-2|z|/\lambda}}{\sqrt{\Delta^2 + z^2}} \tag{49}$$

$$= \frac{1}{C^2} \pi \int dA \phi^2 \left[\mathbf{H}_o\left(\frac{2\Delta}{\lambda}\right) - Y_o\left(\frac{2\Delta}{\lambda}\right) \right]$$

$$\left\langle -\frac{\hbar^2}{2\mu} \frac{\partial^2}{\partial z^2} \right\rangle = \frac{1}{C^2} \frac{\hbar^2}{2\mu\lambda} \int dA \phi^2 . \tag{50}$$

Above, $\mathbf{H}_o$ is the Struve function, and Y_0 the Neumann function (i.e. the Bessel function of the second kind). It is unlikely, except perhaps for some special limiting case, that products of the Struve function and Mathieu functions could be integrated analytically, so the problem reduces to the numerical maximization of $E_{binding}$, which can be written

$$E_{binding}[\lambda] = \frac{e^2}{\kappa} \frac{I[\lambda]}{\lambda} - \frac{\hbar^2}{2\mu\lambda^2} \tag{51}$$

with

$$I[\lambda] = \frac{\pi \int dA \phi^2 \left[\mathbf{H}_0\left(\frac{2\Delta}{\lambda}\right) - Y_0\left(\frac{2\Delta}{\lambda}\right) \right]}{\int dA \phi^2} \tag{52}$$

Since all the functions are well defined and lack singularities apart from at $\Delta=0$ for the Struve and Bessel functions, the integral can be straightforwardly discretized and numerically integrated.

Gaussian Solution

The Gaussian solution is easier to evaluate, lacking the discontinuity in momentum of the exponential solution at $z=0$. Again, with tables the Coulomb term can be integrated with respect to z. The denominator and the kinetic term are standard Gaussian integrals in z.

$$\langle \Psi | \Psi \rangle = \frac{1}{C^2} \int dA \phi^2 \int_{-\infty}^{\infty} e^{-2z^2/\lambda^2} dz = \frac{1}{C^2} \lambda \sqrt{\frac{\pi}{2}} \int dA \phi^2 \tag{53}$$

$$\left\langle \frac{1}{\sqrt{\Delta^2 + z^2}} \right\rangle = \frac{1}{C^2} \int dA \phi^2 \int_{-\infty}^{\infty} dz \frac{e^{-2z^2/\lambda^2}}{\sqrt{\Delta^2 + z^2}} \tag{54}$$

$$= \frac{1}{C^2} \int dA \phi^2 e^{(\Delta/\lambda)^2} K_0\left(\left[\frac{\Delta}{\lambda}\right]^2\right)$$

$$\left\langle -\frac{\hbar^2}{2\mu} \frac{\partial^2}{\partial z^2} \right\rangle = \frac{1}{C^2} \frac{\hbar^2}{2\mu\lambda} \sqrt{\frac{\pi}{2}} \int dA \phi^2 \tag{55}$$

Above K_0 is the Macdonald function (i.e. the modified Bessel function of the third kind). Again, it is unlikely that the Coulomb term can be simplified any further analytically, so

$$E_{binding}[\lambda] = \frac{e^2}{\kappa}\frac{J[\lambda]}{\lambda} - \frac{\hbar^2}{2\mu\lambda^2} \qquad (56)$$

with

$$J[\lambda] = \sqrt{\frac{2}{\pi}}\frac{\int dA\,\phi^2 e^{(\Delta/\lambda)^2} K_0\left(\left[\frac{\Delta}{\lambda}\right]^2\right)}{\int dA\,\phi^2}. \qquad (57)$$

Here, again, the only singularity occurs when the argument of the Bessel function (i.e. Δ) is zero, therefore the integral may simply be discretized.

Numerical Integration

The integrals in *I* (52) and *J* (57) were evaluated using the extended open trapezoidal rule.[95] The integrals are four dimensional, each coordinate was discretized evenly. The calculation time scales for each dimension separately as *O(N)*, with *N* the number of discrete samples, so with each of the four dimensions sampled the same number of times, the process is *~O(N⁴)*. This means there is a practical limit on the number of samples possible if the calculation is to be completed in a reasonable amount of time. Using more than about 40 points per dimension was prohibitively slow.

A Monte Carlo integral may seem easier for such an integral, however, the use of the analytic crescent solution meant that the boundary conditions

are extremely simple. The trapezoidal rule used here converges as $O(1/N^3)$ - much faster than Monte Carlo convergence of $O(1/\sqrt{N})$. A speed advantage is also gained by precalculating many terms on the mesh and reusing them at each step. To estimate the accuracy of the integration the area of the crescent was calculated using the same numerical routines and compared to the analytic solution, and the result was found to agree to within 10%.

The only technical difficulty in these integrals stems from the fact that $K_0(x)$ and $\mathbf{H}_0(x)$-$Y_0(x)$ are logarithmically singular as their arguments approach zero. (The integrals converge, however. See Appendix F.)

The integrals listed formally in (52) and (57) were split up so that $x=0$ was always on a boundary. For example, the double integral boundaries $\eta_c \in [\beta,\rho]$, $\eta_v \in [\beta,\rho]$ could be split into $\eta_c \in [\beta,\rho]$, $\eta_v \in [\beta,\eta_c)$ and $\eta_v \in (\eta_c,\rho]$. The reason for using the open integration routine is that, by definition, the integrand is never evaluated at the boundary.

A final comment regarding both the Gaussian and exponential integrals. In the above form in which they are written, they may appear to be dependent on β, ρ, and s. However, if λ is written in terms of s, the s dependence cancels, and it is seen that I and J depend on β,ρ and, of course λ.

Results

The results of the variational calculation are shown in Figure 50. The Coulombic and kinetic terms are shown separately. It is seen that for very short variational parameters (i.e. a very compact structure) the kinetic term wins out, whereas for large variational parameters the Coulomb term wins. The competition between the two terms determines the maximum binding energy.

Parameters were chosen to compare with the previous chapter. An electron effective mass of

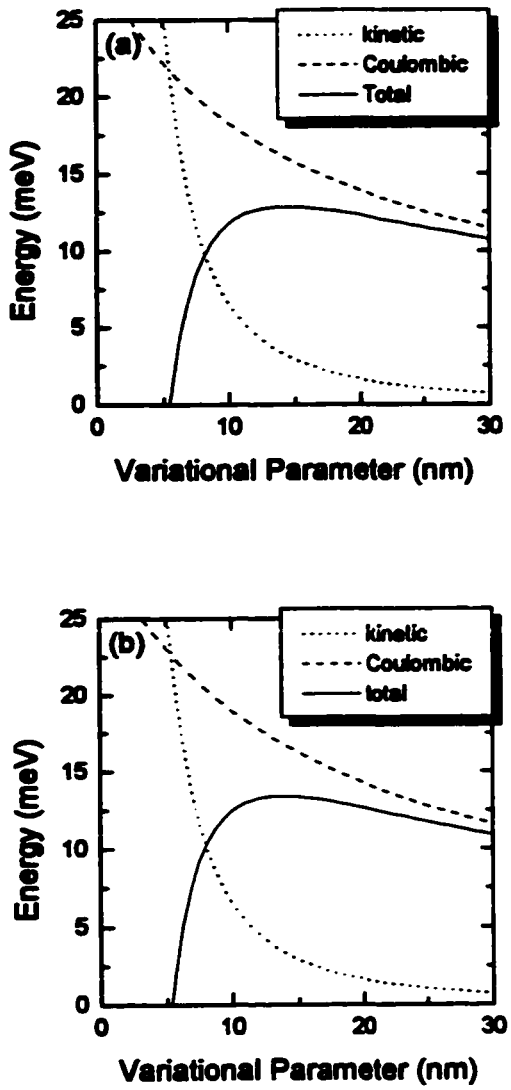


Figure 50. Variational calculation of exciton binding energy

The Coulombic term and the other kinetic term are shown separately, together with the total variational energy for (a) exponential and (b) Gaussian trial solutions. The final variational energy and variational parameter correspond to the maximum of the total energy curve.

0.0665 m_e, a hole effective mass of 0.5 m_e, and a dielectric constant of $13\varepsilon_0$ was used. The inner ellipticity was chosen to be 0.94, outer ellipticity 0.66, and scale factor of 10.0 nm, for a maximum crescent width of 7.8 nm, and a minimum width of 4.6 nm. The constant area eigenstate was $q=4.01$. The crescent was discretized evenly into 40 angular and 40 radial points. The numerical calculation of the area of the ellipse differed from the exact value by less than 8%.

At first glance the two curves appear quite similar, however, it should be cautioned that the two variational parameters in fact have different meanings for the two integrals. For example, the r.m.s. z value of the exponential integral $I[\lambda]$ is $\lambda/\sqrt{2}$, while the Gaussian integral $J[\lambda]$ is $\lambda/2$. The exponential integral gives an excitonic binding energy of 12.9 meV, with a variational parameter of 14.2 nm, while the Gaussian integral gives an excitonic binding energy of 13.3 meV with variational parameter of 13.9 nm.

That both calculations give a similar binding energy is hopeful, suggesting that the calculation may be quite close to the true value. A variational calculation can only give an energy greater than the ground state energy, or in this form, can only give a binding energy lower than the true binding energy. Thus, assuming the numerical calculation is accurate, the Gaussian integral is more valuable since it gives the highest value, and so the binding energy must

be greater than or equal to 13.3 meV. A better bound can be found by including an in-plane Coulombic interaction in the trial wavefunction. The fairly small r.m.s. *z* separation arising from the variational parameters suggests that this may be significant. However, this involves a considerably more complicated calculation, since the fact that the trial function was z-separable eliminated several terms.

Inoshita and Sakaki[83] used an exciton binding energy of ~15 meV, calculated for an infinite barrier circular cylindrical wire.[85] The bulk exciton binding energy was estimated to be 4.2 meV. The exciton binding energy enhancement for a sharp crescent is seen to be quite strong. However, in that work there is no indication of how to determine an effective circular cylinder radius for a given crescent, so it is not possible to attach much meaning to a comparison of the respective wire binding energies.

Finally, the variational calculation can be applied to the idealized indium phosphide structure grown as part of this work. Using again a 6.0 nm bulge, 2.7 nm minimum width, and a 30 nm radius of curvature, using a 40×40 grid, the exponential trial solution gave, with a 23.3 nm variational parameter, a 7.9 meV binding energy and the Gaussian trial solution, with 22.0 nm variational parameter, gave an 8.2 meV exciton binding energy. This is

substantially higher than the ~2 meV exciton binding energy in the bulk for indium rich InGaAs.

Conclusion

The simple analytic formulation for geometric bound states developed in Chapter 4 was exploited to provide a variational estimate on the exciton binding energy for a real crescent structure. The analytic solution makes such calculations considerably simpler and faster than other techniques. It is important to perform the variational calculation of the exciton binding energy in such a realistic coordinate system, rather than the more common circular cylinder coordinates because the energy is unbound as the wire becomes infinitely thin. For very thin crescents in-plane Coulomb interactions should also be included.

CONCLUSION

Summary

The preceding chapters focused on a particular fabrication technique, relying on only a few characterization techniques, and some simple theory. The thesis began with a description of some of the tools available in the fabrication of semiconductor structures, and explained a means of using them to fabricate on scales far smaller than usual. Looking first at gallium arsenide based materials, it was shown that the final structure was very sensitive to growth conditions, in particular, V-III ratio. With some understanding of the growth, it became possible to grow inverted-V crescents, which were seen to be optically active. Although at least as good as can be achieved by other means, it was seen that there was room for improvement in sample roughness, and in the breadth of the photoluminescence. With this in mind, indium phosphide based growth was explored, and, at least on the wider mesas, produced much higher quality structures, both structurally and optically. The need for a means of characterizing crescent structures lead to the

development of the only separable model of geometrical quantum confinement in two dimensions. To make the model more realistic, and to demonstrate its use, the excitonic binding energy, a particularly important quantity in one dimension, was estimated.

Future Possibilities

There are many possibilities for future work on this project.

The simple theoretical model can straightforwardly be extended to finite barriers. It should also be extended to include dispersion in k-space, and multiple valence bands could be introduced. Optical transition selection rules can be obtained relatively easily for an analytic solution and could be calculated. Finally, the Schrodinger equation could be solved self-consistently with the Poisson equation to account for electron-electron interactions at high carrier densities. The Mathieu functions could be used as a basis for such a calculation.

For fabrication of nanostructures, the most fundamental problem here is the V-shaped tip sharpness. This should be minimized. Further studies of mesa tips may allow this size to be reduced from the present ~20 nm radius of curvature to something still sharper. On the other hand, it may suggest that there is a limit to sharpness obtained from this technique. Roughness also

remains an issue, especially for GaAs mesas. Much of the roughness appears to occur after the (100) surface has vanished and only (111)B planes remain. The quality of structures could be improved by studying (111)B growth, and adjusting the growth conditions on completed (111)B mesas accordingly.

There are applications of faceting other than nanostructures that would be a logical outgrowth of the present work. For example, gratings of extremely high precision could be fabricated for applications such as distributed Bragg reflectors.[96] Migration of material from the (111)B to the (100) could be used for bandgap shifting.[97] There may also be applications for the growth of heterostructures on facet planes, for optoelectronic applications (although usually this means facet planes normal to the (100) plane[98]), perhaps, for example, to escape the normal incidence selection rule[99] of intersubband transitions.

Progress in nanostructure fabrication would be faster with better characterization tools. A serious difficulty with quantum wires and dots as compared to quantum wells is that the potential geometry is much more complicated, and compositions may vary in three dimensions. Further use of high resolution microscopy techniques such as TEM, scanning tunneling microscopy (STM) or atomic force microscopy (AFM) could give more detailed structural information. Newer techniques such as near field optical

microscopy suggest the possibility of obtaining information about composition as well. Ideally, characterization would be done *in situ*, while the structure is being fabricated.

In general, it can be said that going from planar to laterally patterned quantum semiconductor structures involves a giant leap in science and technology. In this work some steps were taken in that direction, and, some key issues raised, but the ideal nanostructure is still in the future.

Epilogue

Although the above chapters are concerned primarily with a specific fabrication technique and a specific type of structure, it may suggest some of the broader issues associated with fabrication and characterization of semiconductor nanostructures.

Because human hands grip on the scale of centimetres and human eyes see down scales of microns we have long been denied the possibility of building structures on small scales which have always been accessible to nature. With a variety of techniques, more and more, scientists are able to manipulate structures on the scale of nanometres. Although many technical difficulties arise when trying to work on these scales, the field shows great promise.

Appendix A

FLOW RATES

In CBE, as opposed to MBE, sources are supplied as gases, and the rates with which these gases flow are controlled by mass flow controllers (MFC). It is possible to operate a CBE machine solely by monitoring the flow rates of the gases relative to previous growths. However, it is sometimes useful physically to find the absolute rate of flux or the flux of one source relative to another. This appendix describes the relation between the flow rates as given by the MFC digital readouts and the actual flux of material using simple thermodynamic arguments and using the MFC documentation.[100]

The MFCs measure the gas flow by monitoring the temperature profile of a laminar flow. They are calibrated to give a correct reading for nitrogen gas (N_2). Each mass flow controller is rated at a specific maximum flow rate, given in standard cubic centimeters per minute* (sccm). If nitrogen was the gas, the readout would return the percentage of the full scale - that is the percentage of the maximum flow rate. Thus for a flow of nitrogen,

$$F_{N_2} = M_{MFC} \cdot r \qquad \text{A 1}$$

* The "standard" in sccm refers to Standard Pressure, defined as 760 mm of Hg, and Standard Temperature defined as 0°C.

Table 8. Maximum flow rating for gas lines

Line	*Gas*	M_{MFC}*(sccm)*
AS1	Arsine	20
AS3	Arsine	50
HYD2	Phosphine	50
OM1	TEG + H_2	20
OM2	TMI + H_2	20
OM3	TetraESn	20
OM4	DiEZn	20

Table 9. Gas correction factors

Gas	C_{gas}^{GCF}
Hydrogen	1.011
Arsine	0.67
Phosphine[†]	0.76

where F is the actual flow rate of nitrogen, M_{MFC} is the maximum rating of the MFC, and r is the ratio which corresponds to the percentage of full scale read. (That is, *r=percent/100.*) The maximum rating of the MFC for each gas line on the CBE machine is shown in Table 8.

The sensitivity of the MFC is a function of gas type, however, and a correction must be made for gases other than nitrogen. Typically, the MFC is less sensitive to gases with larger constituent molecules. The flow rate of a gas other than nitrogen may be expressed as

$$F_{gas} = C_{gas}^{GCF} F_{N_2}. \qquad \text{A 2}$$

For some gases the gas correction factors are known. It is tabulated in Table 9 for selected gases. More generally,

[†] The MFC used for phosphine in this work was calibrated for phosphorus, meaning that the gas correction factor was already included in the reading.

$$C_{gas}^{GCF} = 0.3106 \frac{S}{\rho^{gas} C_p^{gas}}. \qquad \text{A 3}$$

where ρ^{gas} is the density of the gas at 0°C, in *g/L*, C_p^{gas} is the specific heat of the gas in *cal/g/°C*, and *S* is a structure factor equal to 1.030 for monatomic gases, 1.000 for diatomic gases, 0.941 for triatomic gases, and 0.880 for polyatomic gases. In summary, combining A 1 and A 2,

$$F_{gas} = C_{gas}^{GCF} M_{MFC}{}^{r}. \qquad \text{A 4}$$

where C^{GCF} and M_{MFC} are found from the above tables or C^{GCF} from relation A 3. For the hydrides the problem is solved. For the organometallics, since a hydrogen carrier gas is used, two gases are involved, and the situation is more complicated.

To understand the situation for the organometallics, it is necessary to consider the thermodynamics of the bubblers (refer also to Figure 4, p15). The bubbler is a heated chamber, containing the solid or liquid which is to be carried, through which flows a carrier gas. The bubbler is maintained at a constant pressure, in this case 40.0 torr. Of that pressure, some fraction may be attributed to the vapor pressure of the organic

Table 10. Vapor pressures of organometallics

Organometallic	*Bubbler T*	*A*	*B*	P_{vapor}*(torr)*
TMI	37.0	3014	10.5	6.27
TEG	30.0	21	8.08	8.86

Source: Morton International CVD Metalorganics

source material. To a very good approximation the vapor pressure of organometallics sources is an exponential in inverse temperature.

$$P_{vapor}(torr) = 10^{B-A/T}. \qquad \text{A 5}$$

Above, the temperature T is in Kelvin. Since the material constants A and B are known for TMI and TEG, knowing the bubbler temperatures the vapor pressures in equilibrium can be calculated. The values are tabulated in Table 10 for the bubbler temperatures used in this work.

The organometallic is carried in "bubbles" which contain vapor and carrier gas in equilibrium. Assuming that both gases are ideal, knowing the volume of a bubble, since its temperature is fixed, the number of molecules it contains can be calculated directly from the ideal gas law. The MFCs are designed so that the fraction of molecules of each gas remains fixed. The MFCs effectively determine the volumetric flow rate dV/dt for a fixed T and P (not necessarily the same as the bubbler T or P). The final result is that the flow rate F is proportional to the pressure in the bubble, with the same proportionality constant for the total flow as well as the partial flow of each constituent. It would seem straightforward to calculate the flow rate of organometallics, except that the MFC has different sensitivity to different gases. This is a peculiarity of the particular system used here, because the MFC is located downstream, after the bubbler. If the MFC were located upstream to the bubbler, the only flow to measure would be hydrogen.

If the MFC reads a total flow R, given by

$$R = \alpha F_{TMI} + \beta F_{H_2}. \qquad \text{A 6}$$

where α and β are the sensitivities of the MFC to *TMI* and H_2 respectively, using the proportionality between flows and pressures,

$$\alpha F_{TMI} = R - \beta\left(1 - P_{TMI}/P_T\right)F_T. \qquad \text{A 7}$$

where the subscript "T" refers to the total flow. Although not exact for mixtures of gases, the sensitivities may be taken to be the gas correction factors, very nearly unity for hydrogen, and typically quite small for a large organometallic molecule.

Since the flow rate of organometallic is typically much less than that of hydrogen gas, a simpler approximation is to assume that the MFC measures only the hydrogen flow rate. Then using a gas correction factor of nearly unity the simpler result of the proportionality is

$$F_{TMI} = \frac{F_{H_2}}{\left(P_T/P_{TMI} - 1\right)}. \qquad \text{A 8}$$

This is the expression used to estimate the organometallic flux in this work.

It may be of interest to convert these flow rates in *sccm* to a number of atoms per second. Using the ideal gas law for Standard conditions

$$\frac{dN}{dt}(atoms/s) = \frac{(101{,}325\ Pa)}{(1.381\times10^{-23}\ J/K)(273.15\ K)} \cdot \frac{1\,m^3}{10^6\ cm^3}\frac{1\,m}{60\ s}F(sccm) \qquad \text{A 9}$$
$$= 4.47\times10^{17}\cdot F(sccm)$$

The molecular beam is wider than the substrate width. For comparison, a 2 inch GaAs (100) wafer requires 5.14×10^{16} atoms to complete a monolayer. Since typical gas flow rates are a few sccm, and typical growth rates are on the order of a monolayer per second, it can be concluded that of the gas entering the chamber, roughly 5 to 10% is incorporated into the sample.

Appendix B

WAFER SPECIFICATIONS

A crystal is a highly ordered structure, having a variety of symmetries, or, alternatively, having a variety of special directions. Since most fabrication processes that a semiconductor crystal is made to undergo are chemical in nature (e.g. epitaxy, etching), they are often orientation dependent. Therefore, for many processes, it is very important to be able to align a wafer uniquely, and be aware of the underlying crystalline structure.

Commercially available (100) GaAs and InP wafers are made of circular slices of a cylindrically ground ingot of bulk crystal.[101] An etch is conventionally used to orient the crystal, and a long major flat is ground along the (0$\underline{1}$1) surface (longitudinally on the cylinder) and a narrower minor flat along the (01$\underline{1}$) (also longitudinally, perpendicular to the major flat). The ingot is sawed into separate wafers and ultimately one side is polished.

Because zinc-blende structure lacks reflection symmetry in the (100) plane, the orientation of the crystal surface is dependent on which side of the wafer is polished. There are two wafer conventions in use: the EJ spec (for $\underline{E}$uropean-$\underline{J}$apanese $\underline{spe}$cification) and the SEMI spec (an American specification). In the EJ spec the minor flat is clockwise from the major flat viewed from above. If the wafer is cut slightly misoriented from the (100) towards this direction planes are preferentially arsenic

terminated. The SEMI spec corresponds to polishing the opposite side of the wafer, and so although the minor flat still lies along the (01$\bar{1}$) plane, misorientation of the top surface towards the minor axis now results in planes that are preferentially gallium terminated. In the SEMI spec, then, the minor flat is located counterclockwise to the major flat. One can tell by inspection of the flats relation to one another which specification is in use and so which direction corresponds to preferentially arsenic or gallium terminated planes.

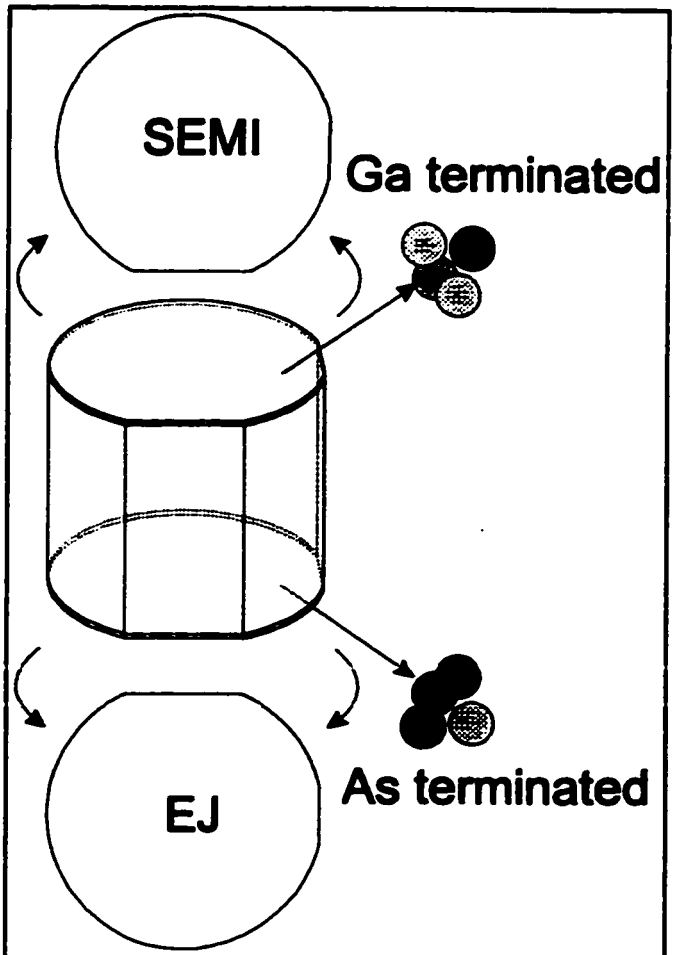


Figure 51. Wafer specifications

An ingot, pictured here is a circular cylinder, with a major (01$\bar{1}$) oriented flat - toward the front - and minor (01$\bar{1}$) flat - to the right - ground into the curved surface. SEMI spec and EJ spec wafers polished surfaces correspond to the top and bottom of this ingot, respectively. Planes tilted toward the minor flat are gallium terminated in the EJ spec and arsenic terminated in the EJ spec, as illustrated schematically by the tetrad of atoms where the dark spheres represent gallium and the light spheres represent arsenic.

Appendix C

PHOTOLUMINESCENCE CHARACTERIZATION CHARTS

InGaAs on InP

A quantum well of InGaAs on InP can be most simply described by two parameters: a width and a composition. Knowing a single photoluminescence (PL) transition energy for a single quantum well alone cannot provide both parameters. Some possibilities to uniquely determine compositions and widths include growing two (or more) quantum wells of different composition or different widths and observing the PL of both, using PL in combination with x-ray diffraction on a superlattice structure, or photoluminescence excitation spectroscopy which provides access to higher level transitions. However, knowing the PL spectrum alone does provide a single constraint on the energy level. Of course, very thick (>20 nm) layers can be considered as the limit of large wells, and in that case one measurement is sufficient to get the composition, with the caveat that care must be taken to account for strain.

In PL, light (typically laser light) of energy larger than the bandgap energy of the barrier material is used to excite carriers from the valence band to the conduction band. Conduction band electrons typically relax non-radiatively into the lowest quantum well level where they can recombine (in a type I semiconductor) by a spatially direct transition into the highest quantum well hole band. The holes likewise relax to the highest hole band. The depth of the well, and hence the energy of the transition depends on the composition

of the well material since InAs is narrower band gap than GaAs. The differences in effective masses and the need for a strained epitaxial layer to accommodate itself to the underlying substrate also affect the transition energy. A complication is that the band offset (the alignment of bands between different materials) is difficult to determine.

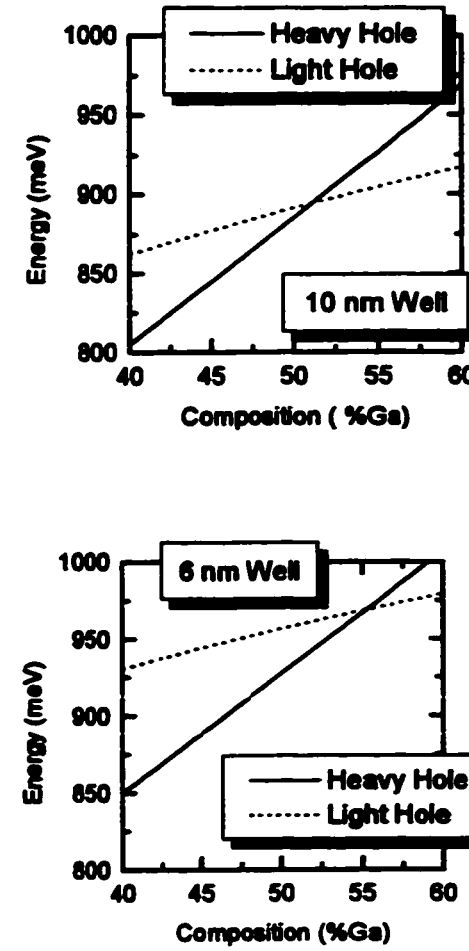


Figure 52 Linear composition dependence of the lowest quantum well transition for InGaAs:InP

Any numbers derived from the PL data for InGaAs on InP rely on a computer program, "Quaternaire", written by Alain Roth, which accounts for strain as a perturbation, and uses phenomenological estimates of bandgaps, offsets, strain coefficients and effective masses from the literature. Quantum well layers are assumed to be unrelaxed. All calculations are for a sample at a temperature of 4.2K. It is possible to demonstrate some rules of thumb which follow from running the program in the regime of interest, namely for compressively strained (high indium content) quantum wells.

The first rule of thumb is that the transition energy varies linearly with composition as long as the heavy hole or light hole transition remains the lowest (Figure 52). Because the material system switches from compressive to tensile strain, there is a transition point

between these two lines which have different slopes. (The heavy hole line shows a larger slope.) This transition does not occur exactly at the transition between tensile and compressive strain however, because lower effective mass of the light hole pushes its transition up.[102](The matching conditions between layers bring effective masses into the calculations of the energies.) The transition point is well width dependent, but is pushed well to higher Ga content than lattice matching, which occurs at 47% Ga.

Throughout the regime of interest for this work (compressively strained layers of

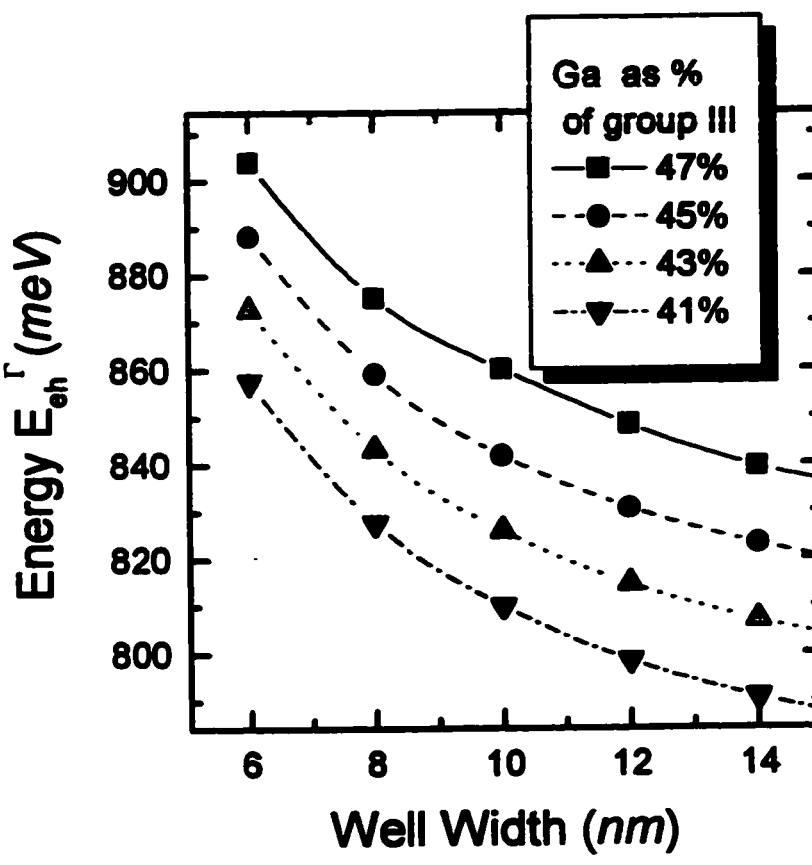


Figure 53. Ground state transition dependence on well width
For InGaAs quantum wells on InP substrates.

InGaAs between ~6 nm and ~12 nm in thickness lattice matched to InP) the transition energy shows nearly an exponential dependence on well width (Figure 53). This second rule of thumb, together with the previous one allows the width and composition dependence to be separated out of what in the full solution is a transcendental equation:

$$E = E_{\infty}(x,w) + E_{QW}e^{-(w-w_0)/\omega} \qquad \text{(C1)}$$

Here x and w represent the fraction of Ga and the well width respectively. E_{∞}, a fitting parameter representing the energy of an unrelaxed infinitely wide layer, is only weakly a function of w and can, to a first approximation, be taken as linear in x. The two fitting parameters w and ω, are roughly constant throughout the regime of interest and represent the width dependence of the ground state energy. In that case, the x and w dependencies are completely separate. Best fits to exponentials yielded the following parameterization for composition as a fraction and width in nanometres: w_0=5.75 nm,ω=4.5 nm,E_{QW}=83 meV, with E_{∞} to be linearly interpolated between 829 meV for 47% Ga and 782 meV for 41% Ga. This can be used to quickly obtain an idea of the composition, but a more precise fitting with the program is necessary where a high level of accuracy is required.

InGaAs on GaAs

To calculate the ground state energy of a quantum well of InGaAs on GaAs a computer program ("Struct.c") was written based on Alain Roth's "Quaternaire". It was possible to run the program once for the entire parameter space available, namely varying composition and well width. Again, all calculations are for a sample at a temperature of

4.2K. Since indium is larger than gallium the quantum wells are always compressively strained, and the heavy hole transition is always the lowest energy transition.

In this program the unstrained lattice constant, strain tensor components, deformation potentials, electron effective mass and dielectric constants were linearly interpolated between known gallium arsenide and indium arsenide values at 5K. The fundamental band gap and spin orbit splitting were quadratically interpolated from known values. The hole effective masses were obtained by linearly interpolating Luttinger parameters and calculating the resulting effective masses. A 66% band offset was applied to the bandgap calculated including strain effects following refs. [103, 104].

The end result of this calculation is shown in Figure 54. Like the InGaAs:InP quantum

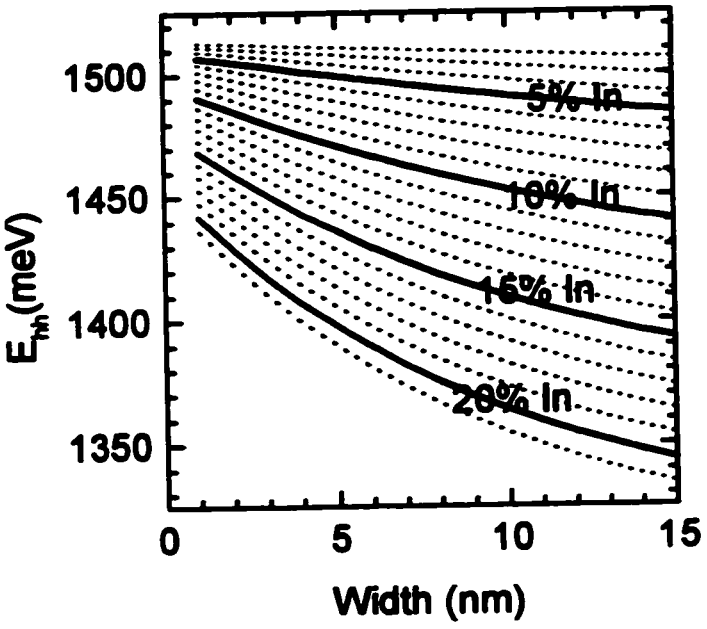


Figure 54 Lowest energy transition of an InGaAs quantum well.

The lowest transition is always corresponds to the conduction band to heavy hole transition at the gamma point in k-space. Lines of constant composition are plotted as a function of well width.

well, throughout the range of interest the can be approximated by an exponential. Furthermore, provided the composition is greater than ~5% less than ~20% (above which the layer is likely to relax even for thin layers, anyway), for a fixed well width the transition very nearly changes linearly with a change in composition. Perhaps this is not surprising since many parameters were obtained via linear interpolation, and those that were obtained quadratically had relatively small bowing parameters.

It should be cautioned that Figure 54 uses simplifying assumptions for the band structure and effective mass following ref. [104] which are not made in the program "quaternaire". Although the results compare favorably for low compositions (<10%), there is an overestimate (~3%) of the composition for high composition (~20%) wells in comparison. Where high accuracy is required, and particularly for high indium content wells, the program "quaternaire" should be used.

Appendix D

MATHIEU FUNCTION CALCULATIONS

There are several ways to calculate these functions.* I will highlight two - first, the way used in this project, which is probably the easiest to code for, and is very efficient if very high precision is not required, and only a few eigenfunctions are required, and second, the eigenvector method, which is more general. (The routines compared well to tabulated values selected at random.)

The Mathieu equation is

$$\partial_\theta^2\Theta + (a - 2q\cos 2\theta)\Theta = 0 \qquad \text{E 1}$$

and its solutions are known as the Mathieu functions. The modified Mathieu equation is the same with the substitution $\theta \rightarrow i\theta$, which is equivalent to changing the "+" to a "-" above, and the "cos" to a "cosh". The important point is that if the Mathieu function is to be periodic in 2π, the characteristic number (a) is a function of q.

In the limit $q \rightarrow 0$ the equation is satisfied by a sinusoid. The first expansions for the Mathieu functions were given as a power series in q with coefficients that include sinusoids in θ. The functions can be divided up into ones that become cosines as $q \rightarrow 0$, and ones that become sines as $q \rightarrow 0$. This can further be subdivided into expansions in even or odd factors of θ. Each type has a different recursion relation to calculate the coefficients

* Note that IMSL provides a FORTRAN library which calculates some Mathieu functions - but at the time of this work it did not calculate modified Mathieu functions.

of the expansion. The recursions can be found in several texts.[105,106,107] One example is given below, for $ce_{2n}(\theta,q)$ which is obtained from the expansion

$$ce_{2n}(\theta,q)=\sum_{r=1}^{\infty}A_{2r}\cos(2r\theta). \qquad \text{E 2}$$

The recursion resulting from substituting (E 2) into Mathieu's equation is

$$\begin{aligned} aA_0-qA_2&=0 \\ (a-4)A_2-q(A_4+2A_0)&=0. \\ (a-4r^2)A_{2r}-q(A_{2r+2}+A_{2r-2})&=0 \end{aligned} \qquad \text{E 3}$$

For arbitrary q the characteristic number can be calculated quickly by substituting each of (E 3) into one another to create an infinite continued fraction for $a(q)$. The continued fraction can be evaluated quickly using the Wallis algorithm.[108] There is an infinite ladder of $a_{2n}(q)$ symbolized by the index $2n$ in (E 2). Fortunately, power series in q and asymptotic expansions in $1/q$ converge rapidly for small and large q respectively. The numerical programs used in this project guess at the initial $a_{2n}(q)$ using these expansions and then use the successive approximations to the continued fraction to refine the estimate.

Knowing $a(q)$, the recursion (E 3) can be used directly to find the coefficients A_{2n}. The rapidly convergent Miller's algorithm[109] was used (in the stable direction of decreasing coefficient $2n$). An extremely small seed number is used for some large A_{2n}'s. The recursion is then followed back to A_0. Coefficients agreed to many decimal places with tabulated values.

One problem remains in the calculations of the Mathieu functions, namely that the sinusoidal terms result in very slow convergence of the series for any θ. Fortunately, the

same recursion can be constructed with several different Bessel function expansions, which converge extremely quickly with θ The most rapidly convergent expansion can be found in the NBS tables.[106] Since the recursion is the same, they rely on exactly the same coefficients. The modified Mathieu functions are calculated using the same coefficients, but with $\theta \rightarrow i\theta$, which amounts to changing some of the Bessel functions J to Y in the most rapidly convergent series.

Having outlined the approach used in this project, a more general method should be described. This has been incorporated into the periodic, unmodified Mathieu function calculations available in the IMSL library. This method stems from the realization that (E 3) and its counterparts for the Mathieu functions other than $ce_{2n}(\theta,q)$ can be written as an eigenvector problem with

$$\underline{\mathbf{T}}(q)\cdot\underline{\mathbf{A}} = a\underline{\mathbf{A}}. \qquad \text{E 4}$$

where the infinite matrix $\underline{\mathbf{T}}$ is tridiagonal, and the elements of the infinite vector $\underline{\mathbf{A}}$ are the coefficients A_{2r}. Very fast routines exist to calculate the spectra of eigenvalues a_{2n} and the components of the eigenvector A_{2r}, for example in the LAPACK library.[110] Some routines can be sped up by a guess at the eigenvalue or components such as the one used in the other calculation method.

Ultimately this eigenvector approach will be faster for obtaining the entire spectrum of eigenvalues simultaneously, as well as determining the components to a higher degree of accuracy. The approach used for this project has lower overhead in addition to being easier to code in its entirety, so it is fast for finding a particular characteristic number, or for finding the coefficients knowing the characteristic number.

There are many references which describe the properties of the Mathieu functions and list identities which they satisfy.[105,106,107,111]

Appendix E

THE USE OF DISCONTINUOUS TRIAL FUNCTIONS

There is a danger in using trial functions with discontinuous derivatives in variational calculations, namely that such functions can make the kinetic energy operator non-Hermitian.[112] Failure to exercise caution with such functions could result in a negative kinetic energy, which would most certainly ruin a variational estimate of an eigenenergy. In Chapter 5 an exponential trial solution with a discontinuous derivative was tested. In this appendix the kinetic energy term will be correctly evaluated.

Quantity	x<0	x>0
$\lvert\varphi\rangle$	$\sqrt{k}e^{kx}$	$\sqrt{k}e^{-kx}$
$\mathbf{K}\lvert\varphi\rangle$	$ik^{3/2}e^{kx}$	$-ik^{3/2}e^{-kx}$
$\langle\varphi\rvert\mathbf{K}$	$-ik^{3/2}e^{kx}$	$ik^{3/2}e^{-kx}$
$(\langle\varphi\rvert\mathbf{K})(\mathbf{K}\lvert\varphi\rangle)$		k^2
$\mathbf{K}^2\lvert\varphi\rangle$	$-k^{5/2}e^{kx}$	$-k^{5/2}e^{-kx}$
$\langle\varphi\rvert$	$\sqrt{k}e^{kx}$	$\sqrt{k}e^{-kx}$
$(\langle\varphi\rvert)(\mathbf{K}^2\lvert\varphi\rangle)$		$-k^2$

Table 11. Application of the momentum operator.

The one-dimensional trial solution of interest is

$$\varphi(x) = \sqrt{k}e^{-k|x|} \qquad \text{E 1}$$

where k is a constant x is the spatial coordinate. Although this solution is continuous, the derivative is not continuous at $x=0$. If the momentum operator is $\mathbf{P} = \hbar\mathbf{K}$ then in the position basis $\mathbf{K} \rightarrow i\frac{\partial}{\partial x}$. The results of applying this operator are tabulated in Table 11.

It can readily be seen from the table that

$$\langle\varphi|(\mathbf{K}^2|\varphi\rangle) \neq (\langle\varphi|\mathbf{K})(\mathbf{K}|\varphi\rangle) \qquad \text{E 2}$$

in fact, they differ only in sign. The left hand side is unphysical since it gives a negative

kinetic energy. But if E 2 holds as an inequality, the operator **K** is not Hermitian. The problem here stems from an incorrect handling of the discontinuity in momentum across $x=0$. A solution is to enforce Hermiticity by always using the right hand side. [112] It is interesting to note that this problem would not have arisen in the momentum basis.[113] The reason is that in the momentum basis **K** becomes a number, not an operator, so the Hermiticity is guaranteed.

The correct result is

$$\langle \varphi | \mathbf{K}^2 | \varphi \rangle = -k \int_{-\infty}^{\infty} e^{-k|x|} \frac{\partial^2}{\partial x^2} e^{-k|x|} dx = k^2 . \qquad \text{E 3}$$

as used in the variational calculation of the exciton binding energy in Chapter 5. Implicit in E 3 is the idea that the derivative operator is such that momentum is Hermitian.

Appendix F

CONVERGENCE OF VARIATIONAL INTEGRALS

The Coulomb interaction term in the Hamiltonian scales as $1/r$ in interparticle separation. Thus it is singular at zero interparticle separation. In two and three dimensions, this singularity is lifted by the r and r^2 dependence of the Jacobian. In one dimension, as discussed physically in the introduction to Chapter 5, this singularity persists after integration. Mathematically, the Jacobian contains no term in r to cancel the singularity. The result is an infinite exciton binding energy, meaning in purely one dimension the electron and hole do not form a stable pair and are compelled to recombine.

The integrals in (52) and (57) are suspect since they involve integration over a singularity. Moreover, they relate to the exciton binding energy in quasi-1D. There are at least two ways to resolve this problem. First, there are physical arguments, and second rigorous bounds can be put on the integrals. Both approaches will be used below.

Physical Argument

The semiclassical physical argument is as follows: In 2D a stable excitonic state can be formed because centrifugal force can balance the Coulombic attraction. There is no reason why a finite 2D slab, say of circular cross-section, could not contain such an orbit and support such a bound state. Such a circle can be inscribed in the crescent, and so it too must support such a bound state, at least for some size of crescent. The binding energy must therefore be finite, and since the variational binding energy must be less than

the true binding energy the integral must converge. But this argument is perhaps circular, so the purely mathematical approach will be followed below.

Exponential Integral

Looking at the numerator of (52), and using the fact that the Mathieu functions are finite and may be normalized to be everywhere less than unity,

$$\int dA\phi^2\left[\mathbf{H}_0\left(\frac{2\Delta}{\lambda}\right)-Y_0\left(\frac{2\Delta}{\lambda}\right)\right]<\int dA\left[\mathbf{H}_0\left(\frac{2\Delta}{\lambda}\right)-Y_0\left(\frac{2\Delta}{\lambda}\right)\right]. \qquad \text{F 1}$$

Changing to a relative coordinate so that

$$\begin{aligned} x &= x_c - x_v \\ y &= y_c - y_v \\ X &= X_c - X_v \\ Y &= Y_c - Y_v \\ x &= \frac{\lambda}{2} r\cos\vartheta \\ y &= \frac{\lambda}{2} r\sin\vartheta \end{aligned} \qquad \text{F 2}$$

The integral may be bound by

$$<\int dX\int dY\int d\vartheta\int rdr\left[\mathbf{H}_0(r)-Y_0(r)\right]. \qquad \text{F 3}$$

where all the bounds on the integral are finite, and should be chosen to encompass the same region as above. Up to a (finite) factor this is

$$<\int_0^c rdr\mathbf{H}_0(r)-\int_0^c rdrY_0(r). \qquad \text{F 4}$$

where the bounds c is finite and has been chosen to include the entire area of integration. Both these integrals are finite, and can be obtained from tables.[114] (If the integrand changes sign on this region the integral can always be split up into regions of different sign

and since each part is finite, the magnitude can then be added to give a finite result.) Therefore the integral converges.

Gaussian Integral

A similar line of reasoning shows that the Gaussian integral converges. Since the range is finite,

$$\int dA \phi^2 e^{\Delta^2/\lambda^2} K_0\left(\frac{\Delta^2}{\lambda^2}\right) < C \int dA K_0\left(\frac{\Delta^2}{\lambda^2}\right). \qquad \text{F 5}$$

for C some finite factor. The same sort of relative coordinate transformation as used in (F 2) puts the following bound, up to a finite factor on the integral:

$$< \int_0^a r dr K_0\left(r^2\right) \propto \int_0^b du K_0(u). \qquad \text{F 6}$$

Here a and b are finite. The rightmost integral is finite can also be found in Prudnikov *et al.*[115] (To be completely rigorous, care must be taken that if the integrand changes sign on the circular region in F 6, it should be split up its magnitude added. However, this can always be done since the integral converges for any b, and since $K_0(x)$ changes sign a finite number of times on that interval.)

[1] J. Maddox, Nature **366** (1993) p297

[2] R. P. Feynman, Engineering and Science, California Institute of Technology (1960) February reprinted at http://nano.xerox.com/nanotech/feynman.html

[3] see for example G. M. Whitesides, Scientific American **273** (September 1995) p146

[4] L. Esaki and R. Tsu, IBM. J. Res. Develop. **14** (1970) p61

[5] reprinted in A. Y. Cho, Molecular Beam Epitaxy, American Institute of Physics Press, New York (1994) p365

[6] A. Y. Cho, Growth of Periodic Structures by the Molecular Beam Method, Appl. Phys. Lett. **19** (1971) p467

[7] S. M. Sze, Physics of Semiconductor Devices, John Wiley and Sons, New York (1991) p65

[8] K. Nakajima in GaInAsP Alloy Semiconductors, T.P. Pearsall, Ed., John Wiley and Sons, New York (1982) Chapter 2

[9] G. B. Stringfellow, Organometallic Vapor-Phase Epitaxy, Academic Press, New York (1989) Section 1.2

[10] G. H. Olson in K. Nakajima in GaInAsP Alloy Semiconductors, T.P. Pearsall, Ed., John Wiley and Sons, New York (1982) Chapter 1

[11] W. T. Tsang, The Growth of GaAs, AlGaAs, InP and InGaAs by Chemical Beam Epitaxy Using Group III and V Alkyls, J. Appl. Phys. (1986) p235

[12] For a historical review of the development of MBE see A. Y. Cho, Molecular Beam Epitaxy, American Institute of Physics Press, New York (1994)

[13] For an extensive review of MBE see M. A. Herman and H. Sitter, Molecular Beam Epitaxy, Springer-Verlag, New York (1989)

[14] F. J. Morris and H. Fukui, J. Vac. Sci. Technol. **11** (1974) p506

[15] M. A. Herman, H. Sitter, Molecular Beam Epitaxy, Springer-Verlag, New York (1989) p17

[16] Special issue: Selective Area Epitaxy, Non-Planar Epitaxy and Integration Technology, Semicond. Sci. Technol. **8** (1993) p959-1185

[17] W. M. Moreau, Semiconductor Lithography, Plenum Press, New York (1988) p1

[18] R. E. Williams, Gallium Arsenide Processing Techniques, Artech House, Inc., Dedham, Massachusetts, USA (1984) Chapters 5-8, p101

[19] for a discussion of the many advantages of postbaking, see R. E. Williams, Gallium Arsenide Processing Techniques, Artech House, Inc., Dedham, Massachusetts, USA (1984) p 139

[20] Yan Feng, Ph.D. Thesis, Department of Electrical Engineering, Imperial College of Science, Technology and Medicine, London, England (1992) Chapter 2.

[21] for a review see E. Kapon, Semiconductors and Semimetals **40** A. C. Gossard, ed., Academic Press, New York, (1994) p259

[22] A. Madhukar, K. C. Rajkumar and P. Chen, Appl. Phys. Lett. **62** (1993) p1547

[23] K. Eberl, P. Grambow, A. Lehmann, A. Kurtenbach, and K. v. Klitzing, D. Heitmann, M. Dilger and M. Hohenstein, Appl. Phys. Lett. **63** (1993) p1059

[24] S. Ando and T. Fukui, J. Crystal Growth **98** (1989) p646

[25] J. A. Lebens, C. S. Tsai, and K. J. Vahala, T. F. Kuech, Appl. Phys. Lett. 56 (1990) p2643

[26] Y. Arakawa, Y. Nagamune, M. Nishioka and S. Tsukamoto, Semicond. Sci. and Technol., 8, (1993) p1082

[27] H. Asai, S. Yamada, and Takashi Fukui, Appl. Phys. Lett., 51 (1987) p1518

[28] S. Koshiba, H. Noge, H. Akiyama, T. Inoshita, Y. Nakamura, A. Shimizu, Y. Nagamune, M. Tsuchiya, H. Kano and H.Sakaki, K. Wada, Appl. Phys. Lett. 64 (1994) p363

[29] P. Finnie, M. Buchanan, C. Lacelle, A. P. Roth, J. Crystal Growth 160 (1996) p220

[30] Y. D. Galeuchet, P. Roentgen, S. Nilsson and V. Graf, in Science and Engineering of One- and Zero-Dimensional Semiconductors, S.P. Beaumont and C. M. Sotomajor Torres, Plenum Press, New York (1990) p1

[31] Y. D. Galeuchet, P. Roentgen and V. Graf, J. Appl. Phys. 68 (1990) p560

[32] H. Sugiura, T. Nishida, R. Iga, T. Yamada and T. Tamamura, Journal of Crystal Growth, 121 (1992) p579

[33] A. P. Roth, P. Finnie, M. Buchanan, C. Lacelle (unpublished).

[34] See e.g. I. Ipatova and V. Mitin, Solid State Electronics, Addison Wesley Publishing Company, Inc., New York (1996) p198

[35] M. Kastner, Physics Today (January 1993) p24

[36] C. G. Smith, Rep. Prog. Phys. 59 (1996) p235

[37] Yia-Chung Chang, L. L. Chang, and L. Esaki, Appl. Phys. Lett. 47 (1985) p1324

[38] A.R. Goñi, L.N. Pfeiffer, K.W. West, A. Pinczuk, H.U. Baranger and H.L. Sturmer, Appl. Phys. Lett. 61 (1992) p1956

[39] P.M. Petroff in Semiconductors and Semimetals, 40 A. C. Gossard, ed., Academic Press, New York, (1994) p233

[40] D. Leonard, M. Krishnamurthy, C. M. Reaves, S. P. Deubeiars and P. M. Petroff, Appl. Phys. Lett. 63 (1993) p3203

[41] J.Y. Marzin, J. M. Gerard, A. Izrael, D. Barrier, G. Bastard, Appl. Phys. Lett. 73 (1994) p716

[42] S. Fafard, R. Leon, D. Leonard, J.L. Mertz and P. M. Petroff, Phys. Rev. B 50 (1994) p8086

[43] for a brief review see R. Bhat, Semicond. Sci. Techol. 8 (1993) p984

[44] H. Heinecke, A. Milde, B. Baur and R. Matz, Semicond. Sci. Technol. 8 (1993) p1023

[45] O. Kayser, Journal of Crystal Growth 107 (1991) p989

[46] J. M. Moisson, C. Guille, F. Houzay, F. Barthe, and M. Van Rompay, Phys. Rev. B 40 (1989) p6149

[47] N. Kobayashi, J.L. Benchimol, F. Alexandre and Y. Gao, Appl. Phys. Lett. 51 (1987) p1907

[48] D. A. Andrews and G.J. Davies. J. Appl. Phys. 67 (1990) p3187

[49] P. Chen, K.C. Rajkumar, and A. Mardakar, Appl. Phys. Lett., 58 (1991) p1771

[50] H. Sugiura, T. Nishida, R. Iga, T. Yamada, and T. Tamamura, Journal of Crystal Growth 121 (1992) p570

[51] H. Asai, J. Crystal Growth, 80 (1987) p425

[52] T. Yamada and Y. Horikoshi, Jpn. J. Appl. Phys. 33 (1994) L1027

[53] G. B. Stringfellow, Organometallic Vapor-Phase Epitaxy, Academic Press, New York (1989) p149

[54] for migration between (111)B and (100) in MBE see S. Koshiba, N. Nakamura, M. Tsuhchiya, H. Noge, H. Kano, Y. Nagamune, T. Noda and H. Sakaki, J. Appl. Phys. 76 (1994) p4138 and X.-Q. Shen and T. Nishinaga, Jpn. J. Appl. Phys. 32 (1993) L1119

[55] O. Kayser, R. Westphalen, B. Opitz and P. Balk, Journal of Crystal Growth 112 (1991) p111

[56] R. Matz, H. Heinecke, B. Baur, R. Primig, C. Crener, Journal of Crystal Growth 127 (1993) p230

[57] H. Heinecke, Journal of Crystal Growth 136 (1994) p18

[58] T. Sudersena Rao, C. Lacelle, R. Benzaquen, S.J. Rolfe, S. Charbonneau, P.D. Berger and A.P. Roth, T.Steiner and M.L.W. Thewalt, J. Appl. Phys. 76 (1994) p5300

[59] see e.g. A.-L. Barabasi and H.E. Stanley, Fractal Concepts in Surface Growth, Cambridge University Press (1995) p131

[60] G. B. Stringfellow, Organometallic Vapor-Phase Epitaxy, Academic Press, New York (1989) p18, Figure 2.2

[61] G. B. Stringfellow, Organometallic Vapor-Phase Epitaxy, Academic Press, New York (1989) p144

[62] see e.g. E. Kapon in Epitaxial Microstructures, Semiconductors and Semimetals 40, A. C. Gossard, ed., Academic Press, New York, (1994), Table I, p278 and references cited within.

[63] Gerald Bastard, Wave Mechanics Applied to Semiconductor Heterostructures, Les Editions de Physique, Les Ulis Cedex, France (1988) Table I, p34

[64] A. P. Roth, P. Levesque, R. W. G. Syme, D.J. Lockwood, G.C. Aers, T.S. Rao and C. Lacelle, J. Appl. Phys. 80 (1996) p4033

[65] NIH Image (public domain Macintosh software), developed by U.S. National Institutes of Health, available at http://rsb.info.nih.gov/NIH-image

[66] thanks to G. C. Aers and R. Williams (private communication) for explaining their diffusion calculations on which this is based

[67] for a more complete treatment of diffusion, see Mitsuru Ohtsuka and Seiichi Miyazawa, J. Appl. Phys. 64 (1988) p3522

[68] for a general review, see for example C. Weisbuch and B. Vinter, Quantum Semiconductor Structures, Academic Press, New York, (1991) or G. Bastard, Wave Mechanics Applied to Semiconductor Heterostructures, Les Editions de Physique (1988) Chapter 3

[69] P. Exner and P. Seba, J. Math. Phys. 30 (1989) p2574

[70] E. Kapon et al. Appl. Phys. Lett. 50 (1987) p327

[71] F. M. Peeters, in Science and Engineering of One- and Zero- Dimensional Semiconductors, S. P. Beaumont and C. M. Sotomajor Torres, Eds., Plenum Press, New York (1990)

[72] J. Goldstone and R. L. Jaffe, Phys. Rev. B 45 (1992) p14100

[73] Y. Nagaoka and M. Ikegami, Springer Series in Solid State Sciences 109 H. Fukuyama and T. Ando Eds., Springer-Verlag, Berlin (1992) p167

[74] Hua Wu, D.W.L. Sprung, and J. Martorell, Phys. Rev. B 45 (1992) p11960

[75] John P. Carini, J.T. Londergan, Kieran Mullen and D. F. Murdock, Phys. Rev B 46 (1992) p15538

[76] C J. Anthony and M. J. Kelly, Semicond. Sci. Technol. 9 (1994) p35

[77] F. M. Arscott and A. Darai, IMA Journal of Applied Mathematics 27 (1981) p33

[78] P. Moon and D. E. Spencer, Field Theory Handbook (Springer-Verlag, New York, 1971), Table 1.03

[79] N. W. McLachlan, Theory and Application of Mathieu Functions (Oxford University Press, New York, 1947), Chapter 2.

[80] IMSL Math/Library Special Functions (IMSL, Inc., Houston Texas, 1991), Chapter 12. This method is used by the commercial library to calculate the separation constant and the periodic Mathieu functions. Unfortunately the library does not yet calculate the aperiodic solutions or the solutions to the modified Mathieu equation.

[81] Tables Relating to Mathieu Functions, 2nd Ed. (US Government Printing Office, Washington, D. C., 1967) Chapter 3

[82] this figure is based on one from D. Halliday and R. Resnick, Fundamentals of Physics, John Wiley and Sons, Inc., New York (1986) p357

[83] T. Inoshita and H. Sakaki, J. Appl. Phys. 79 (1996) p269

[84] G. Bastard, Wave Mechanics Applied to Semiconductor Heterostructures, Les Editions de Physique (1988) p76

[85] L. Banyai, I. Galbraith, C. Ell, H. Haug, Phys. Rev. B 36 (1987) p6099

[86] see for example F. Bassani and G. Pastori Parravincini, Electronic States and Optical Transitions in Solids, The Science of Solid State 8, R. A. Ballinger, Ed., Pergamon Press (1975) p184

[87] G. Bastard, E. E. Mendez, L. L. Chang and L. Esaki, Phys. Rev. B 26 (1982) p1974

[88] Gang Li, Spiros V. Branis, and K.K. Bajaj, J. Appl. Phys 77 (1995) p1097

[89] S. Jaziri, G. Bastard and R. Bennaceur, Semicond. Sci. Technol 8 (1993) p670

[90] R. Shankar, Principles of Quantum Mechanics, 2nd Ed., Plenum Press, New York (1994) p 433

[91] Gerald Bastard, Wave Mechanics Applied to Semiconductor Heterostructures, Les Éditions de Physique, Les Ulis Cedex, France (1988) p130

[92] Gerald Bastard, Wave Mechanics Applied to Semiconductor Heterostructures, Les Éditions de Physique, Les Ulis Cedex, France (1988) p135

[93] see e.g. R. Shankar, Principles of Quantum Mechanics, 2nd Ed., Plenum Press, New York (1994) p 429

[94] I. S. Gradshteyn and I. M. Ryzhik, Table of Integrals, Series and Products, 4th ed., Academic Press (1965), p322, Eq. 3.387.7

[95] W. H. Press, S.A. Teukolsky, W. T. Vetterling, and B. P. Flannery Numerical Recipes in C, 2nd Ed., Cambridge University Press, New York (1992) p135

[96] J. Singh, Semiconductor Optoelectronics, McGraw-Hill, Inc. (1995) p537

[97] A. R. Pratt, R. L. Williams, C. E. Norman, M. R. Fahy, A. Marinopoulou and F. Chantenoud, Appl. Phys. Lett. 65 (1994) p1009

[98] H. Heinecke, A. Milde, B. Baur, and R. Matz, Semicond. Sci. Technol. 8 (1993) p1023

[99] see e.g. H. C. Liu in III-V Semiconductor Quantum Wells and Superlattices for Long Wavelength Infrared Detectors, Artech House, Boston (1994).

[100] Document 111858-P1, Rev A 4/90, Type 1259C/2259C Mass Flow Controllers, MKS Instruments, Inc., Andover, Massachusetts (1990)

[101] For a more thorough description of the preparation of GaAs wafers see R. Williams, Modern GaAs Processing Methods, 2nd ed., Artech House,London, (1990) p38

[102] G. Bastard, Wave Mechanics Applied to Semiconductor Heterostructures, Les Editions de Physique, 1988, Fig. 5, p76

[103] F. H. Pollak, Semiconductors and Semimetals 32 (1991) p26

[104] R. People and Sputz, Phys. Rev. B 41 (1990) p8437.

[105] N. W. McLachlan, Theory and Application of Mathieu Functions (Oxford University Press, New York, 1947), Chapter 2

[106] Tables Relating to Mathieu Functions, Applied Math Series 59 U.S. Department of Commerce, National Bureau of Standards, Washington, D. C. (1967) Introduction

[107] I. S. Gradshteyn and I. M. Ryzhik, Table of Integrals, Series and Products, 4th ed., Academic Press, New York (1965), p992

[108] W. H. Press, S.A. Teukolsky, W. T. Vetterling, and B. P. Flannery Numerical Recipes in C, 2nd Ed., Cambridge University Press, New York, 1992, p170

[109] W. H. Press, S.A. Teukolsky, W. T. Vetterling, and B. P. Flannery Numerical Recipes in C, 2nd Ed., Cambridge University Press, New York, 1992, p181

[110] LAPACK documentation is available at http://www.netlib.org/lapack

[111] G. Blanch in Handbook of Mathematical Functions, Applied Math Series 55 M. Abramowitz and I. E. Stegun, Eds. U.S. Department of Commerce, National Bureau of Standards, Washington, D. C. (1972) p721

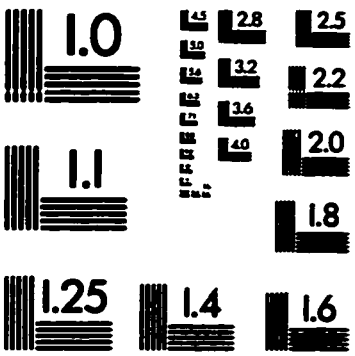
[112] B. Joos (private communication)

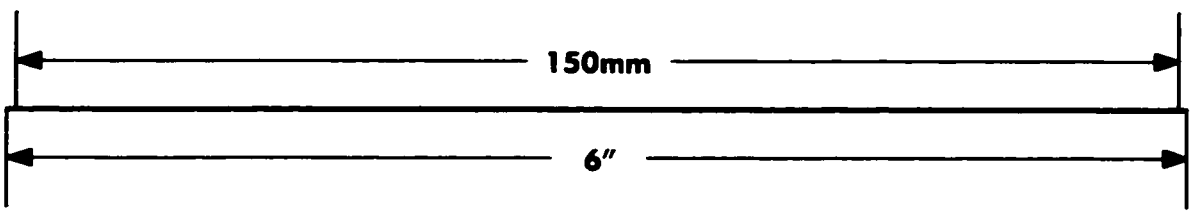
[113] R. Masut (private communication)

[114] A. P. Prudnikov, Yu. A. Brychkov, O. I. Marichev, Integrals and Series, Gordon and Breach Science Publishers, New York (1990) Volumes II p24 and III p44

[115] A. P. Prudnikov, Yu. A. Brychkov, O. I. Marichev, Integrals and Series, Gordon and Breach Science Publishers, New York (1990) Volume II p247

IMAGE EVALUATION
TEST TARGET (QA-3)





APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved